

PERSPECTIVES ON THE 'BILINGUAL ADVANTAGE': CHALLENGES AND OPPORTUNITIES

EDITED BY: Peter Bright and Roberto Filippi

PUBLISHED IN: Frontiers in Psychology and Frontiers in Communication





frontiers

Frontiers Copyright Statement

© Copyright 2007-2019 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88963-017-2

DOI 10.3389/978-2-88963-017-2

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

PERSPECTIVES ON THE 'BILINGUAL ADVANTAGE': CHALLENGES AND OPPORTUNITIES

Topic Editors:

Peter Bright, Anglia Ruskin University, United Kingdom

Roberto Filippi, UCL Institute of Education, United Kingdom

The claim that multilanguage acquisition drives advantages in 'executive function' is currently an issue of vigorous debate in academic literature. Critics argue that evidence for this advantage has been confounded by unsound or questionable methodological practices, with some investigators abandoning research in this area altogether, indicating either that there is no bilingual advantage or that it is impossible to capture and therefore rule out alternative explanations for group differences. Over the past decade, and against this backdrop, theory has developed from a relatively narrow focus on inhibitory control to incorporate theory of mind, rule-based learning, reactive and proactive control, visuo-spatial memory, and control of verbal interference in speech comprehension. Most recently, authors have claimed that the process of becoming bilingual may also impact on metacognitive abilities.

The fundamental issue is whether the limited capacity and goal-directed selectivity of our executive system can somehow be enhanced or otherwise profit from the continuous, intense competition associated with communicating in multilingual environments. However, although this issue has received much attention in academic literature, the question of which cognitive mechanisms are most influenced by the enhanced competition associated with multilingual contexts remains unresolved.

Therefore, rather than dismissing this important topic, we advocate a more systematic approach in which the effects of multilinguistic experience are assessed and interpreted across well-defined stages of cognitive development. We encourage a broad, developmentally informed approach to plotting the trajectory of interactions between multi-language learning and cognitive development, using a convergence of neuroimaging and behavioral methods, across the whole lifespan.

Moreover, we suggest that the current theoretical framing of the bilingual advantage is simplistic, and this issue may limit attempts to identify specific mechanisms most likely to be modulated by multilingual experience. For example, there is a tendency in academic literature to treat 'executive function' as an essentially unitary fronto-parietal system recruited in response to all manner of cognitive demand, yet performance across so called 'executive function' tasks is highly variable and intercorrelations are sometimes low. It may be the case that some 'higher level' mechanisms of 'executive function' remain relatively unaffected, while others are more sensitive to multilingual experience – and that there may be disadvantages as well as advantages, which themselves may be sensitive to factors such as age. In our view, there is an urgent need to take a more fine-grained approach to this issue, so that the strength and direction of changes in diverse cognitive abilities associated with multilanguage acquisition can be better understood.

This book compiles work from psychologists and neuroscientists who actively research whether, how, and the extent to which multilanguage acquisition promotes enhanced cognition or protects against age-related cognitive or neurological deterioration. We hope this collection encourages future efforts to drive theoretical progress well beyond the highly simplistic issue of whether the bilingual cognitive advantage is real or spurious.

Citation: Bright, P., Filippi, R., eds. (2019). Perspectives on the 'Bilingual Advantage': Challenges and Opportunities. Lausanne: Frontiers Media.
doi: 10.3389/978-2-88963-017-2

Table of Contents

- 06 Editorial: Perspectives on the “Bilingual Advantage”: Challenges and Opportunities**
Peter Bright and Roberto Filippi
- 09 Executive Functions and Language: Their Differential Influence on Mono- vs. Multilingual Spelling in Primary School**
Sophia Czapka, Annegret Klassert and Julia Festman
- 27 Sequential Congruency Effects in Monolingual and Bilingual Adults: A Failure to Replicate Grundy et al. (2017)**
Samantha F. Goldsmith and J. Bruce Morton
- 33 Comprehending Non-literal Language: Effects of Aging and Bilingualism**
Shamala Sundaray, Theodoros Marinis and Arpita Bose
- 48 Exploring Different Types of Inhibition During Bilingual Language Production**
Maria Borragan, Clara D. Martin, Angela de Bruin and Jon Andoni Duñabeitia
- 59 Symbiosis, Parasitism and Bilingual Cognitive Control: A Neuroemergentist Perspective**
Arturo E. Hernandez, Hannah L. Claussenius-Kalman, Juliana Ronderos and Kelly A. Vaughn
- 66 Multilingual Language Control and Executive Function: A Replication Study**
Gregory J. Poarch
- 77 Does Extreme Language Control Training Improve Cognitive Control? A Comparison of Professional Interpreters, L2 Teachers and Monolinguals**
Lize Van der Linden, Eowyn Van de Putte, Evy Woumans, Wouter Duyck and Arnaud Szmalec
- 97 Bidialectalism and Bilingualism: Exploring the Role of Language Similarity as a Link Between Linguistic Ability and Executive Control**
Jessica Oschwald, Alisa Schättin, Claudia C. von Bastian and Alessandra S. Souza
- 119 Impact of Language Experience on Attention to Faces in Infancy: Evidence From Unimodal and Bimodal Bilingual Infants**
Evelyne Mercure, Isabel Quiroz, Laura Goldberg, Harriet Bowden-Howl, Kimberley Coulson, Teodora Gliga, Roberto Filippi, Peter Bright, Mark H. Johnson and Mairéad MacSweeney
- 129 Measuring the Timing of the Bilingual Advantage**
Sara Incera
- 138 The Importance of Socioeconomic Status as a Modulator of the Bilingual Advantage in Cognitive Ability**
Kamila Naeem, Roberto Filippi, Eva Periche-Tomas, Andriani Papageorgiou and Peter Bright
- 147 The Role of Cognitive Development and Strategic Task Tendencies in the Bilingual Advantage Controversy**
Esli Struys, Wouter Duyck and Evy Woumans

- 158** *Exploiting Language Variation to Better Understand the Cognitive Consequences of Bilingualism*
Andrea A. Takahesu Tabori, Emily N. Mech and Natsuki Atagi
- 165** *Working Memory With Emotional Distraction in Monolingual and Bilingual Children*
Monika Janus and Ellen Bialystok
- 177** *Do Bilinguals Have an Advantage in Theory of Mind? A Meta-Analysis*
Scott R. Schroeder
- 185** *Bilingual Advantages in Inhibition or Selective Attention: More Challenges*
Kenneth R. Paap, Regina Anders-Jefferson, Lauren Mason, Katerinne Alvarado and Brandon Zimiga
- 195** *Recognizing the Effects of Language Mode on the Cognitive Advantages of Bilingualism*
Ziying Yu and John W. Schwieter



Editorial: Perspectives on the “Bilingual Advantage”: Challenges and Opportunities

Peter Bright^{1*} and Roberto Filippi^{2*}

¹ Department of Psychology, Anglia Ruskin University, Cambridge, United Kingdom, ² Institute of Education, University College London, London, United Kingdom

Keywords: multilingualism, cognitive control, executive functions, bilingual advantage, bilingualism

Editorial on the Research Topic

Perspectives on the “Bilingual Advantage”: Challenges and Opportunities

When we ask our students or members of the general public the question *Is being bilingual/multilingual an advantage?* The answer, invariably, is yes. The reasons provided are intuitively sensible and leave little room for disagreement. Multilingual speakers can communicate with different people, they understand different cultures, they have more job opportunities, they can travel the world with more confidence, and so forth.

However, when we formulate the question in a different way, *Is being bilingual/multilingual an advantage for cognitive development?* Answers are not as straightforward. Some are concerned that second language learning may delay language acquisition in early stages of life, others think that children should focus more on one language to avoid mental confusion. In some cases, and this is probably the most disturbing situation, education professionals advise parents from different cultures to raise their children as monolinguals, advocating that this is more likely to lead to good academic achievement (e.g., Festman et al., 2017). This opinion almost certainly derives in part from early evidence (e.g., Saer, 1923) for a mental delay in bilingual children compared to monolingual peers on a range of tests measuring intelligence quotient (IQ).

The more recent work of pioneer scientists (e.g., Peal and Lambert, 1962; Bialystok and Ellen, 1991), incorporating more rigorous and systematic paradigms and procedures, has underpinned a now widely-held consensus among researchers in the field, that multilanguage learning is not detrimental for cognitive development. Nevertheless, while few—if any—scientists now hold the position that multilanguage acquisition underpins a cognitive *disadvantage*, there is ongoing vigorous debate about whether there are distinct cognitive *advantages* associated with multilingualism that cannot be explained by other candidate explanatory variables. Understanding the cognitive sequelae of bilingualism presents many hurdles that will require continued intense effort.

Collectively, the 17 articles contained herein, reflect the current state of the field, with well-defended positions on opposing sides of current debate. Altogether, 44 leading scientists in the field of multilingualism have contributed with commentaries, meta-analyses, methodological advice, and empirical research. We are most grateful to them, to the independent reviewers and to Frontiers for providing the means to make this happen.

Yu and Schwieter begin this collection with a conceptual analysis of the significance of language mode in bilingual cognition, that is, the degree of co-activation of the two languages at any one time (Grosjean, 1998, 2010). They encourage more robust and systematic consideration of language mode in future studies due to its potential modulatory effect on language activation and also, therefore, on the likely cognitive benefits associated with bilingualism. In a short review, these

OPEN ACCESS

Edited by:

Alain Morin,
Mount Royal University, Canada

Reviewed by:

Kenneth R. Paap,
San Francisco State University,
United States
Ellen Bialystok,
York University, Canada

*Correspondence:

Peter Bright
peter.bright@anglia.ac.uk
Roberto Filippi
r.filippi@ucl.ac.uk

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 26 April 2019

Accepted: 23 May 2019

Published: 06 June 2019

Citation:

Bright P and Filippi R (2019) Editorial:
Perspectives on the “Bilingual
Advantage”: Challenges and
Opportunities.
Front. Psychol. 10:1346.
doi: 10.3389/fpsyg.2019.01346

authors provide a convincing case that the failure to assess and control language mode may, at least in part, explain the contradictory findings reported in the literature. The controversy about whether, and the extent to which, bilingualism confers cognitive benefits is also tackled by Takahesu Tabori et al. in their timely methodological review which, in particular, addresses sample characteristics. They argue that most published studies provide insufficient information on language experience/background, social context of language use and decry the paucity of longitudinal designs which, they argue, offer a greater degree of experimental control. They encourage work toward more widely agreed criteria for terms such as “native language,” “first language,” “second language,” etc., and argue against over-simplification, most obvious in the long-standing dichotomised categorization of monolingual vs. bilingual and bilingual advantage vs. no advantage. Several of the studies in this collection demonstrate a shift to more nuanced and precise conceptualization of bilingual cognition, and this, of course, is to be welcomed and encouraged.

In her excellent review, Incera, considers timing of processing in the bilingual mind as a tool for understanding how bilingual and monolingual cognition may diverge. She offers a range of recommendations for future attempts at resolving conflicting findings, and researchers would do well to act on them. Of these, inclusion of time-sensitive measures and baseline conditions, consideration of bilingualism as a continuous variable and a focus on group by condition interactions over main effects of bilingualism are, in our view, most likely to lead to sustained theoretical advances in this area. Hernandez et al. outline a neuroemergentist approach which, they argue, may also offer a more ambitious and plausible framing of the complex ways in which bilingualism may interact with development of domain-general cognitive control.

Schroeder tackles the possibility that bilingual children have an advantage in theory of mind, presenting a meta-analysis of 16 studies. Small to medium positive effects of bilingualism were observed (contingent on the analysis), indicating that second language learning may have modest implications for the development of social competence, although well-grounded explanations for this association are currently lacking.

Five studies address the impact of multi-language experience on cognitive control in infants or children. Mercure et al. explored attention to still faces in monolingual infants, unimodal bilingual infants (i.e., learning two spoken languages) and bimodal bilingual infants of Deaf mothers (learning British Sign Language and spoken English). Equivalent attention capture and maintenance by face stimuli was observed in monolinguals and bimodal bilinguals, but unimodal bilinguals showed comparatively faster attention capture and maintenance, raising implications of multilanguage learning for social communication during infancy. Poarch provides a replication study with findings partly consistent with the central claim of the bilingual advantage theory, that controlling multiple languages in daily life confers genuine benefits in domain-general cognitive control. Specifically, equivalent performance among monolingual and bilingual children was observed on the Simon task, but the bilinguals demonstrated a significant advantage on

the flanker task, indicating that these tasks may recruit partly distinct mechanisms of cognitive control that are differentially sensitive to language environment and may also follow different developmental trajectories. Struys et al. also employed the Simon and flanker tasks in a comparison of performance among younger and older monolingual and bilingual Dutch-French children. They report equivalent performance across language groups but, crucially, there was marked variation in the actual strategies employed to resolve conflict in the tasks. This finding is consistent with recent (currently unpublished) work from our lab which indicates significant differences in the neural networks recruited among bilingual and monolingual participants when resolving conflict despite the absence of any group effects at the behavioral level.

Janus and Bialystok consider the reported association between executive function and emotion regulation, arguing that bilingual advantages in executive control may, intuitively, also underpin performance benefits in emotional contexts. However, in their study of emotional face N-back task performance in monolingual and bilingual children, there were no group differences in the overall effect of emotional valence on reaction time (despite better accuracy in bilinguals). Czapka et al. present a novel and intriguing study of real word and non-word spelling in monolingual and bilingual third grade (~9 year-old) primary school children in Germany, providing compelling evidence that monolinguals at this age are better able to deploy higher level cognitive control during spelling, most likely due to superior knowledge of the German language. For bilinguals, German lexicon size was a better predictor of spelling ability than executive function. These findings reinforce the importance of adopting a fine-grained, developmentally informed approach to charting interactions between multi-language learning and cognitive development, without which we are unlikely to resolve the contradictory claims and entrenched positions so prevalent in the recent literature.

Seven studies examine bilingual processing in adults, each of which focuses on a key issue in current debate. Naeem et al. address the potential importance of an alternative explanatory variable: socioeconomic status (SES). Employing demonstrably low and high SES monolingual and bilingual participants, these authors found evidence (from Simon task performance) that bilingualism may promote a speed of processing advantage, but only in those with low SES. Furthermore, there was no evidence for a bilingual advantage in executive planning ability (based on Tower of London performance), with monolinguals showing a disproportionate advantage. Van der Linden et al. explore interference suppression, response inhibition, and short-term memory performance in professional simultaneous interpreters. To the extent that bilingual cognitive advantages are associated with the requirement to manage and control simultaneously active languages in daily life, the authors argue that a comparison of such highly skilled bilinguals against monolinguals should increase the likelihood of detecting a bilingual advantage, if it exists. In fact, the two groups performed similarly on all measures (flanker, Simon, and digit span tasks), a finding reinforced in a second experiment which incorporated an additional group of second language

teachers. Despite anecdotal evidence for an STM advantage over monolinguals among interpreters, this evidence is clearly difficult to reconcile with bilingual advantage theory. In their study on the effect of language similarity on the association between linguistic performance and executive function, Oschwald et al. found very limited evidence for benefits in executive function associated with the increased demands of managing more dissimilar languages. These results, therefore, also offer evidence against the claim that managing cross-language interference promotes or enhances executive function. Evidence presented by Borragan et al. provides a possible explanation for lack of transfer from control of language interference to non-verbal executive function. These authors examine performance in highly proficient but unbalanced bilinguals on a multilingual rapid picture naming task incorporating multiple inhibitory demands. Findings are most consistent with the existence of functionally independent inhibitory mechanisms associated with language processing which may not be recruited in non-verbal tasks.

Further evidence against the existence of a genuine bilingual advantage, either in attentional control or response inhibition is presented by Paap et al. In this study, no effects attributable to bilingualism were observed on the tasks whether (i) participants were separated into monolingual or bilingual groups or (ii) degree of bilingualism was treated as a continuous variable, and Bayes factor analyses robustly supported the null hypothesis. The study by Goldsmith and Morton tests recent evidence by Grundy et al. (2017) that bilingual adults show smaller sequential congruency effects than monolingual adults, perhaps consistent with a bilingual efficiency advantage in the disengagement of attention from no longer relevant task stimuli. This new study, offered as a replication, showed statistically equivalent performance in both groups. However, Grundy and Bialystok have published a reply in *Frontiers* (available here), outlining that the study is not a direct replication but differs in several

ways. Perhaps, most importantly, they point out that Goldsmith and Morton employ long rather than short response-to-stimulus intervals, and it is at short intervals that language group differences in the disengagement of attention can most readily be observed.

The possibility that bilingualism may offer protection against age-related cognitive deterioration and/or neural degeneration is an important issue in the literature. Rather than addressing vocabulary, syntax, or comprehension, Sundaray et al. take the novel approach of addressing non-literal language (pragmatic inference making) in young and older monolingual and bilingual participants. With the exception of conventional metaphors (for which an age-related deficit was observed only in monolinguals) no differences between language groups in processing pragmatic inferences were observed. Thus, the evidence here suggests a possible protective effect of bilingualism in comprehension of non-literal language, restricted to conventional metaphors.

There are many challenges in this line of research, but when there is challenge there should also be opportunity to advance knowledge. In collecting these articles within a single volume, we hope readers will take the opportunity to digest the full range of empirically supported inferences, and further develop a well-informed understanding of how (and the extent to which) the process of acquiring a second language confers domain general cognitive benefits.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

This work was supported by the Leverhulme Trust UK [RPG-2015-024]. A special thought goes to Prof. Annette Karmiloff-Smith who inspired our work.

REFERENCES

- Bialystok, E., and Ellen, B. (eds.). (1991). *Language Processing in Bilingual Children*. Cambridge: Cambridge University Press.
- Festman, J., Poarch, G. J., and Dewaele, J. M. (2017). *Raising Multilingual Children*. Bristol, UK: Multilingual Matters.
- Grosjean, F. (1998). Studying bilinguals: methodological and conceptual issues. *Bilingual. Lang. Cogn.* 1, 131–149.
- Grosjean, F. (2010). *Bilingual*. Cambridge: Harvard university press.
- Grundy, J. G., Chung-Fat-Yim, A., Friesen, D. C., Mak, L., and Bialystok, E. (2017). Sequential congruency effects reveal differences in disengagement of attention for monolingual and bilingual young adults. *Cognition* 163, 42–55. doi: 10.1016/j.cognition.2017.02.010
- Peal, E., and Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychol. Monogr.* 76:1. doi: 10.1037/h0093840

Saer, D. J. (1923). The effect of bilingualism on intelligence. *Br. J. Psychol. Gen. Sec.* 14, 25–38.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Bright and Filippi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Executive Functions and Language: Their Differential Influence on Mono- vs. Multilingual Spelling in Primary School

Sophia Czapka^{1,2*}, Annegret Klassert^{1,3} and Julia Festman^{1,4}

¹ Research Group: Diversity and Inclusion, Human Sciences Faculty, University of Potsdam, Potsdam, Germany,

² Leibniz-Center General Linguistics, Berlin, Germany, ³ University of Applied Sciences Clara Hoffbauer, Potsdam, Germany,

⁴ Department of Primary and Secondary Education, Pedagogical University Tyrol, Innsbruck, Austria

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Vicky Chondrogianni,
The University of Edinburgh,
United Kingdom
Maurits W. Van Der Molen,
University of Amsterdam, Netherlands

*Correspondence:

Sophia Czapka
sophia.czapka@uni-potsdam.de

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 23 July 2018

Accepted: 14 January 2019

Published: 06 February 2019

Citation:

Czapka S, Klassert A and
Festman J (2019) Executive Functions
and Language: Their Differential
Influence on Mono- vs. Multilingual
Spelling in Primary School.
Front. Psychol. 10:97.
doi: 10.3389/fpsyg.2019.00097

We aimed at unveiling the role of executive functions (EFs) and language-related skills in spelling for mono- versus multilingual primary school children. We focused on EF and language-related skills, in particular lexicon size and phonological awareness (PA), because these factors were found to predict spelling in studies predominantly conducted with monolinguals, and because multilingualism can modulate these factors. There is evidence for (a) a *bilingual advantage* in EF due to constant high cognitive demands through language control, (b) a smaller mental lexicon in German and (c) possibly better PA. Multilinguals in Germany show on average poorer German language proficiency, what can influence performance on language-based tasks negatively. Thus, we included two spelling tasks to tease apart spelling based on lexical knowledge (i.e., word spelling) from spelling based on non-lexical strategies (i.e., non-word spelling). Our sample consisted of heterogeneous third graders from Germany: 69 monolinguals (age: $M = 108$ months) and 57 multilinguals (age: $M = 111$ months). On less language-dependent tasks (e.g., non-word spelling, PA, intelligence, short-term memory (STM) and three EF tasks testing switching, inhibition, and working memory) performance of both groups did not differ significantly. However, multilinguals performed significantly more poorly on tasks measuring German lexicon size and word spelling than monolinguals. Regression analyses revealed that for multilinguals, inhibition was related to spelling, whereas switching was the only EF component to influence word spelling in monolinguals and non-word spelling performance in both groups. By adding lexicon size and other language-related factors to the regression models, the influence of switching was reduced to insignificant effects, but inhibition remained significant for multilinguals. Language-related skills best predicted spelling and both language groups shared those variables: PA for word spelling, and STM for non-word spelling. Additionally, multilinguals' word spelling performance was also predicted by their German lexicon size, and non-word spelling performance by PA. This study offers an in-depth look at spelling acquisition at a certain point of literacy development. Mono-

and multilinguals have the predominant factors for spelling in common, but probably due to superior language knowledge, monolinguals were already able to make use of EF during spelling. For multilinguals, German lexicon size was more important for spelling than EF. For multilinguals' spelling these functions might come into play only at a later stage.

Keywords: bilingualism, spelling, literacy acquisition, executive functions, lexicon size, primary school

INTRODUCTION

Spelling *per se* is a highly crucial skill, because “[...] struggling with spelling production may result in students being demotivated, running out of time, having less time for planning or writing a shorter text” (Rønneberg and Torrance, 2017, p. 2). Hence, spelling influences the quality of a text, and often has an impact on the reader's judgment of the writer's competences. Reading texts full of spelling mistakes makes the comprehension of the content taxing and tiring, as it interrupts the perception of content. Thus, spelling is important for demonstrating subject-related competence, especially in the school context. Additionally, it has significant consequences, since the ability to spell correctly is also known to be a crucial criterion whether teachers recommend pupils for junior high school in the very selective German school system (Roos and Schöler, 2009).

Known predictors for literacy in general are specific language [such as phonological awareness (PA) and size of the mental lexicon] and cognitive skills [in particular executive functions (EFs)], but their role in spelling is not clear-cut. It also remains unknown if these predictors play the same role for multilinguals. Therefore, it is important to investigate not only the underlying internal factors influencing monolingual children's spelling performance, as multilingual children make up a considerable part of the student body in Germany (e.g., 38% of primary school children in Berlin; Leerhoff et al., 2013) and are considered at risk for school failure (Zöllner et al., 2006; Ditton and Krüsken, 2009; OECD, 2014). A main contributor to this risk is the on average lower proficiency in German, the language of schooling (Niklas et al., 2011). This impacts negatively on their reading performance. From primary school throughout the children's school career it is on average lower compared to their monolingual peers' (Müller and Stanat, 2006; Marx and Stanat, 2012). For spelling, findings showed multilinguals' similar (Roos and Schöler, 2009; Chudaske, 2012) or poorer performance (Schründer-Lenzen and Merckens, 2006) compared to their monolingual peers.

In sum, the influence of language-related and cognitive predictors of spelling, in particular for multilingual children, are largely unknown and findings for multilinguals' spelling competences are mixed. As research on the acquisition of spelling is considered to be rather anglo-centric (Wimmer and Landerl, 1997) and the English orthography is less consistent than German, it is called for an investigation of these different language-related and cognitive predictors for spelling in German.

Spelling

Contemporary accounts of writing include both linguistic and cognitive processes which are functionally integrated when writing (Abbott and Berninger, 1993; Berninger and Amtmann, 2003). These language skills comprise for example PA (i.e., the ability to perceive the phonological structure of words and manipulate elements of spoken language), as well as alphabetical, morphological, and lexical knowledge (Lubin et al., 2016). Being able to produce the correct, orthographic spelling of words requires a number of linguistic processes to be executed (e.g., phonological and morphological analysis of the word, translation into graphemes; Alamargot and Chanquoy, 2001; Treiman, 2017) and orchestrated by cognitive processes prior to graphomotor execution (i.e., the writing process itself).

Beginning writers' spelling in alphabetic orthographies like German is phonemic (see Ziegler and Goswami, 2006) and thus relies on phoneme-grapheme conversion. This is one of two different procedures to spell as suggested by the *dual-route account of orthographic retrieval* (e.g., Barry, 1994). Spelling via this *non-lexical route* relies on the direct translation of phonemes to their corresponding graphemes—so called phoneme-grapheme conversion (e.g., ‘enough’ [ɪˈnʌf] could result in the phonemic spelling ‘inaf’). With increasing training and through experience with printed language (Ehri, 2005, 2014), orthographic learning takes place, meaning the phonology and orthography of a word become connected, and an orthographic representation is stored in memory. These orthographic word forms build a separate orthographic lexicon (that is independent of the phonological forms stored in memory; Berninger et al., 2006), what allows direct retrieval of word forms from that lexicon. This route is the *lexical route* of spelling (the second procedure of the *dual-route account*). The direct access makes spelling more efficient (meaning correct and fast) (Berninger et al., 2006). However, this route can only process familiar words that already have lexical representations. Non-words and unfamiliar words are spelled via the non-lexical route, which is costlier than accessing whole word forms, because phoneme-grapheme conversion and orthographic rules need to be aligned. In conclusion, an extensive orthographic lexicon is essential for efficient word spelling (Roos and Schöler, 2009).

The *componential model of writing* (Schoonen et al., 2002; Harrison et al., 2016) broadens the complexity of the writing skill and includes also more cognitive components. It distinguishes between lower-level writing skills, comprising handwriting, punctuation and spelling, and higher-level writing skills, i.e., planning, formulating/composition, revising (McCutchen, 1996). Lower-level writing skills are acquired first in writing acquisition

and initially lack an automatized execution. Hence, they require conscious control through EF (see section “EF and Spelling”) and use up cognitive resources (Bourdin and Fayol, 1994; Grabowski, 2010). Cognitive resources comprise many mental processes (e.g., attention, memory, motor control and EF; Franconeri et al., 2013) and are limited in their capacity (following the *capacity theory of writing* by McCutchen, 1996, 2011). Hence, children who are still struggling with graphomotor execution have problems in spelling an entire word correctly, since the handwriting process itself takes up too many cognitive resources (Pontart et al., 2013). Similarly, if spelling draws on children’s attention, it draws processing resources away from higher-level processes and few resources are left for high-quality text composition. Therefore, children need to train spelling to reach a more automatic execution of these lower-level writing skills (for an overview, see Gerth et al., 2016). When writers become more advanced (around the age of 14), the automatization of these lower-level processes frees cognitive resources for higher-level processes such as sentence- and text-level processes (McCutchen, 1996). Even for advanced writers, writing draws on cognitive resources, because different lower-level and higher-level processes need to be coordinated concurrently (Alamargot and Chanquoy, 2001).

In sum, spelling is a developing skill especially in primary school children, as spelling is taught mainly in the school context. Spelling draws on a number of different cognitive and language-related processes. In the following sections, we will first describe the links between the three EF components proposed by Miyake and Friedman (2012) (i.e., switching, inhibition and updating) and spelling, before we highlight the impact of bi-/multilingualism on EF. Finally, we will portray the links between the language-related skills and spelling in more detail, with a specific focus on the role of lexicon size and PA for multilinguals.

EF and Spelling

Executive functions are a family of cognitive control mechanisms that regulate thought and action. These effortful, mental processes can be divided into three core functions: first, switching (or shifting) describes the ability to switch flexibly between mental sets or tasks. Secondly, inhibition is the ability to suppress dominant or irrelevant information or behavior to maintain a task goal. Thirdly, working memory (WM or updating) is the ability to manipulate mentally stored information (Miyake and Friedman, 2012; Diamond, 2013).

Executive functions are an important prerequisite for school success: they have been shown to predict school readiness and school performance (St Clair-Thompson and Gathercole, 2006; Diamond, 2013; Zorza et al., 2017), reading and mathematics (Best et al., 2011) and also to affect writing (e.g., Kellogg, 1996; Monette et al., 2011). Despite the clearly crucial contribution of EF to literacy development in general, there is still no consistent picture in regard to the impact of EF on spelling as one important literacy subprocess. To the best of our knowledge, there are only very few studies that investigated the direct influence of concrete EF components (switching, inhibition and WM) on spelling. In the following sections, we will give a short overview of their findings and—due to the very low number of specific studies

on spelling—we occasionally extend the review to literacy more generally.

Influence of Switching on Spelling

During spelling different processes need to be coordinated (as described above) and the writer needs to switch effectively between them to write fluently and correctly. For example, one needs to translate phonemes into graphemes (non-lexical route) or retrieve the spelling of the word from orthographic memory (lexical route), apply orthographic rules, compare phonemic and lexical word forms, and finally plan and execute graphomotor processes (Lubin et al., 2016). Third grade students are at a developmental stage of transition from phonemic spelling to extending their orthographic lexicon allowing the lexical route to be used more frequently (von Suchodoletz et al., 2017). This requires children to switch between these coexisting routes (non-lexical and lexical route) and even include, e.g., morphological strategies to support their spelling activities. Good switching abilities allow for flexible application of these different options to changing task demands during spelling.

Two studies confirmed a direct connection between switching and spelling. Lubin et al. (2016) investigated how switching, inhibition and WM influence spelling in primary school children. Their study was based on a group of native French-speaking fourth graders in France. For the switching task, children had to switch between counting up- and downwards depending on a cue. They found that switching was the only EF component explaining variance in their French spelling task. Von Suchodoletz (2017) conducted a cross-sectional study on attention shifting, assessed with a card sorting task, and spelling in German first, third and eighth graders. They found that better switching abilities were related to superior general spelling skills, in particular in third graders.

In two other studies on literacy in English, switching was found to predict reading, written expression and spelling of English at grade one to three (Altemeier et al., 2008) and also writing in the subsequent 6 years (Berninger et al., 2016). Switching in these studies was measured with a task called *rapid automatized switching* that combines a rapid naming task with increased EF demands through switching categories (e.g., naming alternately numbers and letters). However, because *rapid automatized naming* (RAN) on its own predicts reading and possibly writing (see below), we believe that it is difficult to tease apart the influence of RAN and switching when using rapid automatized switching as predictor.

Influence of Inhibition on Spelling

In our study we refer to inhibition as the ability to control interference at the level of perception; it is an involuntary and automatic reaction to or away from a stimulus (Diamond, 2013). According to Diamond (2013), this definition contrasts with other forms of inhibition like selective attention, which is a voluntary and active focus of one’s attention to or away from a stimulus, or self-control, i.e., the control of one’s behavior and emotions to resist temptation.

Writing involves many competitive processes that require inhibition of an irrelevant or incorrect choice to select the

appropriate alternative. In spelling, inhibition should be involved in choosing between lexical and non-lexical spelling routes, or in the selection of neighborhood word competitors which is suggested to be easier the more precise the orthographic knowledge is (Perfetti, 2017).

There is some evidence that inhibition is involved in reading-related skills (Blair and Razza, 2007) and literacy in general (assessed with general scores for reading and spelling; St Clair-Thompson and Gathercole, 2006; Allan et al., 2014). However, these general findings with respect to literacy do not allow us to draw conclusions on spelling. Concerning spelling more specifically, there seems to be hardly any evidence for an influence of inhibition on spelling: according to Altemeier et al. (2008), who tested English-speaking children longitudinally in the first 4 years of school, inhibition (assessed with a Color-Word interference test) additionally to rapid automatized switching (see section “Influence of Switching on Spelling”) predicted reading and writing, but not spelling. Also, Lubin et al. (2016) reported no impact of inhibition (measured with an Opposite World task) on spelling in French monolingual children. However, Lubin et al. (2016) discussed the possibility that in more speeded conditions or under conditions in which automatic processes need to be inhibited, inhibition abilities might become relevant.

Influence of WM on Spelling

WM is the third EF component identified by Miyake et al. (2000) and a complex structure itself. In Baddeley’s model of WM, the ability to manipulate mentally stored information (what we refer to as WM) is part of the so-called central executive (Repovš and Baddeley, 2006; Baddeley, 2010). Baddeley’s model additionally comprises separate functions for mental short-term storage: one for visual and spatial information, the visuo-spatial sketchpad, and one for verbal material, that is the phonological loop or phonological short-term memory (STM).

In writing research, WM is probably the most investigated EF component. The role of WM in proficient writing is described in Kellogg’s model (Kellogg, 1996), in which Baddeley’s central executive affects all sub-processes of writing (Alamargot and Chanquoy, 2001). WM plays a role in particular for higher-level writing skills, i.e., processes involved in text composition (e.g., planning, formulating, revising). WM is necessary for “active maintenance of multiple ideas, the retrieval of grammatical rules from long-term memory, and the recursive self-monitoring that is required during the act of writing” (Hooper et al., 2002). WM is also involved in spelling to dictation in the following way: while words and decoded sounds are stored temporarily in STM, processing of this information takes place in WM (orthographic information is retrieved from long-term memory, morphological rules are applied or phonemes translated to corresponding graphemes) and concurrently the writer needs to continue listening and updates the STM content, what is the task of WM (Strattman and Hodson, 2005).

Experimental support for the role of WM in spelling acquisition, however, is weak. Research findings suggest that WM and spelling are not connected in primary school children (see Lubin et al., 2016). Swanson and Berninger (1996) reported that in children at the age of 10 to 12, WM was rather connected

to higher-level and STM to lower-level writing skills. Also, for older children, Vanderberg and Swanson (2007) found that in 15-year-olds, WM (the authors refer to the central executive) predicted all levels of writing (planning, translating, revision), except spelling. Berninger et al. (2010) used a global measure of WM that combined a non-word repetition task for STM and backward digit span for WM (testing the central executive). They found that WM influenced spelling from second to sixth grade, but this finding does not allow generalization for specific components of EF or WM, since STM and WM represent distinct cognitive functions. To summarize, the influence of WM on spelling seems to be crucial for higher-level tasks (Swanson and Berninger, 1996), which play only a very minor role for primary school children, since they are mainly acquiring lower-level writing skills.

Influence of Bi- and Multilingualism on EF

As we study predictors for mono- versus multilinguals’ performance, we need to take into account that multilingualism can affect EF positively (Adesope et al., 2010; Bialystok, 2015; Hilchey et al., 2015; Zhang, 2018). The reason for the so-called *bilingual advantage* is that a bilingual’s languages are constantly activated, what puts particular demands on cognitive control to activate the language currently in use and inhibit the non-target language (Green, 1998; Kroll et al., 2012; Spalek et al., 2014). The constant load on the EF system has been suggested to constitute a training and to lead to improved EF (Bialystok, 2015). Strongest evidence for a bilingual advantage has been found for inhibition tasks (Bialystok and Viswanathan, 2009; Poarch and van Hell, 2012; Poarch, 2018). For WM, there is evidence for a positive influence of bilingualism (Luo et al., 2013; Blom et al., 2014), but also against (Namazi and Thordardottir, 2010; Engel de Abreu, 2011). Finally, superior switching performance tested with card sorting tasks could be related to bilingualism (Bialystok and Martin, 2004; Wiseheart et al., 2016).

It is important to note that the existence of a bilingual advantage in EF has been doubted in the current literature. Criticism concerns amongst others a publication bias for positive results and methodological flaws in the literature which increase the chance of false positive results (Paap et al., 2015; Zhou and Krott, 2016). Paap et al. (2015) also criticized the insufficient control of covariates, what could distort results. Therefore, groups need to be comparable to control the impact of other potentially influential variables like SES (Hope, 2015). When comparing mono- with multilinguals, SES is an important factor, because bilingualism and SES independently influence EF and vocabulary in 8-year-olds (Calvo and Bialystok, 2014), reading (Duzy et al., 2014; Maitz et al., 2018), and spelling (Zöller et al., 2006). When investigating multilingual school children in Germany, this variable is of major importance, because multilingual children in a German setting stem more often from families with lower SES, what might hide a potential advantage. Roos and Schöler (2009) illustrated that an initial disadvantage of multilinguals in reading and spelling disappeared when SES and intelligence were controlled for. Certain characteristics of bilinguals can also influence the outcome in research on the bilingual advantage: for example, bilinguals with good language

control (i.e., who rarely experienced unintended intrusion errors in a picture-naming task) showed better cognitive control than bilinguals who often unintentionally switched between languages (Festman and Münte, 2012).

Language-Related Skills and Spelling

Relevant language-related skills, which are known to influence the development of reading and writing, comprise lexicon size, PA, STM and phonological recoding (i.e., RAN). We refer to lexicon size as the number of lexical entries in memory for a specific language, in this paper German. For multilinguals, each language has its own mental lexicon that contains on average a smaller amount of lexical entries compared to monolinguals (Bialystok, 2009). This may not be confused with a conceptual scoring of lexicon size, which captures the number of concepts in the mental lexicon regardless of language. Thus, multilinguals may have larger mental lexicons following this account (Holmström et al., 2016). PA describes the ability to perceive phonological structures, i.e., phonemes and syllables, and manipulate elements of spoken language. As it is essential to recognize which phonemes build a word, it is relevant for non-word spelling and for the analysis of word structures. STM stores language material in memory for a short period of time, which is why it is relevant in spelling tasks involving dictation (when words and sentences need to be remembered) or writing of non-words (unknown sound sequences are stored in STM). Phonological recoding is the ability to recode written symbols into sound-based representations. It is usually assessed with RAN tasks, which measure the ability to fluently and effortlessly access phonological information by requiring participants to name letters, objects or pictures under speeded conditions (Wile and Borowsky, 2004). RAN as an automatized (i.e., fast, accurate, and effortless) skill is an important precondition for fluent and efficient spelling (Meyer et al., 1998).

These four language-related skills are known predictors of reading and spelling (for reviews see Hippmann, 2008; Verhoeven et al., 2011), however their influence depends on the language and its grapheme–phoneme consistency (Moll et al., 2014). Phonological STM has been found to predict spelling accuracy in Norwegian (Lervåg and Hulme, 2010) and in English (Caravolas, 2004). Interestingly, in a comparative study of English first and second language learners, their spelling performance was influenced by different components (Jongejan et al., 2007): verbal WM was most important for first language users, whereas for second language learners it was RAN. Specifically for German primary school, Ennemoser et al. (2012) found PA and RAN to predict both spelling and reading in the first four school years.

Lexicon size plays a special role, since it predicts literacy directly (Verhoeven and Perfetti, 2011), but it also plays an important role in the development of PA and consequently influences literacy indirectly. Different theories assume that phonological representations within an initially small mental lexicon are holistic (Lonigan et al., 2013; Goodrich et al., 2014). At this point, lexicon size is an important predictor for literacy. Only with increasing lexicon size do the phonological representations become gradually more segmented and fine-grained, what is necessary to become aware of the

phonological structure of words (see Metsala and Walley's *lexical restructuring model*, 1998, and Ziegler and Goswami's *psycholinguistic grain size theory*, 2006). This awareness allows children to use PA as resource for literacy development and the influence of lexicon size is being reduced (Metsala and Walley, 1998).

The Role of Lexicon Size for Multilinguals

Compared to monolinguals, multilinguals have a smaller lexicon in each language (for a review see Bialystok, 2009; for receptive skills, Bialystok and Luk, 2012). The root of this disparity does not lie in an impairment or language talent, but rather in the reduced amount of learning opportunities, for example due to a later start of German acquisition or a lower amount of exposure to each language (Pearson et al., 1997; Segerer et al., 2013; Gagarina and Klassert, 2018). Indeed, this is especially pronounced in multilingual children in Germany who have on average a lower level of language proficiency in German—the language of schooling—than their monolingual peers (Niklas et al., 2011; Klassert et al., 2014). Dissimilarities in reading and writing between mono- and multilinguals can be attributed to this difference in German lexical proficiency, as the impact of lexicon size on literacy performance is particularly strong for multilinguals. Especially their lower lexical abilities (Segerer et al., 2013; Limbird et al., 2014) influence writing both directly and indirectly, as described above.

Two types of studies verified the strong effect of lexicon size in multilinguals: first, comparisons between mono- and multilinguals showed that the latter were disadvantaged in tasks that required language-specific lexical knowledge (e.g., word spelling or reading), but not in less language-dependent measures (e.g., non-word spelling or reading; Weber et al., 2007; Segerer et al., 2013; Kormi-Nouri et al., 2015). Secondly, regression analyses showed that lexicon size was the most important predictor of bilinguals' but not of monolinguals' reading skills (Limbird et al., 2014).

The Role of PA for Multilinguals

Some researchers assume that multilinguals have greater PA skills (Campbell and Sais, 1995), because of their refined metalinguistic awareness (the ability to explicitly reflect upon language structure and meaning) built up by the experience of learning and managing two or more languages in life (Adesope et al., 2010). Another explanation might be a superior intrinsic sensitivity to language structure, due to their greater total vocabulary size (taken all languages together), improved attention to language (Bialystok and Herman, 1999) or the transferability of PA from one language to another (see, e.g., Lindsey et al., 2003). However, experimental results rarely support this assumption (for a review, see Jongejan et al., 2007). For example, Bialystok et al. (2003) found no advantage for bilinguals in a number of PA tasks, besides for phoneme segmentation for Spanish–English compared to Chinese–English bilinguals and English monolinguals. Laurent and Martinot (2010) found better PA in fourth and fifth, but not in third grade. Reasons for this discrepancy can be the participants' age, for example because formal literacy education influences PA; also, the specific

languages and language combinations influence PA, as certain linguistic features are more prominent in one language than in another. Finally, PA skills can transfer from one language to another (Jongejan et al., 2007). Due to these mixed findings the advantage of multilinguals in PA remains a contentious topic.

Studies with German samples could not find an advantage for bilinguals either (Duzy et al., 2013; Janssen et al., 2013), besides Limbird and Stanat (2006) who reported an advantage of multilinguals over monolinguals for pseudoword segmentation in German second graders, but no group differences in any other measurements of PA and at other time points. Despite the lack of differences in PA, Limbird and Stanat (2006) hypothesized in their study that the influence of PA on multilinguals' reading should be smaller than for monolinguals. They confirmed this hypothesis and attributed it to the multilinguals' at least numerically higher PA and on average lower reading abilities. In the study by Harrison et al. (2016), English first and second language learners were compared among others on a PA task on which they performed similarly; PA was found to be the only predictor for spelling in English first language learners, while for second language learners, PA and RAN predicted their spelling performance.

Research Focus

The goal of the present study is to determine which variables influence spelling in mono- and multilingual primary school children. Spelling involves cognitive and linguistic factors and is influenced by general background factors, such as SES. However, to our knowledge, there are no studies contrasting the influence of EF and language on spelling between mono- and multilinguals. With respect to spelling, we measured both spelling of words (which relies mainly on the lexical route) and spelling of non-words (which is based on the non-lexical route). This allows us to compare spelling abilities when German knowledge is involved (in word spelling) and when language-specific influences are minimized (in non-word spelling) (cf. Weber et al., 2007).

For cognitive factors we considered EF, namely switching, inhibition and WM (following the seminal study by Miyake et al., 2000). Until now, scientific investigations focused only on monolinguals and showed that these EF components affected spelling performance or literacy more generally. Crucially, we extended the existing research by including bi-/multilingual children.

In the linguistic domain, we focused on two language-related skills, lexicon size in German and PA. These are influential predictors of monolinguals' spelling (Verhoeven et al., 2011). However, these factors seem to influence multilinguals' spelling in different ways (for English, see Jongejan et al., 2007; Harrison et al., 2016), but their role when spelling in German is unknown.

Interestingly, EF and language-related skills (can) develop differently in mono- and multilingual children. Therefore, the experience of acquiring additional languages may alter the impact of both, language and cognition, on literacy development in these groups. For multilinguals this concerns specifically on the one hand benefits in EF (see the so-called "bilingual advantage") and PA, but on the other hand detrimental effects due to a reduced lexicon size in each language, as described above.

Our approach to investigate the role of cognitive and language skills was the following: first, we attempted to include a representative sample of today's heterogeneous school population in Germany and investigated group performance on EF and language-related skills as well as possible differences on background variables (SES, age, intelligence). Second, we wanted to find out if the three distinct EF components exerted a unique influence on spelling in these two language groups. Due to the heterogeneity of our sample, we investigated which factors influence the groups' performances rather than comparing the differences in strength of these relations. Third, we wanted to determine the influence of language-related skills on spelling in both groups. Therefore, we focused on lexicon size and PA, because multilingualism can influence their development. And by analyzing if the impact of EF remained when language-related skills were included in the same analyses, we finally aimed at revealing whether EF or language played the predominant role in third-grade mono- and multilinguals' spelling performance.

Predictions

We predicted similar performance of both groups on tasks which were—in contrast to assessments of lexicon size and word spelling—less language-dependent (e.g., non-word spelling, intelligence, PA; see Kormi-Nouri et al., 2015). Contrarily, we expected between-group differences in German lexicon size, word spelling and SES, as a disadvantage for multilingual children has been repeatedly shown for these factors or was likely (in the case of spelling; cf. Schröder-Lenzen and Merckens, 2006). Moreover, we expected no bilingual advantage on EF task performance, because of detrimental effects especially due to SES.

Despite a lack of a bilingual advantage, we assumed the influence of cognitive skills on spelling to be different between the language groups. Drawing on the literature on monolingual children, we predicted an impact mainly of switching, but no influence of inhibition and WM (as these has been found to influence rather high-level writing processes). Switching might influence word spelling more strongly than non-word spelling, since switching between alternative spelling routes (lexical and non-lexical) might be necessary. For non-words, in contrast, only phoneme-grapheme-conversion can be applied.

Comparing the impact of language-related skills and EF, we hypothesized that EF would still play a less important role for spelling than language-related skills, which are probably more relevant at the children's current developmental stage as writers (focusing on lower-level writing skills). The impact of language-related skills on spelling might be different between the groups. Limbird et al. (2014) found that German lexicon size was a stronger predictor for reading in multilinguals than monolinguals. Accordingly, we also expected lexicon size in German to be the strongest predictor for multilinguals' spelling of words and to influence monolinguals' word spelling less strongly. On non-word spelling, lexicon size should not have an influence, because it requires phonemic spelling only. Concerning PA, we expected it to affect both groups and both spelling tasks similarly, extending Ennemoser et al.'s (2012) observation for monolingual primary school children.

MATERIALS AND METHODS

Participants

Our sample consisted of 69 monolingual (33 female) and 57 multilingual children (30 female). All children attended third grade and their mean age in months was $M = 109.1$, $SD = 7.2$; see **Table 1** for descriptive statistics for all tasks and background variables of both language groups. Monolingual children spoke only German at home and had no further contact with another language. Multilingual children spoke at least one other language at home (referred to as L1) besides German and had at least good verbal proficiency in their L1 (for more detail, see section “Background Variables”). The group to which we refer as *multilinguals* consisted of bilingual ($n = 50$) and trilingual students ($n = 7$). Multilinguals spoke 21 different languages as L1 with Turkish ($n = 18$) or Arabic ($n = 8$) forming the largest subgroups. Overall, 83% of the multilingual group and 7% of the monolingual group indicated a migration background (meaning at least one parent or the child was born outside of Germany). Data for this study were selected from a larger data set of 168 third graders from different schools in Germany. Children needed to be excluded from this analysis if they could not be unambiguously assigned to one language group according to their language background or had an IQ (measured with the Culture Fair Intelligence Test Scale 1, Weiß and Osterland, 2013) below 70 indicating intellectual disabilities (DIMDI, 2016).

TABLE 1 | Descriptive statistics for spelling, EF, language-related skills, and other background factors per language group.

	Monolinguals		Multilinguals		Significance
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)	
<i>n/females</i>	69/48%		57/53%		
Age	107.6	(4.8)	111.0	(9.0)	*
Word spelling	9.5	(3.9)	11.0	(4.2)	*
Non-word spelling	4.8	(2.7)	5.3	(3.0)	
Switching	3.8	(1.2)	3.8	(1.3)	
Inhibition	1083.8	(176.5)	1032.9	(144.0)	
WM	1407.8	(275.4)	1438.2	(262.2)	
Lexicon	25.1	(6.1)	16.3	(8.4)	***
PA	5.8	(2.9)	6.0	(2.7)	
STM	15.4	(3.4)	15.2	(3.5)	
RAN	2.1	(0.4)	2.0	(0.4)	
Intelligence	35.3	(5.5)	35.3	(5.0)	
SES (mother)	3.8	(1.0)	3.1	(1.5)	**
SES (father)	3.8	(1.0)	3.2	(1.7)	*
German proficiency	3.8	(0.4)	3.6	(0.6)	*

Mean scores (with standard deviations) and group differences calculated with two-tailed *t*-tests (Significance *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$) between mono- and multilinguals. [Mean values refer to: Age – in months; word spelling, non-words spelling – word errors; switching – number of experienced rules; inhibition – interference effect in ms; WM – mean reaction times in ms; lexicon, PA, STM, RAN, intelligence – number of correct answers; SES – ISCED; German proficiency – rating between 1 (none) and 4 (very good)].

The study and the protocol were approved by the ethics committee of the University of Potsdam (approval nr. 11/2015), the head of the schools, the *Ministry of Education, Youth and Health* (Land Brandenburg) and the *Senate for Education, Youth and Science* (Berlin). The study was carried out in accordance with the ethical standards laid down in the Declaration of Helsinki. We recruited participants through class teachers who distributed written information for the parents about the study in all third-grade classes at three inclusive primary schools. The parents gave written informed consent before the beginning of the study. All participants had normal or corrected-to-normal vision, were naive as to the purpose of the study and received small gifts for their participation.

Materials

Spelling

Word spelling

In the word spelling task, children were asked to produce the correct orthographic writing of single words. We used the respective subtest of the German standardized test battery BUEGA (Esser et al., 2008). It comprises 17 items with varying phonological and morphological complexity. Items were prerecorded and presented via loudspeaker to guarantee equal testing conditions for all participants. The number of incorrectly spelled words was used as measure of task performance.

Non-word spelling

To contrast children's word spelling performance (which is confounded with their knowledge of German) with spelling purely based on phoneme-grapheme conversion, we designed a non-word spelling task. Here, children were required to produce the phonologically correct spelling of a non-word. This test comprised 12 non-words of varying length (2–3 syllables; see **Supplementary Material**). Quasi-universal stimuli were constructed by using (a) CV and CVC syllable structures (to avoid German-specific structures like consonant clusters), and (b) phonemes that are shared in most languages (to minimize the influence of rare or language-specific sounds such as [x] in German). The selection of phonemes, based on Maddieson (2013a,b), resulted in the following consonants: voiceless fricatives (i.e., [f], [s]), voiceless plosives (i.e., [p], [t], [k]), nasals (i.e., [m], [n]), the lateral approximant [l], and the three vowels [a], [i], and [u] that represent the three extreme points on the vowel diagram (International Phonetic Association, 2015). To guarantee equal testing conditions for all children, all non-words were prerecorded and presented via loudspeaker. The number of incorrectly spelled non-words served as a measure of non-word spelling.

Executive Functions

Switching

The Wisconsin Card Sorting Test (WCST) was chosen to assess mental set shifting in cognitive control, i.e., to infer the implied rule in card sorting guided by feedback from the environment, to change the rule when necessary and to apply the correct rule

continuously (Spreen and Strauss, 1998). We used an online version of the WCST (adapted from Piper et al., 2012). The task required children to sort 64 cards according to one of three different rules (color, number, or shape) by tapping on the corresponding digital card stack on the tablet screen. As mentioned before, the current rule was not revealed but needed to be derived from visual feedback, which followed each answer. After 10 correct trials in a row, a rule change was signaled by three pink exclamation marks appearing for 750 ms. There was no time limit for answers, but children were encouraged to answer as quickly as possible. Instructions and examples were given orally to the whole group before the beginning of the training. The number of experienced rules was used in this study as a measure of switching, since the more successfully the children answered, the more rule changes they experienced.

Inhibition

To test participants' ability to inhibit interfering information (i.e., color) within visually presented stimulus material, we administered the Bivalent Shape Task (BST, Mueller and Esposito, 2014). In this test, children sorted shapes (circles and squares) by pressing corresponding buttons on a tablet screen. The more salient color of the shapes defined to which condition a stimulus belonged: (a) In the neutral condition, black outlines were presented, (b) in the congruent condition, shape and color (red and blue) matched the respective answer button, and (c) in the incongruent condition, colors were interchanged, so that the shape corresponded to the correct button, but the color corresponded to the incorrect button. The incongruent condition required inhibition to refrain from pressing the incorrect button. The experiment consisted of three uniform blocks in a fixed order (neutral, followed by congruent and incongruent) comprising 20 randomized trials, and one mixed block comprising 30 randomized trials, 10 of each condition. The test began with oral instructions and examples, followed by a practice block (12 randomized trials) with visual feedback. In the experimental blocks, no feedback was given, and each item appeared immediately after button press or after a time limit of 3 s. We adapted the BST (Mueller and Esposito, 2014) to counterbalance the position (right and left) and color (blue and red) of the answer buttons over participants. As measure of inhibition, we used the interference effect (reaction time difference between incongruent and congruent trials) of the mixed blocks.

Working memory

To assess participants' WM, we chose the n-back task that requires the ability to temporarily store information in memory, process it and continuously update the stored information. In this task, single letters (A, B, O, R, and S) were presented one-by-one in the center of the tablet screen. Children were asked to press a button at the bottom of the screen when the displayed letter was the same as two trials ago (two-back). Thus, they had to compare the newly perceived information with older already stored information. The task comprised two blocks with 20 pseudorandomized items of which eight were critical items (no more than two critical items in a sequence). Response times were

restricted to 4 s in which the stimulus was presented for 2 s and an inter-stimulus interval with a blank screen appeared for 2 s. The task started with 10 practice trials including visual feedback, but afterward no feedback was given. The mean reaction time was used as a measure of WM.

Language-Related Skills

Lexicon size

To capture the size of the expressive lexicon in German, we used a standardized online picture naming task, the WWT (short version test 2, Glück, 2011). This test consisted of 40 colored items eliciting nouns, verbs or opposites of adjectives. It was self-paced, but time restricted, and administered on tablet computers. The number of correct answers was used as a measure of lexicon size.

Phonological awareness

As the participants were already in third grade, we chose a task which measured the level of PA at a higher degree of difficulty (rather than phoneme identification, segmentation, and syntheses; see Anthony and Francis, 2005). We wanted to assess complex phonological skills measured by sub-syllabic tasks and used an inversion test which relied on manipulative abilities of phonemic knowledge, namely the phoneme-inversion subtest of the standardized German test BAKO (subtest 4, Stock et al., 2003). The prerecorded items were presented one-by-one. The task required children to invert the first two sounds in 11 items (six German words and five non-words) and pronounce it out loud. The output measure was the number of correct answers.

Phonological short-term memory

A non-word repetition task was used to measure the children's STM, since the ability to retain and repeat increasingly longer verbal stimuli serves as indicator of STM capacity. STM was tested with the non-word repetition task from the ZLT-II (Petermann et al., 2013). This task consists of 30 non-words built of meaningless CV-syllables and thereby avoids language-specific structures like consonant clusters. Items were prerecorded to standardize testing conditions and pronounced with a neutral prosody (equal stress on each syllable) to reduce German-specific prosodic patterns. The non-words were presented with increasing length, from two to six syllables. Children were asked to repeat each non-word after its presentation. The number of correctly repeated non-words served as a measure of STM.

Rapid automatized naming

Rapid automatized naming of letters was used to measure phonological recoding. More specifically, we assessed naming speed of an array of letters, which indexes the ability to effectively access and retrieve phonological entries of graphemes. RAN is considered to be a measure of automaticity (Strattman and Hodson, 2005). We chose the RAN test from the standardized German test battery TEPHOBE (Mayer, 2013). Participants received a sheet of paper with 50 letters and were required to name them in sequence as fast and as accurately as possible. The number of correct answers per second indexed naming speed of letters.

Background Variables

Questionnaires

Background information on the children and their family were derived from a paper–pencil questionnaire filled out by the parents at home (a German and a Turkish version of the questionnaire were available). To measure the family's SES we asked for each parent's highest school and professional qualification, which was then categorized by the ISCED (International Standard Classification of Education, UNESCO Institute for Statistics, 2012) on a scale from 0 to 6. Information on the child's language background, the beginning of his/her acquisition of German and the languages spoken at home and in kindergarten were obtained in the questionnaire. The parents' rating of their child's oral language skills (i.e., mean rating for speaking and comprehension on a scale of 1 = *none* to 4 = *very good*) served as a measure of language proficiency.

Intelligence

To assess the participants' non-verbal intelligence, in particular their ability to recognize and continue figurative relationships and logical sequences, we used the Culture Fair Intelligence Test Scale 1 (CFT 1-R, Weiß and Osterland, 2013). We assessed these skills with three subtests, namely Matrix, Series, and Classification. This standardized, non-verbal test was specifically chosen, because it is a culture-independent and language-free test, and consequently, it should not disadvantage children from different cultural backgrounds or with poorer German skills. The sum of correct answers served as a measure of intelligence.

Procedure

This study is part of a larger project comprising the parents' questionnaire and three experimental sessions. We report here only the tasks relevant for this study: in a first group session, children completed the word spelling, inhibition, and switching task on a tablet computer (Microsoft Surface Pro 2 Tablet, display size: 25.5 cm × 17 cm, resolution: 2160 px × 1440 px). In the second group session, we administered the test for non-word spelling and WM on the same tablet, and the intelligence test with a paper–pencil version. In the third, individual session, we included all tasks which required recording children's verbal responses, i.e., German lexicon size, PA, STM, and RAN.

In group sessions, up to 14 children participated and were supervised by trained experimenters who ensured a quiet atmosphere, correct administration of tests, and understanding of instructions despite language, comprehension, or processing speed difficulties.

Data Analysis

Data preparation for the EF tasks included exclusion of single participants whose performance was below chance or who did not participate in one of the tasks (switching: two monolinguals; inhibition: two multilinguals; WM: five multilinguals). Afterward, for the BST and N-back, reaction times of correct responses were log-transformed to normalize distributions, and outliers in form of single data points were removed by visual inspection of the reaction time distribution for each task (inhibition: single blocks with less than four correct

answers and in total, 0.1% of all data points; WM: in total 0.9% of all data points). For all other tasks that included verbal responses (that concerns lexicon, PA, STM, and RAN), audio-files were transcribed. These data and the paper–pencil tests (i.e., intelligence, word and non-word spelling) were rated by one- and double-checked by another person to obtain correct ratings for all items in every tasks.

All calculations were run on R version 3.2.2 (R Core Team, 2015). In a first step, we compared all measures between the language groups with two-tailed *t*-tests to reveal performance differences in cognitive and language measures as well as in background variables. Then, we calculated correlations between the tasks measuring spelling, EF, language-related skills and the potentially influential factors intelligence and SES, for each language group separately. Spearman correlation coefficients were computed with the *rcorr* function from the *Hmisc* package (Harrell, 2017). Finally, we calculated linear mixed effects models with *glmer* function from the *lme4* package (Pinheiro et al., 2015) separately for each spelling task and for each group to find out which predictors influenced mono- and multilinguals in these tasks. We decided not to calculate one model with multilingualism as a predictor besides EF and language tasks, due to the heterogeneity of our sample and group differences (e.g., in age and migration status) that we could not control without overfitting this model. Independent variables were the raw values of the spelling tasks (correct/incorrect rating for every item) and random intercepts for participants and item were calculated to control individual differences and effects on an item-level (e.g., increasing fatigue or the impact of a preceding trial that was answered correctly or wrong; Baayen et al., 2008). As dependent variables, we used the same variables in every step for both groups and all outcome measures. These were z-transformed to allow for comparison between the predictors. First, the influence of the three EF components was tested. Secondly, those EF with significant influence were entered into the models together with the linguistic predictors. These comprised German lexicon size and PA, since multilingualism potentially influences their development or impact on spelling. In a third step, SES was added (note that we do not report these last results in detail, since model fit did not improve by adding SES. For more detail, see below). Other factors like intelligence were not entered into the models to avoid overfitting them and due to intercorrelations between cognitive, language and background factors (see below). Model fit for the generalized mixed-effect models was estimated with the marginal (R^2_m ; variance explained by the fixed effects) and conditional coefficient (R^2_c ; variance explained by fixed and random effects). Both were calculated with the *r.squaredGLMM* function from the package *MuMIn* version 1.41.0 (Burnham and Anderson, 2002; Nakagawa and Schielzeth, 2013).

RESULTS

Group Comparisons

Group comparisons between mono- and multilinguals for all variables are displayed in **Table 1**. (For results of error rates in the EF tasks, see **Supplementary Table 1**.) As can be seen in

Table 1, monolinguals and multilinguals did not differ on most variables. They performed equally on the non-word spelling test, on all three EF tasks (switching, inhibition, working memory), on most tasks testing language-related skills, i.e., PA, STM and RAN, as well as on intelligence.

Contrastingly, monolinguals produced significantly less errors in word spelling. Regarding language-related factors, they had a significantly larger expressive lexicon than their multilingual peers. Moreover, the groups differed on a number of other background variables such that monolinguals were significantly younger and showed significantly higher SES (mother's and father's ISCED), and parents rated their monolingual children's proficiency in German significantly higher than did parents of multilingual children.

Correlations

The correlation coefficients of mono- and multilinguals' performance in spelling, cognitive, and language tasks as well as background factors are presented in **Table 2**. First, the results show that the correlation between the two spelling tasks (i.e., more errors in spelling words correlated with more errors in spelling non-words) was much higher for monolinguals ($r = 0.61$, $p < 0.001$) than for multilinguals ($r = 0.38$, $p < 0.001$).

The three EF tasks did not correlate with each other, as they represent distinct EF-subfunctions. Regarding correlations of spelling with EF, we observed that for monolinguals switching correlated with spelling words (higher number of experienced rules correlated with fewer errors in spelling words; $r = -0.25$, $p < 0.05$), whereas for multilinguals it was inhibition which correlated with word spelling (larger interference effect correlated with fewer errors in word spelling; $r = -0.27$, $p < 0.05$). Apart from that, inhibition and WM hardly correlated with any of the other factors. Note that in both groups switching correlated with lexicon and intelligence (more experienced rules correlated with a larger lexicon and a higher intelligence score; monolinguals: $r = 0.27$, $p < 0.05$ and $r = 0.45$, $p < 0.001$; multilinguals: $r = 0.31$, $p < 0.05$ and $r = 0.28$, $p < 0.05$).

The four language-related measures (lexicon, PA, STM, and RAN) correlated with spelling of words and non-words in both groups (with the exception for multilinguals regarding correlations of lexicon with spelling non-words, $r = -0.23$, and STM with spelling words, $r = -0.24$): fewer errors in the spelling tasks correlated with a larger lexicon, and better PA, STM, and RAN. The language-related measures also correlated with one another in both groups, for example better performance in RAN with better PA (monolinguals: $r = 0.41$, $p < 0.01$, multilinguals: $r = 0.32$, $p < 0.05$), and better STM with greater lexicon size (but only for monolinguals: $r = 0.41$, $p < 0.001$).

Furthermore, we found relatively high intercorrelations of intelligence with several measures (for monolinguals a higher intelligence score correlated with fewer errors in both spelling tasks, more experienced rules in the switching task, larger lexicon, and better PA, whereas for multilinguals a higher intelligence score correlated with fewer errors when spelling words, better switching, and a larger lexicon; see **Table 2**). The correlation of higher SES with a larger lexicon was very high for multilinguals ($r = 0.57$, $p < 0.001$), but moderate for monolinguals ($r = 0.36$, $p < 0.001$). Furthermore, higher SES correlated with better STM for monolinguals and with better RAN for multilinguals.

Predictors of Word Spelling

Our first regression model (see **Table 3**, upper part) explores the influence of the three EF components on word spelling. For the monolingual group (left column), switching ($b = -0.36$, $SE = 0.16$, $p < 0.05$) influenced word spelling significantly. Better performance on this EF task (i.e., more experienced rules in the WCST) was associated with less errors in the spelling of words. For multilinguals (right column), inhibition predicted performance in word spelling ($b = -0.46$, $SE = 0.21$, $p < 0.05$), with a larger interference effect being associated with fewer errors. The fixed effects in both models explained only a small amount of variance ($R^2m = 0.03$ for monolinguals, $R^2m = 0.04$ for multilinguals).

The second regression model (see **Table 3**, lower part) includes inhibition and switching, because of their significance in the first model, and lexicon and PA. Word spelling in monolinguals

TABLE 2 | Correlation coefficients of monolinguals (below diagonal) and multilinguals (above diagonal) for spelling, EF, language-related skills, and other background factors.

	1	2	3	4	5	6	7	8	9	10	11
(1) Words		0.38***	-0.15	-0.27*	0.06	-0.37***	-0.38***	-0.24	-0.42***	-0.28*	-0.17
(2) Non-words	0.61***		-0.20	-0.10	0.02	-0.23	-0.4***	-0.48***	-0.44***	-0.13	-0.16
(3) Switching	-0.25*	-0.19		-0.02	0.27	0.31*	0.17	0.32*	-0.07	0.28*	0.16
(4) Inhibition	-0.07	0.17	0.02		-0.11	0.19	-0.11	0.20	0.12	0.16	-0.01
(5) WM	0.17	0.07	-0.20	-0.15		0.08	0.09	0.01	-0.02	0.07	0.18
(6) Lexicon	-0.31*	-0.38***	0.27*	-0.07	0.22		0.15	0.26	0.26	0.38***	0.57***
(7) PA	-0.46***	-0.37***	0.19	0.18	-0.23	0.24		0.34*	0.32*	0.19	0.24
(8) STM	-0.42***	-0.5***	0.12	-0.08	-0.01	0.41***	0.34***		0.16	-0.05	0.14
(9) RAN	-0.49***	-0.53***	0.16	0.05	-0.15	0.16	0.41***	0.26*		-0.02	0.33*
(10) Intelligence	-0.28*	-0.32*	0.45***	0.05	-0.17	0.37***	0.35***	0.12	0.24		0.29
(11) SES (mother)	0.01	0.00	0.10	0.01	0.16	0.36***	-0.03	0.28*	-0.11	0.02	

Spearman correlation coefficients (Significance *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$).

TABLE 3 | Fixed effects of the linear mixed effects models predicting number of errors in word spelling.

		Monolinguals				Multilinguals			
		<i>b</i>	<i>SE</i>	<i>z</i>	Significance	<i>b</i>	<i>SE</i>	<i>z</i>	Significance
(1)	(Intercept)	0.14	0.35	0.40		0.66	0.41	1.60	
	Switching	−0.36	0.16	−2.19	*	−0.27	0.23	−1.16	
	Inhibition	−0.08	0.16	−0.49		−0.46	0.21	−2.16	*
	WM	0.13	0.16	0.81		0.23	0.23	1.01	
	R^2m/R^2c				0.03/0.5				0.04/0.58
(2)	(Intercept)	0.16	0.34	0.48		0.63	0.38	1.68	+
	Switching	−0.22	0.14	−1.53		0.06	0.18	0.34	
	Inhibition	−0.04	0.14	−0.27		−0.37	0.17	−2.17	*
	Lexicon	−0.23	0.15	−1.52		−0.47	0.18	−2.58	*
	PA	−0.59	0.15	−4.00	***	−0.67	0.18	−3.77	***
	R^2m/R^2c				0.09/0.5				0.12/0.55

Regression models were calculated for mono- and multilinguals separately. In model (1) the three EF components are included as predictors, and predictors in model (2) are lexicon size, PA and the significant EF from (1). R^2m represents the variance explained by the fixed effects and R^2c by fixed and random effects. (Significance *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$).

was significantly predicted only by PA (i.e., the better their PA, the fewer spelling errors they made; $b = -0.59$, $SE = 0.15$, $p < 0.001$). For the multilingual group PA ($b = -0.67$, $SE = 0.18$, $p < 0.001$) and lexicon size ($b = -0.47$, $SE = 0.18$, $p < 0.05$) were significant predictors with better PA and a greater expressive lexicon predicting fewer spelling errors. Additionally, the impact of inhibition remained for multilinguals' spelling performance ($b = -0.37$, $SE = 0.17$, $p < 0.05$). Both models explained now a larger amount of variance for mono- and multilinguals ($R^2m = 0.09$ and $R^2m = 0.12$, respectively).

Note that in a third model, we added the mothers' ISCED to control for the impact of SES, but this did not improve model fit: for monolinguals, R^2m declined from $R^2m = 0.09$ to $R^2m = 0.07$, and for multilinguals from $R^2m = 0.12$ to $R^2m = 0.11$. Therefore, we report only the models without SES here (see the **Supplementary Table 3** for these additional models).

Predictors of Non-word Spelling

For both language groups, we found an effect of switching on non-word spelling in our first model that is significant for monolinguals ($b = -0.28$, $SE = 0.13$, $p < 0.05$) and marginally significant for multilinguals ($b = -0.36$, $SE = 0.20$, $p < 0.1$). These effects indicate that the more categories the children could master in the WCST, the fewer spelling errors they produced (see **Table 4**, upper part). The explained variance of the fixed effects is equal in both groups, but very low ($R^2m = 0.02$).

In the second model (see **Table 4**, in the middle), only switching was included, since WM and inhibition did not influence non-word spelling in either group in the first model. For monolinguals, we found lexicon size ($b = -0.36$, $SE = 0.12$, $p < 0.01$) and PA ($b = -0.27$, $SE = 0.12$, $p < 0.05$) to significantly predict non-word spelling performance such that the larger the lexicon and the better the PA abilities, the lower the error rate when spelling non-words. For the multilingual group, the only significant predictor of non-word spelling was PA ($b = -0.49$, $SE = 0.17$, $p < 0.01$). The explained variance of the fixed effects

improved for both groups ($R^2m = 0.06$ for monolinguals and $R^2m = 0.08$ for multilinguals).

As for word spelling, adding the mothers' ISCED to the regression models reduced model fit (for monolinguals: to $R^2m = 0.05$, and for multilinguals: to $R^2m = 0.07$). Hence, we report only the models without SES (but these additional models can be found in **Supplementary Table 4**).

Since the second models explained relatively little variance, we decided to run another model including STM (see **Table 4**, lower part). STM is responsible for short-term storage of verbal information and therefore essential to successfully perform on our non-word spelling task as children had to memorize non-words of increasing length. When STM was entered into the model, it was the strongest predictor for monolinguals ($b = -0.38$, $SE = 0.12$, $p < 0.01$) and multilinguals ($b = -0.47$, $SE = 0.18$, $p < 0.01$) revealing that the better the STM abilities, the fewer errors were made when spelling non-words. For multilinguals, additionally PA influenced non-word spelling ($b = -0.36$, $SE = 0.16$, $p < 0.05$). In sum, STM is the best predictor for non-word spelling. That neither lexicon nor PA maintained their significance for the monolinguals' performance is partially caused by the intercorrelations between STM and lexicon ($r = 0.41$, $p < 0.001$), and STM and PA (monolinguals: $r = 0.34$, $p < 0.001$). The models with STM fit the data best, since they explained the largest amount of variance (for monolinguals $R^2m = 0.08$ and for multilinguals $R^2m = 0.1$).

DISCUSSION

We investigated the influence of EF and language-related skills on spelling in a group of mono- and multilingual third graders. By including a word and a non-word spelling task, we were able to contrast spelling based on lexical knowledge (word spelling) and phonemic spelling (non-word spelling via the non-lexical route). We assessed naturally heterogeneous groups of mono- and multilingual third graders with inherent differences

TABLE 4 | Fixed effects of the linear mixed effects models predicting number of errors in non-word spelling.

		Monolinguals				Multilinguals			
		<i>b</i>	<i>SE</i>	<i>z</i>	Significance	<i>b</i>	<i>SE</i>	<i>z</i>	Significance
(1)	(Intercept)	−0.50	0.27	−1.87	+	−0.20	0.32	−0.62	
	Switching	−0.28	0.13	−2.12	*	−0.36	0.20	−1.80	+
	Inhibition	0.15	0.13	1.17		−0.15	0.19	−0.81	
	WM	−0.02	0.13	−0.12		0.12	0.20	0.63	
	R^2m/R^2c				0.02/0.36				0.02/0.26
(2)	(Intercept)	−0.48	0.26	−1.86	+	−0.35	0.31	−1.15	
	Switching	−0.13	0.12	−1.06		−0.18	0.18	−1.06	
	Lexicon	−0.36	0.12	−2.92	**	−0.28	0.18	−1.61	
	PA	−0.27	0.12	−2.24	*	−0.49	0.17	−2.90	**
	R^2m/R^2c				0.06/0.36				0.08/0.46
(3)	(Intercept)	−0.49	0.26	−1.91	+	−0.33	0.30	−1.11	
	Switching	−0.15	0.11	−1.36		−0.09	0.17	−0.56	
	Lexicon	−0.19	0.13	−1.54		−0.22	0.17	−1.34	
	PA	−0.16	0.12	−1.37		−0.36	0.16	−2.16	*
	STM	−0.38	0.12	−3.05	**	−0.47	0.18	−2.63	**
	R^2m/R^2c				0.08/0.36				0.1/0.46

Regression models were calculated for mono- and multilinguals separately. In model (1) the three EF components are included as predictors, in model (2) predictors are lexicon size, PA and the (marginally) significant EF from (1), and in (3) STM was added to model (2). R^2m represents the variance explained by the fixed effects and R^2c by fixed and random effects. (Significance *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$).

in SES and lexicon size to the detriment of multilinguals. Despite these differences, the groups did not differ in their performance on non-word spelling, PA and EF, whereas monolinguals outperformed multilinguals in the word spelling and German lexicon test. Our regression analyses revealed that switching explained a small amount of variance in word and non-word spelling for monolinguals. Contrastingly for multilinguals, inhibition influenced word spelling and a trend indicated an influence of switching in non-word spelling. When we added lexicon size and PA—two important predictors of literacy—to our models, the influence of switching disappeared in both groups and tasks, but inhibition remained as predictor for multilinguals' word spelling. Overall, language was a better predictor for spelling and both language groups shared the most influential factors in each spelling task: PA for spelling words and STM for spelling non-words. As predicted, some language-related skills influenced only multilinguals: lexicon size predicted their word spelling, and PA influenced their non-word spelling performance.

The Role of EF in Spelling

Since our sample consisted of very heterogeneous groups of children including multilinguals with disadvantages in SES and lexicon size, we neither expected nor found a bilingual advantage in any of the three EF components. From the literature on the bilingual advantage we know that differences in EF between language groups tend to appear only when groups are well-matched to reduce the influence of confounding variables (Hope, 2015). After all, many variables influence EF (Diamond, 2013), and bi- or multilingualism is only one of them. As Morton and Harper (2007) showed, an apparent bilingual advantage could stem from a hidden advantage in SES from which bilinguals' EF benefited. In our study, the opposite was the case,

that is, multilinguals' disadvantage in SES most likely balanced out the possible positive effects of multilingualism, since both influenced EF (Calvo and Bialystok, 2014).

In the word spelling task, we found a differential role of EF on mono- versus multilinguals. Monolinguals' word spelling performance was influenced by switching, as we predicted, and multilinguals' performance by inhibition. However, the impact of EF on monolinguals disappeared, when language-related skills were controlled for in our models. This can be explained by the role of the lexicon size. In contrast to multilinguals, monolinguals likely have already built up a large number of orthographic entries in their orthographic lexicon, probably linked to their better proficiency in German in general, and the larger lexicon in particular. The large orthographic lexicon enabled them to rely on their refined lexical skills for efficient processing during the spelling tasks, what freed mental resources and EF could come into play. We conclude that in primary school language skills initially play the dominant role for spelling. We further speculate that the major impact of language skills on spelling lasts until the developing language processing skills reach a certain threshold of proficiency (these might relate to the size of the lexicon, the efficiency and automaticity of processing, the well-established use of the lexical route). After this threshold has been passed, freed cognitive resources could be redirected to EF to improve coordination of parallel processes and develop toward higher-level processes of writing for which the influence of EF has been documented in the literature.

For multilinguals, the explanation of the impact of EF on spelling that we found is more difficult. We did not expect this effect of inhibition, since earlier studies (Altemeier et al., 2008; Lubin et al., 2016) with monolinguals did not find an effect of inhibition on spelling. The discrepancy between our finding

and the literature might be caused by different tasks involving different processes apart from inhibition: Altemeier et al. (2008) used a Color-Word interference test, including word reading, and Lubin et al. (2016) tested inhibition with an Opposite World task, requiring children to name digits. The BST that we administered in our study is language-independent, because it involves shape identification and inhibition of color (as the more salient feature). Another possibility for our different finding is that we tested multilinguals. Additionally, the direction of the effect is somewhat counterintuitive, since a larger interference effect—indicating poorer inhibition skills—predicted less spelling errors. This relation even remained significant when language skills were statistically controlled for. The fact that we found this effect only in word spelling indicates that it is connected to processing German, the language of schooling, and not to general writing processes. To clarify this relation and replicate this finding, however, further studies are necessary.

In non-word spelling, both groups rely on the same mechanisms, because German language knowledge is not relevant for task performance: we found a significant effect of switching for monolinguals and the trend in the same direction for multilinguals. The similarity between the groups is likely caused by the minimized need for German specific processing in this task allowing multilinguals to exploit their potential regarding the use of EF. As we mentioned above, monolinguals might be able to use more EF due to their advanced German skills. Multilinguals, however, need more cognitive resources for language processing (like phoneme analysis that is necessary in non-word spelling) resulting in less influence of EF. Here, the marginally significant effect might be caused by individual differences in this group, because some advanced spellers might already have the capacity to utilize switching for spelling. It may be used as an additional resource to make spelling more efficient, since basic language processing skills proceed more automatically than in less proficient spellers.

Why is switching relevant for both spelling tasks? The influence of switching on spelling seems to be replicable for spelling in different languages. This concerns the studies by Lubin et al. (2016) for French 9-year-olds, by von Suchodoletz et al. (2017) for German pupils, and the longitudinal study by Altemeier et al. (2008) showing that rapid automatized switching predicted spelling in English in the first 3 years of school. Spellers need to switch between the ongoing parallel processes during spelling, like language processing, accessing lexical entries or phoneme-grapheme correspondences, graphomotor and output control (Lubin et al., 2016). These demands are universal for word and non-word spelling, which is why switching is the EF component relevant for word and non-word spelling.

Our finding that WM did not influence spelling is in accordance with the literature. WM likely comes into play at a later age and during more complex writing tasks, since WM influences higher-level tasks in writing, like text reviewing process that take place in advanced writers who have managed low-level writing processes (Swanson and Berninger, 1996). The spelling of single words and non-words does not require children to master these higher-level processes for successful task performance.

Similar Main Components for Mono- and Multilinguals' Spelling

Despite some differences in EF involvement, mono- and multilinguals shared the main components during spelling: first, PA was the most influential factor in spelling performance of words (see Table 3) and non-words (if STM is not considered, see second model in Table 4). PA is essential for phonemic writing (for non-words and unknown real words) and spelling via the lexical-route (Pfost, 2015). The ability to recognize syllables and phonemes is necessary to identify phonemes for phoneme-grapheme conversion and to determine morphemes. With regard to spelling words, our results add to the evidence provided by Ennemoser et al. (2012) who found that monolingual German primary school children relied mostly on PA for spelling in first to fourth grade (for Norwegian, see Lervåg and Hulme, 2010; for English, Caravolas, 2004) and extends this to multilinguals in German primary school.

The second main component both language groups shared was STM as strongest predictor for non-word spelling (see third model in Table 4). Non-word spelling necessarily relies on STM, because phonological information must be stored in memory during mental processing (Martin and Gupta, 2004; Repovš and Baddeley, 2006) and during non-word dictation. The relation between STM and spelling is well-documented in the literature: for example, Swanson and Berninger (1996) have determined the influence of STM on spelling of letters and words in a sample of 10- to 12-year-olds, or Wimmer and Mayringer (2002) found children with a spelling deficit to have a smaller STM capacity. Initially, the relation between STM and spelling was not in the focus of our study, since STM is not a component of EF (Miyake et al., 2000). But the poor model fit with only EF and language-related skills as predictors (see Table 4 middle part) made it necessary to investigate further influential variables.

The Impact of Language Proficiency on Multilinguals' Spelling

Our results confirm the greater role of lexicon size for multilinguals' literacy. Limbird et al. (2014) found this link between lexicon size and reading in the first 3 years of primary school and we expand this association to word spelling of third graders. According to the *lexical restructuring model*, lexicon size influences spelling to a certain threshold after which it does not play a role anymore (Metsala and Walley, 1998). Only when words are stored as phonologically detailed representations, the role of lexicon size in word spelling decreases, as we could see in our monolingual group. However, many children in our multilingual group seem to be below that threshold due to their smaller lexicon size in German, meaning that they have not yet developed phonologically fine-grained lexical representations for German, their language of schooling (also found by Niklas et al., 2011; Segerer et al., 2013; Limbird et al., 2014).

Another indication of the strong impact of language-related skills on multilinguals is that language-dependent tests posed a greater challenge for multilinguals in our study. In comparison

to their monolingual peers, multilinguals performed more poorly in word spelling and lexicon, but similarly on all other tasks (i.e., non-word spelling, EF, PA, STM, RAN, and intelligence) that relied as little as possible on German language knowledge. Our results are in line with other studies that found similar result patterns: Kormi-Nouri et al. (2015) found Iranian bilinguals in grade one to five to show poorer word reading, but similar non-word reading performance compared to their monolingual peers. Studies with German school children also revealed that multilinguals underperformed in language-dependent tasks but performed similarly in tests that relied less on German knowledge (Weber et al., 2007; Duzy et al., 2013; Segerer et al., 2013). To avoid an undesired impact of lexicon in testing situations with multilinguals (Messer, 2010; Parra et al., 2011), tests should involve single letters, non-words or visual items. One concrete example is our non-word spelling task, which was designed to minimize language-specific influences. In contrast to other tests using pseudowords (e.g., Hasselhorn et al., 2012), the items were based on quasi-universal structures, with common vowels and consonants, a simple syllable structure (without language-specific structures like consonant clusters) and without morphology (see task description in section “Materials”). We consider this approach essential to assess actual phoneme-based spelling performance in groups with diverse language experience and proficiency (Schöppe et al., 2013).

Interestingly, PA played a greater role in non-word spelling for multilinguals than for monolinguals. For word reading, Limbird and Stanat (2006) found the opposite pattern, as described above. PA exerts a greater impact on multilinguals in our study, because they likely have less fine-grained phonological representations due to their smaller lexicon size (Metsala and Walley, 1998). Non-word spelling relies entirely on the non-lexical route which requires correct phoneme identification to translate phonemes into graphemes. Consequently, multilinguals with more holistic lexical representations might have problems identifying phonemes correctly, what makes PA a more important predictor for their non-word spelling performance compared to monolinguals.

Limitations

Studying the interplay between further potentially important factors on spelling is still necessary; this concerns for example SES, migration background and the lower language status of migrant languages in Germany (Plewnia and Rothe, 2011). In our study, the impact of SES on multilinguals' language skills is supported by the higher correlation between lexicon and SES for multilinguals than monolinguals, but SES did not add to explaining spelling performance beyond EF and language skills. However, multilinguals' disadvantage in SES has negative repercussions on language skills (Calvo and Bialystok, 2014), literacy (Roos and Schöler, 2009) and school success in general (Zöller et al., 2006), because parents with higher SES are more likely to provide early literacy activities, a stimulating educational input, like access to media, experiences and multiple and diverse language learning opportunities. The

latter are especially important for multilingual children who need to acquire German often outside their home (Zöller et al., 2006).

Our choice of EF tasks relied on the tripartite model of EF (Miyake et al., 2000), but we need to acknowledge that the strict division of EF in three separate components has been questioned in the literature (Friedman and Miyake, 2017). Moreover, the precise measurement of EF components with one task has been criticized due to *task impurity* (Friedman, 2016). Task impurity stems from superficial factors like stimuli characteristics (e.g., words versus pictures) or response modality (e.g., motoric versus verbal) that might alter characteristics of an EF task. This concerns especially the WCST as measure of switching, because it is a quite complex task (Best and Miller, 2010): it comprises three stimulus categories, three possible rules, and it requires problem solving strategies to reveal the new rules, inhibition of inappropriate responses and of irrelevant stimuli characteristics, etc. (Miyake et al., 2000; Diamond, 2013). It is therefore possibly a more general measure of EF, but further investigations should verify the role of switching in spelling with other experimental designs.

Our analyses do not allow us to compare the strength our predictors have on spelling in mono- versus multilinguals. For example, Limbird et al. (2014) found PA to influence reading more strongly in monolinguals than multilinguals, but we cannot draw this kind of conclusions from our data with regard to spelling. Therefore, more comparable groups need to be investigated and the differential influence of for example EF needs to be calculated in one model. We refrained from this strategy, because of the heterogeneity of our language groups concerning the differences in migrations status and age that could not be controlled for without risking to overfit regression models.

CONCLUSION

We studied a naturally heterogeneous sample of mono- and multilingual third graders in Germany, with multilinguals having on average a lower SES and smaller German lexicon size. In our study, we contrasted the influence of cognitive (i.e., the EF components switching, inhibition and WM) and language factors (i.e., lexicon size and PA) on word and non-word spelling in these groups. EF explained only a small amount of variance in both spelling tasks in both groups. Switching predicted monolinguals' word spelling, whereas word spelling in multilinguals was predicted by inhibition. In non-word spelling, both groups shared switching as the only predictor (the impact for multilinguals was only marginally significant). Since the effect of switching disappeared when language was controlled for, we postulate that language processing initially takes up more cognitive resources that are not available for EF. This is the case for multilinguals for whom language factors play a predominant role in spelling due to their smaller German lexicon size. Beyond this threshold, language processing (e.g., lexical access, phoneme-grapheme conversion) is so fluent that cognitive resources are freed and

EF become more influential—the monolinguals in our study are likely at this developmental stage.

Comparing the impact of language and cognition, we found that language-related skills exerted a greater influence on spelling than EF and mono- and multilinguals shared the main predictors for spelling: PA in word spelling and STM in non-word spelling. Our study also replicated the strong role of lexicon size for multilinguals' word spelling. This relation needs to be considered when comparing mono- with multilinguals, since many tests use verbal stimuli or depend on language in other ways, what potentially disadvantages multilinguals.

AUTHOR CONTRIBUTIONS

SC, AK, and JF contributed equally to the design and preparation of the study as well as to data acquisition. SC carried out the statistical analysis, and SC and JF composed the article. All authors read and approved the final manuscript. This study was carried out in collaboration between the authors.

REFERENCES

- Abbott, R. D., and Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *J. Educ. Psychol.* 85, 478–508. doi: 10.1037/0022-0663.85.3.478
- Adesope, O. O., Lavin, T., Thompson, T., and Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Rev. Educ. Res.* 80, 207–245. doi: 10.3102/0034654310368803
- Alamargot, D., and Chanquoy, L. (2001). *Through the Models of Writing. Studies in Writing*, Vol. 9. London: Kluwer Academic Publishers. doi: 10.1007/978-94-010-0804-4
- Allan, N. P., Hume, L. E., Allan, D. M., Farrington, A. L., and Lonigan, C. J. (2014). Relations between inhibitory control and the development of academic skills in preschool and kindergarten: a meta-analysis. *Dev. Psychol.* 50, 2368–2379. doi: 10.1037/a0037493
- Altemeier, L. E., Abbott, R. D., and Berninger, V. W. (2008). Executive functions for reading and writing in typical literacy development and dyslexia. *J. Clin. Exp. Neuropsychol.* 30, 588–606. doi: 10.1080/13803390701562818
- Anthony, J. L., and Francis, D. J. (2005). Development of phonological awareness. *Curr. Direct. Psychol. Sci.* 14, 255–259. doi: 10.1111/j.0963-7214.2005.00376.x
- Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. doi: 10.1016/j.jml.2007.12.005
- Baddeley, A. D. (2010). Working memory. *Curr. Biol.* 20, R136–R140. doi: 10.1016/S0079-7421(08)60452-1
- Barry, C. (1994). “Spelling routes (or roots or rutes),” in *Handbook of Spelling: Theory, Process and Intervention*, eds G. D. A. Brown and N. C. Ellis (Chichester: Wiley & Sons), 27–49.
- Berninger, V. W., Abbott, R. D., Cook, C. R., and Nagy, W. (2016). Relationships of attention and executive functions to oral language, reading, and writing skills and systems in middle childhood and early adolescence. *J. Learn. Disabil.* 50, 1–16. doi: 10.1177/0022219415617167
- Berninger, V. W., Abbott, R. D., Jones, J., Wolf, B. J., Gould, L., Anderson-Youngstrom, M., et al. (2006). Early development of language by hand: composing, reading, listening, and speaking connections; three letter-writing modes; and fast mapping in spelling. *Dev. Neuropsychol.* 29, 61–92. doi: 10.1207/s15326942dn2901_5
- Berninger, V. W., Abbott, R. D., Swanson, H. L., Lovitt, D., Trivedi, P., Lin, S.-J. C., et al. (2010). Relationship of word- and sentence-level working memory to reading and writing in second, fourth, and sixth grade. *Lang. Speech Hear. Serv. Sch.* 41, 179–193. doi: 10.1044/0161-1461(2009/08-2)
- Berninger, V. W., and Amtmann, D. (2003). “Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: research into practice,” in *Handbook of Learning Disabilities*, eds H. L. Swanson, K. R. Harris, and S. Graham (New York, NY: Guilford Press), 345–363.
- Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x
- Best, J. R., Miller, P. H., and Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5–17 in a large, representative national sample. *Learn. Ind. Differ.* 21, 327–336. doi: 10.1016/j.lindif.2011.01.007
- Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Bilingualism* 12, 3–11. doi: 10.1017/S1366728908003477
- Bialystok, E. (2015). Bilingualism and the development of executive function: the role of attention. *Child Dev. Perspect.* 9, 117–121. doi: 10.1111/cdep.12116
- Bialystok, E., and Herman, J. (1999). Does bilingualism matter for early literacy? *Bilingual. Lang. Cognit.* 2, 35–44. doi: 10.1017/S1366728999000139
- Bialystok, E., and Luk, G. (2012). Receptive vocabulary differences in monolingual and bilingual children. *Bilingual. Lang. Cognit.* 15, 397–401. doi: 10.1017/S1366728909990423
- Bialystok, E., Majumder, S., and Martin, M. M. (2003). Developing phonological awareness: is there a bilingual advantage? *Appl. Psychol.* 24, 27–44. doi: 10.1017/S014271640300002X
- Bialystok, E., and Martin, M. M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Dev. Sci.* 7, 325–339. doi: 10.1111/j.1467-7687.2004.00351.x
- Bialystok, E., and Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition* 112, 494–500. doi: 10.1016/j.cognition.2009.06.014
- Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x
- Blom, E., Kuntay, A. C., Messer, M. H., Verhagen, J., and Leseman, P. (2014). The benefits of being bilingual: working memory in bilingual turkish-dutch children. *J. Exp. Child Psychol.* 128, 105–119. doi: 10.1016/j.jecp.2014.06.007
- Bourdin, B. B., and Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *Int. J. Psychol.* 29, 591–620. doi: 10.1080/00207599408248175

FUNDING

This research was funded by the Land Brandenburg, Germany.

ACKNOWLEDGMENTS

We thank all members of the *Research Group: Diversity and Inclusion* for their help preparing the study, collecting and processing the data and discussing results in group meetings. We are especially grateful to the students who participated, their parents and the schools for their cooperation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00097/full#supplementary-material>

- Burnham, K. P., and Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York, NY: Springer-Verlag.
- Calvo, A., and Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition* 130, 278–288. doi: 10.1016/j.cognition.2013.11.015
- Campbell, R., and Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *Br. J. Dev. Psychol.* 13, 61–68. doi: 10.1111/j.2044-835X.1995.tb00664.x
- Caravolas, M. (2004). Spelling development in alphabetic writing systems: a cross-linguistic perspective. *Eur. Psychol.* 9, 3–14. doi: 10.1027/1016-9040.9.1.3
- Chudaske, J. (2012). *Sprache, Migration und Schulfachliche Leistung. Einfluss Sprachlicher Kompetenz auf Lese-, Rechtschreib- und Mathematikleistungen*. Wiesbaden: VS Verlag für Sozialwissenschaften. doi: 10.1007/978-3-531-93495-2
- Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750
- DIMDI (ed.) (2016). “ICD-10-GM. version 2016,” in *Internationale Statistische Klassifikation der Krankheiten und Verwandter Gesundheitsprobleme, 10. Revision* (Cologne: DIMDI).
- Ditton, H., and Krüskens, J. (2009). Denn wer hat, dem wird gegeben werden? Eine Längsschnittstudie zur Entwicklung schulischer Leistungen und den Effekten der sozialen Herkunft in der Grundschulzeit. *J. Bildungsforschung* 1, 33–61.
- Duzy, D., Ehm, J. H., Souvignier, E., Schneider, W., and Gold, A. (2013). Prädiktoren der Lesekompetenz bei Kindern mit Deutsch als Zweitsprache. *Z. Entwicklungspsychol. Pädagogis. Psychol.* 45, 173–190. doi: 10.1026/0049-8637/a000093
- Duzy, D., Souvignier, E., Ehm, J. H., and Gold, A. (2014). Early decoding speed and later reading competencies in children with German as a second language. *Child Indic. Res.* 7, 787–804. doi: 10.1007/s12187-014-9242-x
- Ehri, L. C. (2005). Learning to read words: theory, findings, and issues. *Sci. Am. Mind* 9, 167–188. doi: 10.1007/s00167-011-1795-3
- Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Sci. Stud. Read.* 18, 5–21. doi: 10.1080/10888438.2013.819356
- Engel de Abreu, P. M. J. (2011). Working memory in multilingual children: is there a bilingual effect? *Memory* 19, 529–537. doi: 10.1080/09658211.2011.590504
- Ennemoser, M., Marx, P., Weber, J., and Schneider, W. (2012). Spezifische Vorläuferfertigkeiten der Lesegeschwindigkeit, des Leseverständnisses und des Rechtschreibens. Evidenz aus zwei Längsschnittstudien vom Kindergarten bis zur 4. Klasse. *Z. Entwicklungspsychol. Pädagogis. Psychol.* 44, 53–67. doi: 10.1026/0049-8637/a000057
- Esser, G., Wyschkon, A., and Ballaschk, K. (2008). *BUEGA. Basisdiagnostik Umschriebener Entwicklungsstörungen im Grundschulalter*. Göttingen: Hogrefe.
- Festman, J., and Münte, T. F. (2012). Cognitive control in Russian-German bilinguals. *Front. Psychol.* 3:115. doi: 10.3389/fpsyg.2012.00115
- Franconeri, S. L., Alvarez, G. A., and Cavanagh, P. (2013). Flexible cognitive resources: competitive content maps for attention and memory. *Trends Cognit. Sci.* 17, 134–141. doi: 10.1016/j.tics.2013.01.010
- Friedman, N. P. (2016). Research on individual differences in executive functions: implications for the bilingual advantage hypothesis. *Linguist. Approaches Biling.* 6, 535–548. doi: 10.1075/lab.15041.fri
- Friedman, N. P., and Miyake, A. (2017). Unity and diversity of executive functions: individual differences as a window on cognitive structure. *Cortex* 86, 186–204. doi: 10.1016/j.cortex.2016.04.023
- Gagarina, N., and Klassert, A. (2018). Input dominance and development of home language in Russian-German bilinguals. *Front. Commun.* 3:40. doi: 10.3389/fcomm.2018.00040
- Gerth, S., Dolk, T., Klassert, A., Fliesser, M., Fischer, M. H., Nottbusch, G., et al. (2016). Adapting to the surface: a comparison of handwriting measures when writing on a tablet computer and on paper. *Hum. Move. Sci.* 48, 62–73. doi: 10.1016/j.humov.2016.04.006
- Glück, C. W. (2011). *Wortschatz- und Wortfindungstest für 6-bis 10-Jährige (WWT 6-10)*, 2nd Edn. München: Urban & Fischer.
- Goodrich, J. M., Lonigan, C. J., and Farver, J. M. (2014). Children's expressive language skills and their impact on the relation between first- and second-language phonological awareness skills. *Sci. Stud. Read.* 18, 114–129. doi: 10.1080/10888438.2013.819355
- Grabowski, J. (2010). Speaking, writing, and memory span in children: Output modality affects cognitive performance. *Int. J. Psychol.* 45, 28–39. doi: 10.1080/00207590902914051
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingual. Lang. Cognit.* 1, 67–81. doi: 10.1017/S1366728998000133
- Harrell, F. E. (2017). *Package 'Hmisc'*. Vienna: R Foundation for Statistical Computing.
- Harrison, G. L., Gogan, L. D., Jalbert, R., McManus, K., Sinclair, K., and Spurling, J. (2016). Predictors of spelling and writing skills in first- and second-language learners. *Read. Writ.* 29, 69–89. doi: 10.1007/s11145-015-9580-1
- Hasselhorn, M., Schumann-Hengsteler, R., Gronauer, J., Grube, D., Mähler, C., Schmid, I., et al. (eds.) (2012). *AGTB 5-12. Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12 Jahren*. Göttingen: Hogrefe.
- Hilchey, M. D., Saint-Aubin, J., and Klein, R. M. (2015). “Does bilingual exercise enhance cognitive fitness in traditional non-linguistic executive processing tasks?” in *The Cambridge Handbook of Bilingual Processing*, ed. J. W. Schwieter (Cambridge: Cambridge University Press), 286–613.
- Hippmann, K. (2008). *Prädiktoren des Schriftspracherwerbs im Deutschen*. Aachen: Rheinisch-Westfälische Technische Hochschule Aachen.
- Holmström, K., Salameh, E. K., Nettelbladt, U., and Dahlgren-Sandberg, A. (2016). Conceptual scoring of lexical organization in bilingual children with language impairment. *Commun. Disord. Quart.* 38, 24–34. doi: 10.1177/1525740115601999
- Hooper, S. R., Swartz, C. W., Wakely, M. B., de Kruijff, R. E. L., and Montgomery, J. W. (2002). Executive functions in elementary school children with and without problems in written expression. *J. Learn. Disabil.* 35, 57–68. doi: 10.1177/002221940203500105
- Hope, T. M. H. (2015). The bilingual cognitive advantage: no smoke without fire? *AIMS Neurosci.* 2, 58–65. doi: 10.3934/Neuroscience.2015.2.58
- International Phonetic Association (2015). IPA Chart. Available under a Creative Commons Attribution-ShareAlike 3.0 Unported License. Copyright © 2015 International Phonetic Association. Available at: <http://www.internationalphoneticassociation.org/content/ipa-chart>
- Janssen, M., Bosman, A. M. T., and Leeman, P. P. M. (2013). Phoneme awareness, vocabulary and word decoding in monolingual and bilingual Dutch children. *J. Res. Read.* 36, 1–13. doi: 10.1111/j.1467-9817.2011.01480.x
- Jongejan, W., Verhoeven, L., and Siegel, L. S. (2007). Predictors of reading and spelling abilities in first- and second-language learners. *J. Educ. Psychol.* 99, 835–851. doi: 10.1037/0022-0663.99.4.835
- Kellogg, R. T. (1996). “A model of working memory in writing,” in *The Science of Writing: Theories, Methods, Individual Differences and Applications*, eds C. M. Levy and S. Ransdell (Mahwah, NJ: L. Erlbaum Associates), 57–72.
- Klassert, A., Gagarina, N., and Kauschke, C. (2014). Object and action naming in Russian- and German-speaking monolingual and bilingual children. *Bilingualism* 17, 73–88. doi: 10.1017/S1366728913000096
- Kormi-Nouri, R., Jalali-Moghadam, N., and Moradi, A. (2015). The dissociative effects in lexical and non-lexical reading tasks for bilingual children. *Br. J. Educ. Soc. Behav. Sci.* 8, 47–62. doi: 10.9734/BJESBS/2015/16760
- Kroll, J. F., Dussias, P. E., Bogulski, C. A., and Kroff, J. R. V. (2012). Juggling two languages in one mind: what bilinguals tell us about language processing and its consequences for cognition. *Psychol. Learn. Motiv.* 56, 229–262. doi: 10.1016/B978-0-12-394393-4.00007-8
- Laurent, A., and Martinot, C. (2010). Bilingualism and phonological awareness: the case of bilingual (French-Occitan) children. *Read. Writ.* 23, 435–452. doi: 10.1007/s11145-009-9209-3
- Leierhoff, H., Rehkämper, K., Rockmann, U., Brunner, M., Gärtner, H., Wendt, W., et al. (2013). *Bildung in Berlin und Brandenburg 2013*, ed. Autorengruppe Regionale Bildungsberichterstattung Berlin-Brandenburg. Berlin: ISQ-Bericht zur Schulqualität.
- Lervåg, A., and Hulme, C. (2010). Predicting the growth of early spelling skills: are there heterogeneous developmental trajectories? *Sci. Stud. Read.* 14, 485–513. doi: 10.1080/10888431003623488
- Limbird, C. K., Maluch, J. T., Rjosk, C., Stanat, P., and Merckens, H. (2014). Differential growth patterns in emerging reading skills of Turkish-German bilingual and German monolingual primary school students. *Read. Writ.* 27, 945–968. doi: 10.1007/s11145-013-9477-9
- Limbird, C. K., and Stanat, P. (2006). “Prädiktoren von Leseverständnis bei Kindern deutscher und türkischer Herkunftssprache: Ergebnisse einer

- Längsschnittstudie," in *Veränderungsmessung und Längsschnittstudien in der Empirischen Erziehungswissenschaft* eds A. Ittel and H. Merckens (Wiesbaden: VS Verlag für Sozialwissenschaften), 93–123. doi: 10.1017/CBO9781107415324.004
- Linsey, K. A., Manis, F. R., and Bailey, C. E. (2003). Prediction of first-grade reading in spanish-speaking english-language learners. *J. Educ. Psychol.* 95, 482–494. doi: 10.1037/0022-0663.95.3.482
- Lonigan, C. J., Farver, J. M., Nakamoto, J., and Eppe, S. (2013). Developmental trajectories of preschool early literacy skills: a comparison of language-minority and monolingual-English children. *Dev. Psychol.* 49, 1943–1957. doi: 10.1037/a0031408
- Lubin, A., Regrin, E., Boulc'h, L., Pacton, S., and Lanoë, C. (2016). Executive functions differentially contribute to fourth graders' mathematics, reading, and spelling skills. *J. Cognit. Educ. Psychol.* 15, 444–463. doi: 10.1891/1945-8959.15.3.444
- Luo, L., Craik, F. I. M., Moreno, S., and Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: evidence from aging. *Psychol. Aging* 28, 28–34. doi: 10.1037/a0030875
- Maddieson, I. (2013a). "Consonant inventories," in *The World Atlas of Language Structures Online*, eds M. S. Dryer and M. Haspelmath (Leipzig: Max Planck Institute for Evolutionary Anthropology).
- Maddieson, I. (2013b). "Vowel quality inventories," in *The World Atlas of Language Structures Online*, eds M. S. Dryer and M. Haspelmath (Leipzig: Max Planck Institute for Evolutionary Anthropology).
- Maitz, K., Paleczek, L., Seifert, S., and Gasteiger-Klicpera, B. (2018). Zusammenhang der Leseverständnisleistungen mit sozialen Herkunftsfaktoren bei Schülerinnen der dritten Schulstufe. *Z. Grundschulforsch.* 11, 147–160. doi: 10.1007/s42278-018-0012-6
- Martin, N., and Gupta, P. (2004). Exploring the relationship between word processing and verbal short-term memory: evidence from associations and dissociations. *Cognit. Neuropsychol.* 9, 213–229. doi: 10.1080/02643290342000447
- Marx, A. E., and Stanat, P. (2012). Reading comprehension of immigrant students in Germany: research evidence on determinants and target points for intervention. *Read. Writ.* 25, 1929–1945. doi: 10.1007/s11145-011-9307-x
- Mayer, A. (2013). *TEPHOBE Test zur Erfassung der Phonologischen Bewusstheit und der Benennungsschwierigkeit*, 2nd Edn. München: Ernst Reinhardt Verlag.
- McCutchen, D. (1996). A capacity theory of writing: working memory in composition. *Educ. Psychol. Rev.* 8, 299–325. doi: 10.1007/BF01464076
- McCutchen, D. (2011). From novice to expert: implications of language skills and writing-relevant knowledge for memory during the development of writing skills. *J. Writ. Res.* 3, 51–68. doi: 10.17239/jowr-2011.03.01.3
- Messer, M. (2010). *Verbal Short-Term Memory and Vocabulary Development in Monolingual Dutch and Bilingual Turkish-Dutch Preschoolers*. Enschede: Ipskamp Drukkers.
- Metsala, J. L., and Walley, A. C. (1998). "Spoken vocabulary growth and the segmental restructuring of lexical representations: precursors to phonemic awareness and early reading ability," in *Word Recognition in Beginning Literacy*, eds J. L. Metsala and L. C. Ehri (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 89–120.
- Meyer, M. S., Wood, F. B., Hart, L. A., and Felton, R. H. (1998). Longitudinal course of rapid naming in disabled and nondisabled readers. *Ann. Dyslexia* 48, 91–114. doi: 10.1007/s11881-998-0005-6
- Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Direct. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognit. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734
- Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., et al. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learn. Instruct.* 29, 65–77. doi: 10.1016/j.learninstruc.2013.09.003
- Monette, S., Bigras, M., and Guay, M. C. (2011). The role of the executive functions in school achievement at the end of Grade 1. *J. Exp. Child Psychol.* 109, 158–173. doi: 10.1016/j.jecp.2011.01.008
- Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x
- Mueller, S. T., and Esposito, A. G. (2014). Computerized testing software for assessing interference suppression in children and adults: the bivalent shape task (BST). *J. Open Res. Softw.* 2:e3. doi: 10.5334/jors.ak
- Müller, A. G., and Stanat, P. (2006). "Schulischer Erfolg von Schülerinnen und Schülern mit Migrationshintergrund: Analysen zur Situation von Zuwanderern aus der ehemaligen Sowjetunion und aus der Türkei," in *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit*, eds J. Baumert, P. Stanat, and R. Watermann (Wiesbaden: Verlag für Sozialwissenschaften), 221–256. doi: 10.1007/978-3-531-90082-7_6
- Nakagawa, S., and Schielzeth, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods Ecol. Evol.* 4, 133–142. doi: 10.1111/j.2041-210x.2012.00261.x
- Namazi, M., and Thordardottir, E. (2010). A working memory, not bilingual advantage, in controlled attention. *Int. J. Biling. Educ. Bilingual.* 13, 597–616. doi: 10.1080/13670050.2010.488288
- Niklas, F., Schmiedeler, S., Pröstler, N., and Schneider, W. (2011). Die Bedeutung des Migrationshintergrunds, des Kindergartenbesuchs sowie der Zusammensetzung der Kindergartengruppe für sprachliche Leistungen von Vorschulkindern. *Z. Pädagogis. Psychol.* 25, 115–130. doi: 10.1024/1010-0652/a000032
- OECD (2014). *PISA 2012 Ergebnisse: Exzellenz durch Chancengerechtigkeit (Band II): Allen Schülerinnen und Schülern die Voraussetzungen zum Erfolg sichern*. Bielefeld: W. Bertelsmann Verlag. doi: 10.1787/9789264208858-de
- Paap, K. R., Johnson, H. A., and Sawi, O. M. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69, 265–278. doi: 10.1016/j.cortex.2015.04.014
- Parra, M., Hoff, E., and Core, C. (2011). Relations among language exposure, phonological memory, and language development in Spanish-English bilingually developing 2-year-olds. *J. Exp. Child Psychol.* 108, 113–125. doi: 10.1016/j.jecp.2010.07.011
- Pearson, B. Z., Fernández, S., Lewedeg, V., and Oller, K. (1997). The relation of input factors to lexical learning by bilingual infants. *Appl. Psycholinguist.* 18, 41–58. doi: 10.1017/S0142716400009863
- Perfetti, C. (2017). "Lexical quality revisited," in *Developmental Perspectives in Written Language and Literacy: In honor of Ludo Verhoeven*, eds E. Segers and P. van den Broek (Amsterdam: John Benjamins Publishing), 51–67.
- Petermann, F., Daseking, M., Linder, M., Grisseman, H., and Weid, T. (2013). *Zürcher Lesetest-II: ZLT-II; Weiterentwicklung des Zürcher Lesetests (ZLT) von Maria Linder und Hans Grisseman*, 3rd Edn. Bern: H. Huber.
- Pfost, M. (2015). Children's phonological awareness as a predictor of reading and spelling. A systematic review of longitudinal research in German-speaking countries. *Z. Entwicklungspsychol. Pädagogis. Psychol.* 47, 123–138. doi: 10.1026/0049-8637/a000141
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team. (2015). *nlme: Linear and Nonlinear Mixed Effects Models. R Package Version 3.1-137*. Available at: <https://CRAN.R-project.org/package=nlme>
- Piper, B. J., Li, V., Eiwaz, M. A., Kobel, Y. V., Benice, T. S., Chu, A. M., et al. (2012). Executive function on the psychology experiment building language tests. *Behav. Res. Methods* 44, 110–124. doi: 10.3758/s13428-011-0173-x
- Plewnia, A., and Rothe, A. (2011). "Spracheinstellungen und Mehrsprachigkeit, wie Schüler über ihre und andere Sprachen denken," in *Sprache und Integration. Über Mehrsprachigkeit und Migration*, eds L. M. Eichinger, A. Plewnia, and M. Steinle (Tübingen: Narr), 215–253.
- Poarch, G. J. (2018). Multilingual language control and executive function: a replication study. *Front. Commun.* 3:46. doi: 10.3389/fcomm.2018.00046
- Poarch, G. J., and van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: evidence from second-language learners, bilinguals, and trilinguals. *J. Exp. Child Psychol.* 113, 535–551. doi: 10.1016/j.jecp.2012.06.013
- Pontart, V., Bidet-Ildei, C., Lambert, E., Morisset, P., Flouret, L., and Alamargot, D. (2013). Influence of handwriting skills during spelling in primary and lower secondary grades. *Front. Psychol.* 4:818. doi: 10.3389/fpsyg.2013.00818

- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Core Team.
- Repovš, G., and Baddeley, A. (2006). The multi-component model of working memory: explorations in experimental cognitive psychology. *Neuroscience* 139, 5–21. doi: 10.1016/j.neuroscience.2005.12.061
- Rønneberg, V., and Torrance, M. (2017). Cognitive predictors of shallow-orthography spelling speed and accuracy in 6th grade children. *Read. Writ.* 32, 1–20. doi: 10.1007/s11145-017-9751-3
- Roos, J., and Schöler, H. (2009). *Entwicklung des Schriftspracherwerbs in der Grundschule*. Wiesbaden: VS Verlag für Sozialwissenschaften. doi: 10.1007/978-3-531-91574-6
- Schoonen, R., Van Gelderen, A., Gloppe, K., de, Hulstijn, J., Snellings, P., et al. (2002). “Linguistic knowledge, metacognitive knowledge and retrieval speed in L1, L2, and EFL writing,” in *Studies in Writing: New Directions for Research in L2 Writing*, Vol. 11, eds G. Rijlaarsdam, S. Ransdell, and M. Barbier (Dordrecht: Kluwer Academic Publishers), 101–122.
- Schöppe, D., Blatter, K., Faust, V., Jäger, D., Stanat, P., Artelt, C., et al. (2013). Effekte eines Trainings der phonologischen Bewusstheit bei Vorschulkindern mit unterschiedlichem Sprachhintergrund. *Z. Pädagog. Psychol.* 27, 241–254. doi: 10.1024/1010-0652/a000110
- Schründer-Lenzen, A., and Merckens, H. (2006). “Differenzen schriftsprachlicher Kompetenzentwicklung bei Kindern mit und ohne Migrationshintergrund,” in *Risikofaktoren Kindlicher Entwicklung. Migration, Leistungsangst und Schulübergang*, A. Schründer-Lenzen (Wiesbaden: VS Verlag für Sozialwissenschaften), 15–44. doi: 10.1007/978-3-531-90075-9
- Segerer, R., Marx, A. E., Stanat, P., Schneider, W., Roick, T., and Marx, P. (2013). “Determinanten der Lesekompetenz bei Jugendlichen nicht deutscher Herkunftssprache. Zur Bedeutung der Spracherwerbsreihenfolge im Falle von Mehrsprachigkeit,” in *PISA 2009 – Impulse für die Schul- und Unterrichtsforschung*, eds N. Jude and E. Klieme (Weinheim: Beltz), 111–131.
- Spalek, K., Hoshino, N., Wu, Y. J., Damian, M., and Thierry, G. (2014). Speaking two languages at once: unconscious native word form access in second language production. *Cognition* 133, 226–231. doi: 10.1016/j.cognition.2014.06.016
- Spreen, O., and Strauss, E. (1998). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. New York, NY: Oxford University Press.
- St Clair-Thompson, H. L., and Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quart. J. Exp. Psychol.* 59, 745–759. doi: 10.1080/17470210500162854
- Stock, C., Marx, P., and Schneider, W. (2003). *Basiskompetenzen für Leserechtschreibleistungen (BAKO 1 - 4)*. Göttingen: Hogrefe.
- Strattman, K., and Hodson, B. W. (2005). Variables that influence decoding and spelling in beginning readers. *Child Lang. Teach. Ther.* 21, 165–190. doi: 10.1191/0265659005ct287oa
- Swanson, H. L., and Berninger, V. W. (1996). Individual differences in children's working memory and writing skill. *J. Exp. Child Psychol.* 63, 358–385. doi: 10.1006/jecp.1996.0054
- Treiman, R. (2017). Learning to spell: Phonology and beyond. *Cognit. Neuropsychol.* 34, 83–93. doi: 10.1080/02643294.2017.1337630
- UNESCO Institute for Statistics (2012). “The international standard classification of education, ISCED 2011” in *Comparative Social Research*, Vol. 30 (Montreal: UNESCO Institute for Statistics). doi: 10.1108/S0195-631020130000030017
- Vanderberg, R., and Swanson, H. L. (2007). Which components of working memory are important in the writing process? *Read. Writ.* 20, 721–752. doi: 10.1007/s11145-006-9046-6
- Verhoeven, L., and Perfetti, C. A. (2011). Introduction to this special issue: vocabulary growth and reading skill. *Sci. Stud. Read.* 15, 1–7. doi: 10.1080/10888438.2011.536124
- Verhoeven, L., Reitsma, P., and Siegel, L. S. (2011). Cognitive and linguistic factors in reading acquisition. *Read. Writ.* 24, 387–394. doi: 10.1007/s11145-010-9232-4
- von Suchodoletz, A., Fäsche, A., and Skuballa, I. T. (2017). The role of attention shifting in orthographic competencies: cross-sectional findings from 1st, 3rd, and 8th grade students. *Front. Psychol.* 8:1665. doi: 10.3389/fpsyg.2017.01665
- Weber, J., Marx, P., and Schneider, W. (2007). Die Prävention von Leserechtschreibschwierigkeiten bei Kindern mit nichtdeutscher Herkunftssprache durch ein Training der phonologischen Bewusstheit. *Z. Pädagog. Psychol.* 21, 65–75. doi: 10.1024/1010-0652.21.1.65
- Weiß, R. H., and Osterland, J. (2013). *CFT 1-R Grundintelligenztest Skala 1 – Revision*. Göttingen: Hogrefe.
- Wile, T. L., and Borowsky, R. (2004). What does rapid automatized naming measure? A new RAN task compared to naming and lexical decision. *Brain Lang.* 90, 47–62. doi: 10.1016/S0093-934X(03)00419-X
- Wimmer, H., and Landerl, K. (1997). “How learning to spell German differs from learning to spell English,” in *Learning to Spell: Research, Theory, and Practice across Languages*, eds C. A. Perfetti, L. Rieben, and M. Fayol (Mahwah: Lawrence Erlbaum Associates), 81–96.
- Wimmer, H., and Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties: a specific disability in regular orthographies. *J. Educ. Psychol.* 94, 272–277. doi: 10.1037/0022-0663.94.2.272
- Wiseheart, M., Viswanathan, M., and Bialystok, E. (2016). Flexibility in task switching by monolinguals and bilinguals. *Bilingual. Lang. Cognit.* 19, 141–146. doi: 10.1017/S1366728914000273
- Zhang, M. (2018). An overview of the bilingual advantage: history, mechanisms, and consequences. *Western Undergrad. Psychol. J.* 6, 1–9.
- Zhou, B., and Krott, A. (2016). Data trimming procedure can eliminate bilingual cognitive advantage. *Psychon. Bull. Rev.* 23, 1221–1230. doi: 10.3758/s13423-015-0981-6
- Ziegler, J. C., and Goswami, U. (2006). Becoming literate in different languages: similar problems, different solutions. *Dev. Sci.* 9, 429–436. doi: 10.1111/j.1467-7687.2006.00509.x
- Zöller, I., Roos, J., and Schöler, H. (2006). “Einfluss soziokultureller Faktoren auf den Schriftspracherwerb im Grundschulalter,” in *Risikofaktoren kindlicher Entwicklung. Migration, Leistungsangst und Schulübergang*, ed. A. Schründer-Lenzen (Wiesbaden: VS Verlag für Sozialwissenschaften), 45–65. doi: 10.1007/978-3-531-90075-9_2
- Zorza, J. P., Marino, J., and Mesas, A. A. (2017). Executive functions as predictors of school performance and social relationships: primary and secondary school students. *Span. J. Psychol.* 19:E23. doi: 10.1017/sjp.2016.23

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Czapka, Klassert and Festman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Sequential Congruency Effects in Monolingual and Bilingual Adults: A Failure to Replicate Grundy et al. (2017)

Samantha F. Goldsmith and J. Bruce Morton*

Department of Psychology, Brain and Mind Institute, University of Western Ontario, London, ON, Canada

Previous research suggests bilingual adults show smaller sequential congruency effects than monolingual adults. Here we re-examined these findings by administering an Eriksen flanker task to monolingual and bilingual adults. The task produced robust conventional and sequential congruency effects. Neither effect differed for monolingual and bilingual adults. Results are discussed in terms of current debates concerning differences in cognitive control between monolingual and bilingual adults.

Keywords: cognitive control, Eriksen flanker task, bilingualism, bilingual advantage, sequential congruency effects

OPEN ACCESS

Edited by:

Roberto Filippi,
Institute of Education, University
College London, United Kingdom

Reviewed by:

Scott R. Schroeder,
Hofstra University, United States

Wei Li,
University College London,
United Kingdom

*Correspondence:

J. Bruce Morton
jbrucemorton@gmail.com

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 16 July 2018

Accepted: 21 November 2018

Published: 11 December 2018

Citation:

Goldsmith SF and Morton JB (2018)
Sequential Congruency Effects in
Monolingual and Bilingual Adults: A
Failure to Replicate Grundy et al.
(2017). *Front. Psychol.* 9:2476.
doi: 10.3389/fpsyg.2018.02476

INTRODUCTION

Bilingualism and Cognitive Control: Are There Differences?

One longstanding and rather vexing question in the study of human psychology concerns whether a lifetime of bilingualism leads to measurable changes in cognitive control. Several accounts predict that it should. According to Green (1998), for example, everyday language use is challenging for bilinguals as it requires the selection of words and meanings from a target language amidst competition from translation equivalents of a non-target language. Because managing cross-language interference relies on general control processes, bilinguals become highly practiced—and thus advantaged—in problems of cognitive control relative to monolinguals.

Mixed Evidence in Adults

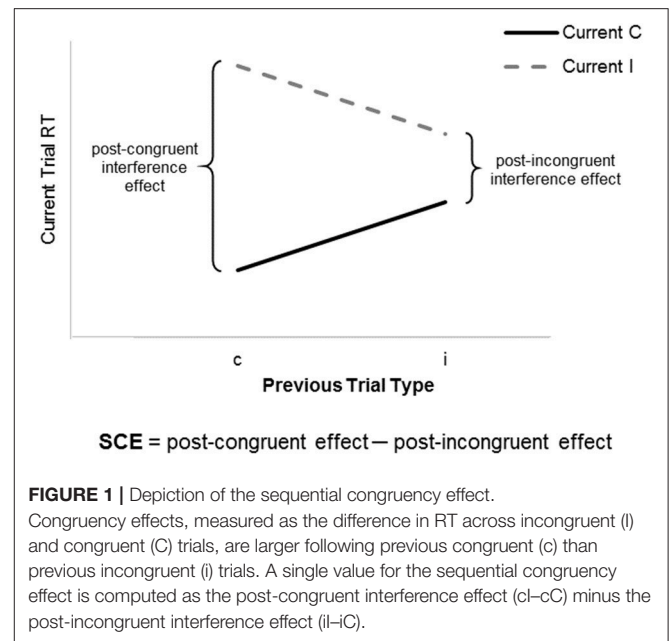
Decades of research have yielded some empirical support for the bilingual advantage hypothesis, mostly in the form of evidence that the distracting effect of irrelevant stimuli is typically smaller for bilinguals than monolinguals (e.g., Bialystok et al., 2004). One aspect of the available evidence that is difficult to reconcile with a simple formulation of the bilingual advantage hypothesis is the fact that the bilingual advantage is more consistently observed in studies of monolingual and bilingual children than it is in studies of monolingual and bilingual adults. Several large-scale adult studies have failed to find any differences between monolinguals and bilinguals across a wide range of cognitive control tasks (Paap and Greenberg, 2013). And in cases where adult differences have been reported, these differences disappear after only a few blocks of trials (Bialystok et al., 2004). If the bilingual advantage reflects a lifetime of experience managing cross-language interference, why is the advantage more pronounced (not less) in young children than in adults? The growing number of large-scale replication failures has led a number of vocal critics to claim there is no coherent evidence for a bilingual advantage in cognitive control.

In defense of the bilingual advantage hypothesis, some have dismissed concerns about the null effects of adult studies. One argument is that adult response times in cognitive

control tasks are quite small (on average, ~ 500 ms), and therefore group differences need to be large for statistically significant differences to emerge. For children, response times are considerably larger, and therefore group differences are easier to detect (see Grundy et al., 2017, p. 43). This argument is obviously flawed, as it is the variance of two distributions, rather than the difference in their means, that determines whether or not a group difference will be statistically significant. Moreover, because response time *variability* is greater in children than in adults, it is typically *harder* to detect group differences in children, even when the absolute value of those differences is larger.

A more interesting suggestion is that differences between monolingual and bilingual adults do exist, but are evident only given careful choice of cognitive control measures and analyses. Following this line of reasoning, Grundy et al. (2017) administered an Eriksen flanker task to groups of monolingual and bilingual adults. Across repeated trials, participants responded to the direction of a centrally presented arrow (press left key for “<”; press right key for “>”). On congruent trials, the target arrow was flanked by arrows pointing the same direction (< < < < or > > > >); on incongruent trials, the target arrow was flanked by arrows pointing the opposite direction (> > < > or < < > <). Groups were compared in two ways. First, they were compared in terms of a conventional congruency or interference effect, computed as the difference in response time on incongruent vs. congruent trials. Consistent with other findings (e.g., Paap and Greenberg, 2013), this conventional analysis revealed no difference between monolingual and bilingual adults. However, a second more advanced analysis compared groups in terms of a sequential congruency effect, computed as the difference in interference effects *following* congruent vs. incongruent trials (refer to **Figure 1**). Although relatively easy to estimate from flanker data, sequential congruency effects of monolingual and bilingual adults had not hitherto been compared. Interestingly, bilinguals showed a *smaller* sequential congruency effect than monolinguals: for bilinguals, interference effects measured after congruent trials were comparable to interference effects measured after incongruent trials, whereas for monolinguals, interference effects measured after congruent trials were larger than interference effects measured after incongruent trials. The findings provide a nice illustration of the idea that differences between monolingual and bilingual adults are subtle and may require careful choice of methods to reveal.

But what do these differences mean? According to Grundy et al., differences in the sequential congruency effect suggest that bilinguals more efficiently disengage attention from previous stimuli (both congruent and incongruent), affording them an advantage of greater attentional focus on current trials, relative to monolinguals. This claim is partially supported by evidence that greater practice on stimulus-response compatibility tasks is associated with smaller sequential congruency effects (e.g., van Steenbergen et al., 2015). That said, the claim that smaller conflict adaptation effects reflect some form of enhanced processing cuts against the grain of virtually every other model of sequential congruency effects. And while it is true that these alternative



models are quite varied, there is at least a consensus among these accounts that the sequential congruency effect is fundamentally an expression of learning (for discussion, see Egner, 2014). The sequential congruency effect, after all, reflects an adaptation of current processing by prior experience. From this standpoint then, smaller sequential congruency effects for bilinguals than monolinguals point to a disadvantage in learning for bilinguals, and are difficult to reconcile with the view that bilinguals are advantaged in cognitive control (Green, 1998). Furthermore, contrary to various claims (Grundy et al., 2017; Bialystok and Grundy, 2018), evidence reported by Grundy et al. (2017) is equivocal on the issue of whether bilinguals show diminished influence of prior congruence, prior incongruence, or both, because there was no measurement of these effects relative to a neutral trial baseline. Prevailing models attribute the sequential congruency effect to an effect of prior conflict (e.g., Botvinick et al., 2001), but there is some evidence suggesting adaptation of current trial performance may be driven more by prior congruence than by prior incongruence (Compton et al., 2012; see **Figure 2**). Whatever the underlying basis of the sequential congruency effect, the fact that Grundy et al.’s data lacked a prior neutral trial baseline, it impossible to draw any conclusions about whether bilinguals show smaller adaptation effects following congruent trials, incongruent trials, or both.

The Current Study

The present study therefore examined sequential congruency effects in monolingual and bilingual adults more closely, by comparing interference effects following congruent and incongruent trials with interference effects following neutral baseline trials. There were three alternative predictions. First, if bilingualism is associated with an advantage in learning and cognitive control (Green, 1998), bilingual adults should show a

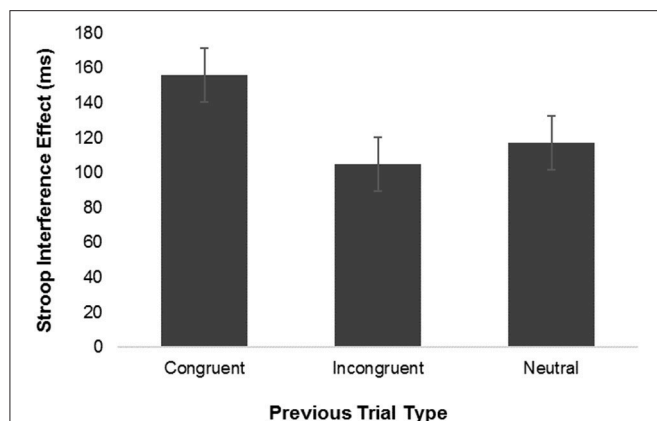


FIGURE 2 | Results of Compton et al. (2012) showing that post-incongruent interference effects are comparable in magnitude to post-neutral interference effects, whereas post-congruent effects are greater than post-neutral. Findings suggest sequential congruency effects are driven more by adaptations to prior congruent than prior incongruent trials.

larger sequential congruency effect than monolingual adults, with effects being driven by prior congruence, prior incongruence, or both. Second, if bilingualism is associated with a *disadvantage* in learning and cognitive control (Grundy et al., 2017), bilingual adults should show a smaller sequential congruency effect than monolingual adults. Finally, if bilingualism is unrelated to learning and cognitive control (Paap and Greenberg, 2013), there should be no difference in the magnitude of the sequential congruency effect for monolingual or bilingual adults.

METHODS

Participants

Seventy-three undergraduate students were recruited from Western University to participate in the study in exchange for course credit. Of these, 65 participants (26 males; mean age = 19.1 years, $SD = 2.526$) were included in the final sample. Data from seven participants were excluded owing to lower than 80% accuracy on the flanker task. Forty-four participants were bilingual (i.e., self-reported as fluent in at least two languages) and 21 were monolingual. Twenty-one bilinguals reported English as their first language, with others reporting Arabic, Chinese, Farsi, Korean, and Vietnamese. Nineteen monolinguals reported English as their first language and two reported Chinese.

Measures

Demographic Questionnaire

Participants completed an eight-item demographic questionnaire that solicited information about participant age, gender, household income, parental education, and parental occupation.

Daily Language Use Questionnaire

Following procedures used elsewhere for assessing bilingual vs. monolingual language status (e.g., Grundy et al., 2017), participants completed a 7-item questionnaire that solicited information about participant first language, knowledge of

other languages (if any), and typical day-to-day language use. Participants indicated the language(s) they typically use with family and friends, at school, when engaging with media, and when performing mental math. Responses to these items were selected from five options: “Only my first language,” “Mostly my first language,” “Both my first and other language(s),” “Mostly my other language(s),” and “Only my other language(s).”

Non-verbal Intelligence

Participants completed five computer-based measures of non-verbal intelligence including a forward digit span task, two spatial memory tasks, a pattern comparison task, and a mental rotation task.

Flanker Task (Eriksen and Eriksen, 1974)

The primary task was an Eriksen flanker task implemented in Python. Trials began with a white fixation cross centered on a black screen for 1,000 ms, followed immediately by a target stimulus embedded in flankers. On congruent trials, flankers pointed in the same direction as the target; on incongruent trials, flankers pointed in the opposite direction of the target; and on neutral trials, flankers consisted of two non-directional horizontal dashes. Stimuli were presented in the center of the screen for 1,500 ms or until a response was made. Participants were instructed to indicate as quickly and accurately as possible the direction the target stimulus. Participants responded by pressing the left- or right-most button on a five-button response box. To ensure response time was measured with the highest possible fidelity, we employed a Chronos button-box (Psychology Software Tools®) with sub-millisecond temporal resolution. The entire task consisted of 420 trials divided into four equal blocks. Participants completed the task in two two-block segments.

Procedure

All procedures were reviewed and approved by the Western University Research Ethics Board. Participants were provided with a letter of information concerning the study and provided signed written consent to their participation.

All measures were completed on a desktop computer with a 15-inch color monitor. A research assistant remained in the testing room throughout testing to oversee the protocol administration. After providing consent, participants completed the demographic and language questionnaires. Participants then completed two 120-trial blocks of the flanker task, the five computer-based measures of non-verbal intelligence, and then two final 120-trial blocks of the flanker task. Testing took on average 45 min to complete.

RESULTS

Demographics and Language Status

Most participants came from middle- or upper-class socioeconomic backgrounds with university-educated parents. Monolingual and bilingual participants had comparable socioeconomic backgrounds. Monolingual participants reported proficiency in only one language; bilingual participants

TABLE 1 | Mean response times (ms) and associated standard deviations for sequential flanker pairs in monolinguals vs. bilinguals.

Prior trial	Current trial	Language status	<i>M</i>	<i>SD</i>
Congruent	Congruent	Monolingual	404.596	43.364
		Bilingual	420.700	49.057
	Incongruent	Monolingual	481.998	44.853
		Bilingual	503.616	56.978
Incongruent	Congruent	Monolingual	419.919	43.864
		Bilingual	432.725	57.656
	Incongruent	Monolingual	483.197	41.907
		Bilingual	492.619	48.862
Neutral	Congruent	Monolingual	420.578	43.060
		Bilingual	426.600	50.903
	Incongruent	Monolingual	491.115	42.322
		Bilingual	505.066	52.893

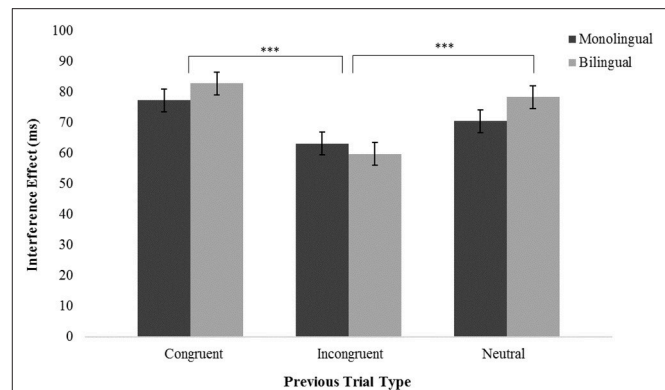
reported balanced daily use of both languages (refer to **Supplementary Table 1**).

Non-verbal Intelligence

Individual scores on each of the five non-verbal intelligence tasks were transformed into *z*-scores and summed to create an aggregate non-verbal intelligence score for each participant. Results of an independent samples *t*-test revealed no significant difference between aggregate scores of monolinguals ($M = 0.542$, $SD = 2.438$) and bilinguals ($M = -0.259$, $SD = 3.036$), $t_{(63)} = 1.056$, $p = 0.295$.

Eriksen Flanker Task and Sequential Congruency Effects

Response times across all flanker trial types are presented in **Table 1** separately for monolingual and bilingual participants. Response times were submitted to a 3-way mixed Analysis of Variance (ANOVA) with Current Trial (congruent, incongruent) and Previous Trial (congruent, incongruent) as within-subjects factors, and Group (monolingual, bilingual) as a between-subjects factor. There was an overall effect of Current Trial, $F_{(1, 63)} = 351.5$, $p < 0.001$, with response times on incongruent trials ($M = 497.4$ ms, $SD = 47.8$) significantly slower than response times on congruent trials ($M = 423.9$ ms, $SD = 48.4$). Current Trial congruency interacted with Previous Trial congruency, as reflected in a significant 2-way Current Trial \times Previous Trial interaction, $F_{(1, 63)} = 14.6$, $p < 0.001$. This interaction reflects a sequential congruency effect and was driven by fact that Current Trial interference effects were greater following congruent trials ($M = 81.1$ ms; $SD = 35.9$) than following incongruent trials ($M = 61.0$ ms; $SD = 31.3$). No other effects or interactions were significant.

**FIGURE 3** | Post-incongruent interference effects were smaller than both post-congruent and post-neutral interference effects. There was no effect of Group and no interaction. *** $p < 0.001$.

Comparison of Post-congruent and Post-incongruent Interference Effects

To examine whether sequential congruency effects are driven more by prior congruent or prior incongruent trials and whether these effects differ for monolinguals and bilinguals, we compared post-congruent and post-incongruent interference effects with a post-neutral trial baseline, shown separately for monolinguals and bilinguals in **Figure 3**. A 2-way mixed ANOVA with Previous Trial (congruent, neutral, incongruent) as a within-subjects factor and Group (monolingual, bilingual) as a between-subjects factor, revealed an effect of Previous Trial on the current trial interference effect, $F_{(2, 63)} = 17.2$, $p < 0.001$, but no effect of Group and no Previous Trial \times Group interaction. *Post-hoc* analyses indicated that current trial interference effects were smaller following incongruent compared to congruent trials ($M_D = 20.1$ ms, $p < 0.001$) and smaller following incongruent compared to neutral trials ($M_D = 14.9$ ms, $p < 0.001$). Current trial interference effects following previous congruent trials were not different than interference effects following previous neutral trials. No other effects or interactions were significant.

DISCUSSION

Monolingual and bilingual adults were administered an Eriksen flanker task. Participants exhibited a conventional congruency effect, as reflected by slower responses on incongruent compared to congruent trials, and a sequential congruency effect, as reflected by a larger congruency effect following congruent than following incongruent trials. There were however no differences in either conventional or sequential congruency effects of monolingual and bilingual adults.

The present findings contrast with evidence suggesting sequential congruency effects differ for bilingual and monolingual adults. Examination of sequential congruency effects have drawn some attention of late given mounting evidence that conventional measures of cognitive control

fail to reveal differences between monolingual and bilingual adults (Paap and Greenberg, 2013). One recent study, for example, reported *smaller* sequential congruency effects for bilingual compared to monolingual adults (Grundy et al., 2017). According to received models of the sequential congruency effect (see Egner, 2014), such group differences point to a possible learning disadvantage for bilingual vs. monolingual adults. Others, however, have interpreted smaller sequential congruency effects for bilinguals as evidence that bilinguals disengage attention from congruent and incongruent stimuli more effectively than monolinguals (Bialystok and Grundy, 2018). We tested this idea directly by measuring interference effects following congruent and incongruent trials relative to a post-neutral trial baseline. Consistent with conflict-adaptation models of the sequential congruency effect (e.g., Gratton et al., 1992; Botvinick et al., 2001; but see Compton et al., 2012), adaptation of conflict processing in the current trial was influenced more by prior incongruent trials than by prior congruent trials. That said, we found no difference in the size of sequential adaptation effects of any kind—post-incongruent or post-congruent—evidenced by monolingual vs. bilingual adults. As such, our findings are inconsistent with the view that relative to monolinguals, bilinguals more effectively disengage attention from previous stimuli or exhibit disadvantages in learning. Instead, the present findings are most consistent with the idea that monolingual and bilingual adults are indistinguishable in terms of sequential adaptation specifically and cognitive control more broadly (Paap and Greenberg, 2013).

Of course, the present study has several important limitations. One critical limitation is that there was very little in the present data that allows us to even speculate why we found no differences between monolinguals and bilinguals whereas other groups have (e.g., Grundy et al., 2017). Comparisons of monolingual and bilingual adults are always challenging because group differences in language status typically encompass differences in other factors, such as socio-economic status, immigration status, and culture, that confound the basic influence of language status. Indeed, controlling for these factors has been shown to attenuate differences between monolingual and bilinguals, at least in studies of children (see Morton and Harper, 2007). In the present case, it is unclear whether cross-study differences in sample composition could explain differences in findings, as only basic demographic variables were

measured. Similarly, we only used very rudimentary survey-based measures of daily language use to assess language status. Although these methods remain well-utilized in studies of monolinguals and bilinguals (see Grundy et al., 2017 as an example), they are ill-equipped to identify subtle differences between monolinguals and bilinguals or differences between different sorts of bilinguals (for discussion, see Baum and Titone, 2014). Clearly, advancing our understanding of language status effects on cognitive control will require adherence to higher methodological standards (for discussion, see Morton, 2015).

As a final note, our findings pertain only to possible differences between monolingual and bilingual adults. Identifying differences in adult samples has been a key challenge in bilingual advantage research and is what motivated Grundy et al. to examine sequential congruency effects more closely in the first place. Although recent large-scale studies of children also present negative evidence for the bilingual advantage hypothesis (see Dick, 2018), research in this area should remain a high priority given the wealth of previously published positive evidence and its enormous influence on the field.

ETHICS STATEMENT

Western University Non-Medical Research Ethics Board. Participants provided written voluntary consent to their participation.

AUTHOR CONTRIBUTIONS

SG and JBM designed the study and wrote the manuscript. SG collected and analyzed the data.

FUNDING

This research was supported by a Social Sciences and Humanities Research Counsel (SSHRC) grant to JBM.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02476/full#supplementary-material>

REFERENCES

- Baum, S., and Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Appl. Psycholinguist.* 35, 857–894. doi: 10.1017/S0142716414000174
- Bialystok, E., Craik, F. I., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290
- Bialystok, E., and Grundy, J. G. (2018). Science does not disengage. *Cognition*, 170, 330–333. doi: 10.1016/j.cognition.2017.10.019
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652. doi: 10.1037/0033-295X.108.3.624
- Compton, R. J., Huber, E., Levinson, A. R., and Zheutlin, A. (2012). Is “conflict adaptation” driven by conflict? Behavioral and EEG evidence for the underappreciated role of congruent trials. *Psychophysiology* 49, 583–589. doi: 10.1111/j.1469-8986.2012.01354.x
- Dick, A. (2018). No bilingual advantage for executive function: evidence from a large sample of children in the Adolescent Brain and Cognitive Development (ABCD) Study. Available online at: <https://doi.org/10.31234/osf.io/gjh95>

- Egner, T. (2014). Creatures of habit (and control): a multi-level learning perspective on the modulation of congruency effects. *Front. Psychol.* 5:1247. doi: 10.3389/fpsyg.2014.01247
- Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267
- Gratton, G., Coles, M. G., and Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *J. Exp. Psychol. Gen.* 121, 480–506. doi: 10.1037/0096-3445.121.4.480
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism Lang. Cogn.* 1, 67–81. doi: 10.1017/S1366728998000133
- Grundy, J. G., Chung-Fat-Yim, A., Friesen, D. C., Mak, L., and Bialystok, E. (2017). Sequential congruency effects reveal differences in disengagement of attention for monolingual and bilingual young adults. *Cognition* 163, 42–55. doi: 10.1016/j.cognition.2017.02.010
- Morton, J. B. (2015). Still waiting for real answers. *Cortex* 73, 352–353. doi: 10.1016/j.cortex.2015.07.010
- Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x
- Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- van Steenbergen, H., Haasnoot, E., Bocanegra, B. R., Berretty, E. W., and Hommel, B. (2015). Practice explains abolished behavioural adaptation after human dorsal anterior cingulate cortex lesions. *Sci. Rep.* 5:9721. doi: 10.1038/srep09721

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Goldsmith and Morton. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Comprehending Non-literal Language: Effects of Aging and Bilingualism

Shamala Sundaray^{1*}, Theodoros Marinis^{1,2*} and Arpita Bose¹

¹ School of Psychology and Clinical Language Sciences, University of Reading, Reading, United Kingdom, ² Department of Linguistics, University of Konstanz, Konstanz, Germany

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Antonella Sorace,
University of Edinburgh,
United Kingdom
Evy Adèle Woumans,
Ghent University, Belgium

*Correspondence:

Shamala Sundaray
shamala.sundaray@gmail.com
Theodoros Marinis
t.marinis@uni-konstanz.de

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 29 July 2018

Accepted: 29 October 2018

Published: 22 November 2018

Citation:

Sundaray S, Marinis T and Bose A
(2018) Comprehending Non-literal
Language: Effects of Aging and
Bilingualism. *Front. Psychol.* 9:2230.
doi: 10.3389/fpsyg.2018.02230

A pressing issue that the twenty-first century is facing in many parts of the developed world is a rapidly aging population. Whilst several studies have looked at aging older adults and their language use in terms of vocabulary, syntax and sentence comprehension, few have focused on the comprehension of non-literal language (i.e., pragmatic inference-making) by aging older adults, and even fewer, if any, have explored the effects of bilingualism on pragmatic inferences of non-literal language by aging older bilinguals. Thus, the present study examined the effects of age(ing) and the effects of bilingualism on aging older adults' ability to infer non-literal meaning. Four groups of participants made up of monolingual English-speaking and bilingual English-Tamil speaking young (17–23 years) and older (60–83 years) adults were tested with pragmatic tasks that included non-conventional indirect requests, conversational implicatures, conventional metaphors and novel metaphors for both accuracy and efficiency in terms of response times. While the study did not find any significant difference between monolinguals and bilinguals on pragmatic inferences, there was a significant effect of age on one type of non-literal language tested: conventional metaphors. The effect of age was present only for the monolinguals with aging older monolinguals performing less well than the young monolinguals. Aging older bilingual adults were not affected by age whilst processing conventional metaphors. This suggests a bilingual advantage in pragmatic inferences of conventional metaphors.

Keywords: aging, bilingualism, executive control, metaphors, pragmatic inferences

INTRODUCTION

Everyday communication involves not only literal language, but also the use of non-literal language, such as idioms, proverbs, metaphors, indirect requests, and conversational implicatures. To comprehend non-literal language, pragmatic inferences have to be made: the listener has to go beyond the literal meaning of the utterance and draw upon the situational context of the utterance as well as the listener's and speaker's knowledge of the world to arrive at the implied (non-literal) meaning. Pragmatic inferences are also thought to be cognitively more demanding because the listener has to both access their theory of mind to realize the speaker's communicative intentions (Champagne-Lavau and Joannette, 2009) and inhibit the literal meaning (Glucksberg et al., 2001) which becomes activated together with the implied meaning (Stewart and Heredia, 2002) during the processing of the non-literal language. Given that a great part of our daily conversations includes

non-literal utterances, it is remarkable that listeners are able to comprehend them effortlessly and in great speed in spite of the high cognitive demands. This is true of healthy young adults who are in the peak of their cognitive abilities. However, it is unclear whether this is the case also for aging older adults, whose cognitive abilities are on the decline. Moreover, it is unclear whether the aging process affects the comprehension of non-literal language in monolingual and bilingual aging older adults in the same way given recent findings that show bilinguals having a cognitive reserve (Craig et al., 2010; Bialystok et al., 2013). The present paper fills these gaps by addressing how monolingual and bilingual healthy young and aging older adults comprehend non-literal language.

The general perception has been that the language abilities of aging older adults regress with each decade. However, research has revealed that regression is not in all language areas. Healthy aging older adults may face difficulty in understanding spoken discourse, experience problems retrieving words from the mental lexicon while speaking or increasingly suffer from tip-of-the-tongue state (Gollan and Brown, 2006; Thornton and Light, 2006; Burke and Shafto, 2008). On the other hand, they have been found to have a larger vocabulary size (Burke and Shafto, 2008; Bialystok and Luk, 2012; Kavé and Halamish, 2015), and to create more complex narratives than younger adults (Thornton and Light, 2006; Burke and Shafto, 2008). Healthy aging older adults have also been reported to use “high-level vocabulary and complex syntax” (Ulatowska et al., 1998, p. 628). In addition, sentence comprehension has been reported to be intact in old age (Tyler et al., 2009).

While much research has been aimed at aging older adults’ understanding and production of vocabulary and grammatical structures at the sentential level and at times, discourse level (see Thornton and Light, 2006 for a comprehensive review), research into the pragmatic language abilities of aging older adults is comparatively rather scattered, if not impoverished. Thus, it is unclear whether or not aging older adults’ pragmatic inferential abilities, which lead to correct meaning formation of non-literal languages, regresses much like some other aspects of the aging older adults’ language.

Of the few studies that have investigated the comprehension of non-literal language by aging older adults, the focus has been on idioms (Westbury and Titone, 2011), proverbs (Nippold et al., 1997; Ulatowska et al., 1998; Uekermann et al., 2008) and metaphors (Newsome and Glucksberg, 2002; Qualls and Harris, 2003; Mashal et al., 2011). These studies, discussed below, have revealed contradictory or questionable findings in terms of the aging older adults’ pragmatic inferential abilities.

A few of the aforementioned studies point to regression in aging older adults’ pragmatic inferential abilities. Nippold et al. (1997) investigated the proverb comprehension abilities of 353 people aged between 13 and 79 years in Oregon using a Proverb Explanation Task. This task consisted of 24 proverbs which had received low familiarity ratings in Nippold and Haq (1996, cited in Nippold et al., 1997). The adolescents and adults read short stories with the proverbs appearing at the end and wrote down the meanings of the proverbs. While the study found proverb comprehension ability to decline in adults in their 60s (Nippold

et al., 1997), the stories, based on one out of the two examples provided by the authors, required connective inferences. A failure to make the connective inference could potentially impede understanding of the proverbs under study. Uekermann et al. (2008) study of 105 healthy adults, 35 of whom were aging older adults between the ages of 60 and 79, led to a similar conclusion that aging older adults were impaired in proverb comprehension. The participants in this study had to, firstly, rate the familiarity of 32 German proverbs on a five-point Likert scale, and secondly, had to determine the non-literal meaning of these proverbs from four options which varied along “degree of abstraction” and “meaningfulness” (p. 35). On the other hand, other studies did not find any regression in aging older adults’ non-literal language comprehension. Ulatowska et al. (1998), who had looked at 16 normally aging older monolingual speakers of American English in their 80s and 90s over a period of three years, found that there was no decline in proverb understanding and interpretation; instead there was an improvement for familiar proverbs and no significant changes for unfamiliar proverbs on the second testing after three years.

Metaphor comprehension too does not seem to regress with age. Aging older adults have been found to have access to metaphorical meaning (Morrone et al., 2010). Morrone et al. (2010) found their aging older participants aged 65 to 75 years making more errors and taking a longer time to reject the non-literal meaning of metaphors than the younger participants aged 21 to 30 years. This was believed to indicate that the aging older adults had access to the non-literal meanings of the metaphors. They posit that the non-literal meanings of the metaphors were likely activated and arrived at immediately, and thus needed to be inhibited; a decline in the inhibitory abilities of the aging older adults was deemed to lead to longer rejection times and more errors. Similarly, Newsome and Glucksberg (2002) found that the metaphor comprehension processes of aging older adults between the ages of 70 to 79 were not only seemingly intact, but also that the aging older adults were “as efficient as the younger adults (aged 17–21) in filtering out metaphor-irrelevant information” (p. 262). Newsome and Glucksberg presented the non-reversible metaphors and literal phrases in sentences as primes which were followed by metaphor-relevant and metaphor-irrelevant sentence probes with the last word of each prime beginning each sentence probe; participants had to judge whether the sentences made sense. Both young adults and aging older adults were better able to appreciate metaphor-relevant material after being primed by the metaphors and metaphor-irrelevant materials after being primed by the literal sentence primes.

In some instances, older adults have been found to possess superior pragmatic inferential abilities to young adults. Qualls and Harris (2003) investigated both younger (17–31 years) and older (54–73 years) African American adults’ comprehension of non-literal language. This study revealed that the older adults have better comprehension of idioms and metonyms than the younger adults. However, Qualls and Harris (2003) had a number of important confounds in their study: the answer options for metonyms included metaphors, which themselves require pragmatic inferring. In addition, the metaphor items included

both conventional and novel metaphors, both under the umbrella term of metaphors. This is problematic because processing of conventional and novel metaphors employ different cognitive mechanisms and appreciation of novel metaphors has been shown to be affected by age (Mashal et al., 2011). Lastly, the authors had included adults who were between 50 and 59 in their group of older adults. Whilst this definition of older adults is applicable to most African countries (World Health Organisation, 2002), it should not apply to African Americans who experience a longer life expectancy than and differ socially from the people in Africa; adults between 50 and 59 years of age would have better cognitive abilities than older adults, thus confounding the results.

Another important study on metaphors and aging older adults is the study by Mashal et al. (2011). Mashal et al. (2011) compared young and aging older adults in their appreciation of conventional and novel metaphoric expressions. Their first experiment, which was aimed at rating the plausibility of metaphors and literal expressions, revealed that the young adults regarded more metaphoric expressions as plausible than the aging older adults, with both groups not showing any significant difference for the plausibility rating of the literal and unrelated expressions. However, it is unclear whether the aging older adults found more of the novel metaphoric expressions as less (or more) plausible than the conventional ones; this they address in their second experiment that used different groups of young and aging older adults to examine if there was any age effect in terms of appreciating conventional versus novel metaphors. In this second experiment, the young and aging older adults had to rate the familiarity level of the 79 metaphoric expressions that were appreciated as plausible in the first experiment. Interestingly, the aging older adults rated more of the metaphoric expressions as being more familiar, appreciating them as being conventional. This was unlike the young adults who regarded the metaphoric expressions as being more novel. Expressions that were deemed as being highly novel by the young adults, were rated as being highly meaningless by the aging older adults. The study by Mashal et al. (2011) alludes to novel metaphor processing, unlike conventional metaphor processing, to be problematic in aging older adults.

The aforementioned studies, besides highlighting the contradictory findings with regard to aging older adults' non-literal language comprehension, also point to the possibility that different pragmatic inference-making strategies are employed depending upon the type of non-literal language encountered (Garcia, 2004). In addition, these studies either did not present the non-literal utterances within a situational context or presented them in texts that require connective inferences to be made. In our everyday social interactions, literal and non-literal utterances do not occur in isolation. These utterances are produced within specific contexts, and we unpack the meaning of these utterances based on these contexts. Thus, the failure to comprehend non-literal language in some of the studies looked at earlier could be due to the lack of context. To address these shortcomings, the present study focused on the comprehension of a range of non-literal language in the same groups of participants and included a situational context

for each target utterance to increase the ecological validity of the task.

All the studies mentioned above have focused on monolingual aging older adults. Although an estimated 50% or more of the world's population is either bilingual or multilingual (Grosjean, 2010), there is a lack of studies investigating bilingual aging older adults' comprehension of non-literal language. Given the current debate about whether or not bilinguals have better cognitive abilities than monolinguals and, as established earlier, the cognitive demands of pragmatic inferring during non-literal language comprehension, it is important to investigate the comprehension of non-literal language by bilingual aging older adults. In the present study, 'bilinguals' are defined based on Grosjean (2010), according to whom bilinguals are people "who use two or more languages (or dialects) in their everyday lives." (p. 4).

A number of studies have found that bilinguals have better cognitive abilities than monolinguals in terms of better executive control functions across the lifespan (Bialystok et al., 2006; Bialystok and Craik, 2010; Luk et al., 2011) and working memory (Bialystok et al., 2004). Moreover, aging adults who might otherwise succumb to dementia or neurodegenerative disease(s) earlier are now being diagnosed later due to their bilingualism (Craik et al., 2010). This has led to the hypothesis that the accrued neurocognitive differences arising from bilingual language processing over the lifespan lead to neuroplastic changes in the bilingual brain which attenuate age-related cognitive decline (Bak et al., 2014; Baum and Titone, 2014, p. 859). In addition, studies have also found that the frontal and temporal lobes, where language functions take place, are of greater volume in bilinguals than monolinguals (Olsen et al., 2015).

However, several other studies were not able to find a bilingual cognitive advantage (Paap and Greenberg, 2013; Zahodne et al., 2014; Bogulski et al., 2015). For example, in contrast to researchers who found bilinguals to be in possession of superior inhibitory abilities, Kousaie and Phillips (2012), using the Color Stroop task, did not find a bilingual advantage for inhibitory control for either their young bilinguals or their old bilinguals in comparison to their monolingual counterparts. Likewise, Colzato et al. (2008) did not find any difference between the young monolinguals and young bilinguals in the Stop Signal inhibition task, although they did find the bilinguals to be better able to maintain action goals and use them to differentiate goal-related information leading to "more pronounced reactive inhibition of irrelevant information" (p. 302). Similarly, de Bruin et al. (2015), who had controlled for a number of variables such as education, socioeconomic status, intelligence, age of acquisition and immigration status, did not find a bilingual cognitive advantage for inhibitory control in their aging older adults regardless of whether they were active or inactive bilinguals. Yet other studies have found the age of acquisition of the second language to influence the bilingual cognitive advantage; Vega-Mendoza et al. (2015) found late acquisition of second language having a positive effect on inhibition. Given that the comprehension of non-literal language is cognitively more demanding, examining monolingual and bilingual aging older

adults' comprehension of non-literal language can shed light on the debate surrounding the cognitive advantage in bilinguals.

The present study addresses the issues highlighted earlier by investigating the comprehension of non-literal utterances by monolingual and bilingual young and aging older adults. It aims to answer two research questions: (1) Is there an age effect on pragmatic inference-making? and (2) Is there a bilingual advantage in pragmatic inference-making?

This study focuses on three types of frequently occurring non-literal language: non-conventional indirect requests, conversational implicatures, and metaphors which are further divided into conventional and novel metaphors. The inclusion of different types of non-literal language will allow for greater insight to the pragmatic inferential abilities of healthy aging older adults. It is predicted that aging older adults will have pragmatic inferential abilities on par with young adults for some, but not all, non-literal language types.

Given that a number of studies have argued that L1 and L2 proficiency, age of L2 acquisition, language dominance, and L1 or L2 dominant linguistic environment that the bilinguals live in ought to be taken into account when studying bilinguals (van Hell and Poarch, 2014; Dong and Li, 2015; Mishra, 2015; Titone et al., 2015), the present study controls for age of acquisition, vocabulary knowledge, verbal fluency (see Perani et al., 2003), education, socioeconomic status, inhibition, intelligence, and processing speed, which is known to slow down with age (Salthouse, 1996) as well as verbal short-term memory and working memory, which are believed to play vital roles in discourse processing and comprehension (Hasher and Zacks, 1988).

MATERIALS AND METHODS

Participants

Seventy-three healthy adults participated in this study: 19 monolingual English-speaking young adults (mean age = 19.47, $SD = 0.7$) and 20 monolingual English-speaking aging older adults (mean age = 69.9, $SD = 6.8$) from the United Kingdom as well as 19 bilingual English-Tamil-speaking young adults (mean age = 21.02, $SD = 1.58$) and 15 bilingual English-Tamil-speaking aging older adults (mean age = 67.01, $SD = 4.39$) from Singapore. **Table 1** shows the demographic information of all four groups. All aging older adults were screened with the Mini Mental State Examination (MMSE) to rule out the onset of dementia or mild cognitive impairment; the cut-off of 27 was used based on a study conducted by O'Bryant et al. (2008) on the sensitivity of the MMSE. **Table 1** shows the groups' mean scores on the MMSE. None of the aging older adults had a score of <27 on the MMSE.

All participants completed the Language History and Use Questionnaire (LHUQ), an adaptation of the Language History Questionnaire of the Brain, Language, and Computation Lab, Penn State University (Li et al., 2006). The LHUQ consisted of 22 items which gather information such as the age of language acquisition, self-assessed language proficiency, and L1 and L2 frequency of use and code switching among other questions that elicit the participants' age, sex and socioeconomic status (SES) (years of formal education as an indication of SES). **Table 2**

provides the results of the LHUQ pertaining to age of language acquisition and language usage.

All monolingual participants were native speakers of British English. Some of the monolingual participants indicated on the LHUQ that they were aware of one or more foreign languages; these were learnt in a classroom setting around the age of 11 and later at school or after the age of 19 for work. Only two young monolinguals reported using their additional language. The use was only for half an hour out of 24 per day and not on a daily basis and therefore they were included in the monolingual group based on Grosjean's (2010) definition of bilinguals. All bilingual participants were speakers of Standard Singapore English and Standard Spoken Tamil; both English and Tamil were used in the homes of all bilingual participants. All, but four, of the young bilinguals reported that English was acquired from birth; two of the young bilinguals acquired English at the age of five, while the other two began acquiring English once in school at ages six and seven when they started school. Most of the older bilinguals began acquiring English from around the age of six, except for three older bilinguals who began learning English at the age of 12 in a formal school setting before migrating to Singapore as young adults. Given that English is widely used in public life in Singapore, all learners were exposed to English in a naturalistic environment, including these three older bilinguals. To address the potential role of age of acquisition acting as a confounding factor, it was included as a covariate in the analyses of the pragmatic tasks.

The Complex Ideational Materials Subtest (CIMS) of the Boston Diagnostic Aphasia Evaluation (BDAE) (short version) was used to test participants' auditory English sentence comprehension. The task includes a total of six pairs of yes-no questions. Each question answered correctly was awarded 1 point giving rise to a total possible score of 12. Only the aging older adults were tested in the CIMS because of the significant difference between the aging older monolinguals' and bilinguals' age of acquisition of English.

The monolingual young adults were undergraduates from the Department of Psychology, University of Reading, and received course credits for their participation. The monolingual aging older adults were recruited via the University of Reading's Aging Research Panel and were reimbursed £10 for their transport. The bilingual young adults were recruited from the National University of Singapore, the Nanyang Technological University and Ngee Ann Polytechnic in Singapore. The bilingual aging older adults were recruited through visits at temples in Singapore and through personal contacts and were given gifts of fruits and biscuits for cultural reasons.

Materials

Background Tests

To be able to control for potential confounding factors resulting from differences between the groups on verbal and non-verbal abilities as well as processing speed, a large battery of background tests was carefully selected to record the participants' lexical and semantic knowledge, and cognitive abilities, including fluid intelligence, verbal short-term memory and working memory as well as processing speed. In terms of verbal abilities, the battery

TABLE 1 | Demographic statistics of all participants.

Demographic characteristics		Monolinguals		Bilinguals	
		Young (<i>n</i> = 19)	Old (<i>n</i> = 20)	Young [#] (<i>n</i> = 18)	Old (<i>n</i> = 15)
Gender (M, F)		3, 16	10, 10	7, 11 [7, 12]	6, 9
Age	Mean (<i>SD</i>)	19.47 (0.7)	69.9 (6.8)	20.93 (1.57) [21.02 (1.58)]	67.01 (4.39)
	Min-Max	18–21	60–83	17–23	60–78
Education	Mean (<i>SD</i>)	14.97 (0.63)	14.4 (3.58)	15.83 (1.54) [15.89 (1.52)]	13.3 (3.63)
	Min-Max	14–16	10–20	14–19	7–18
MMSE	Mean (<i>SD</i>)	NA	28.8 (1.24)	NA	28.67 (1.05)
	Min-Max	NA	27–30	NA	27–30
CIMS	Mean (<i>SD</i>)	NA	11.65 (0.67)	NA	11.33 (0.98)
	Min-Max		10–12		9–12

MMSE, Mini Mental State Examination; CIMS, Complex Ideational Materials Subtest; YM, Young monolinguals; YB, Young bilinguals; OM, Old monolinguals; OB, Old bilinguals. [#]One bilingual young adult was excluded from the final analysis of the English pragmatic task because of equipment failure during this task. [] indicates data for *n* = 19 for young bilinguals.

TABLE 2 | Linguistic characteristics of participants derived from the LHUQ according to groups.

Linguistic characteristics		YM (<i>N</i> = 19)	OM (<i>N</i> = 20)	YB (<i>N</i> = 19)	OB (<i>N</i> = 15)
Age of Acquisition of English (in years)	0–5	19	20	17	2
	6–10	0	0	2	10
	11–19	0	0	0	3
Age of Acquisition of Tamil or other language (in years)	0–5	0	0	18	15
	6–10	0	0	1	0
	11–19	2	5	0	0
	20>	0	3	0	0
Conversing in English [^] (hours/day)	Mean (<i>SD</i>)	13.95 (4.2)	10.73 (3.45)	10.08 (4.19)	5.2 (3.9)
	Min-Max	2.5–16	1.5–14	3–17	0.3–12
Conversing in Tamil or other language (hours/day)	Mean (<i>SD</i>)	0.5 (0.0)	0 (0)	4.4 (3.52)	6.12 (5.48)
	Min-Max	0.5–0.5	0	0–11	0.3–16

[^] Monolingual young and older participants, who chose to state “English only” or “English All Day” when asked on the LHUQ to state the number of hours (out of 24 h per day) that they communicate with various groups of people in the languages they know, were assigned 16 and 12 h, respectively to match the total hours stated by their age cohorts.

focused on lexical and semantic rather than grammatical abilities because the experimental pragmatic tasks relied heavily on lexical and semantic information and did not have any grammatical manipulations. Of course, grammatical abilities are relevant for all tasks involving the sentence and discourse level, but the battery was already very long.

Lexical and semantic measures

The Raven's Short Vocabulary Scale (RVS), consisting of 17 words increasing in difficulty in an ascending order, was used to measure lexical knowledge. All participants had to give the meanings of the words on the list; their answers were audio recorded, and later scored with a 0 if outright wrong, 1 if partially correct and 2 when totally correct. Because vocabulary acquisition is positively related to SES (Hoff, 2003; Fernald et al., 2013), the RVS was used as a covariate together with education to control for the SES of the participants.

A Tamil vocabulary list (TVL) was created with the help of a native Singapore Tamil speaker. The TVL, like the RVS, had 17 vocabulary words and increased in its level of difficulty as the bilingual participants progressed down the list. The TVL was scored in a similar manner to the RVS.

The English Verbal Fluency (EVF) test comprised of the English Letter Fluency (ELF) task and the English Semantic Category Fluency (ESCF) task. The ELF task measures vocabulary retrieval, and together with the ESCF task, also detects neuropsychological impairments and frontal disorders (Gladsjo et al., 1999). In the ELF task, all participants were instructed to provide as many words as possible that began with the letters F, A, and S in one minute each. They were also instructed to exclude proper nouns, such as names of people and places. In the ESCF task, the participants were instructed to state as many animals as they could in one minute; they were specifically instructed to leave out breeds of the same animal (e.g., Alsatian, German Shepard, and Pomeranian all being breeds of the animal “dog”).

The Tamil Verbal Fluency (TVF) test comprised of a Tamil Letter Fluency (TLF) task and a Tamil Semantic Category Fluency (TSCF) task. In the Tamil LF task, the bilingual participants were given the Tamil letters ஸ [p^], ஐ [^], and ஶ [s^] and were similarly instructed as the English LF task, to provide as many words as possible that began with these letters in one minute each. They were also instructed to exclude proper nouns, such as names of people and places, and were provided with additional instructions where they were allowed to substitute the vowel sound [^] in the syllabic consonants, ஸ [p^] and ஶ [s^], with any of the other 11 vowels found in the Tamil alphabet.

The bilingual participants were required to complete both the EVF and the TVF. However, owing to the fact that Tamil speakers in Singapore seldom distinguish most animals by their breeds whilst speaking in Tamil, they were not instructed in the Tamil SCF to refrain from naming animals of the same breed.

Measures of cognitive abilities

The Stroop Arrow task (Blumenfeld and Marian, 2011) was used to measure participants' inhibitory abilities. The Stroop Arrow task has two stimulus dimensions: arrow direction and arrow location. These are either congruent, with right-facing arrow (or left-facing arrow) appearing on the right (or left) of the screen, or incongruent, with right-facing arrow (or left-facing arrow) appearing on the left (or right) of the screen. Participants had to respond to the direction of the arrow and ignore the location. For instance, for a right-facing arrow on the left screen, participants had to inhibit the reflex to press the key on the left for two accounts, one being the location of the arrow on screen and the other being the direction of the arrow. The Stroop Arrow task consisted of 40 congruent trials and 40 incongruent trials which were preceded by 12 practice trials. Each trial began with a black fixation cross which remained on the white screen for 800 ms and was followed by a blank white screen for 250 ms, before the stimulus appeared either on the left or the right of the white screen. The stimulus remained on screen for 1,000 ms or until a response key was hit. The trial ended with a blank screen that lasted for 500 ms, before a new trial began. The response keys were a "left-facing arrow" and a "right-facing arrow" which were overlaid on the "A" and "L" keys of a standard US keyboard, respectively. The Stroop Effect was obtained by subtracting the congruent reaction time from the incongruent reaction time for correct trials; a smaller Stroop Effect implies greater inhibitory control.

The Wechsler Adult Intelligence Scale (WAIS-III) Block Design was used to measure fluid intelligence and to control for between group differences on non-verbal IQ (de Oliveira et al., 2014). The WAIS-III Block Design required the participants to physically manipulate blocks to resemble the image shown to them. There was a total of nine images to reproduce using the blocks with five images being a two-by-two with a maximum time limit of 60 s and the remaining being a three-by-three with a maximum time limit of 120 s. Participants were scored according to the scoring system found in the WAIS-III Block Design where scores range between 4 and 7 for reproducing each image correctly within the time limit; for each image, the score obtained was inversely proportional to the time taken.

The forward and backward Digit Span (DS) tasks from the Wechsler Memory Scale (Revised) were used to test verbal short-term memory and working memory (Woods et al., 2011) because according to Hasher and Zacks (1988) they play vital role in discourse processing. In the forward digit span, participants were required to recall the digits in the order they were presented. In the backward digit span, participants were required to recall the sequence in the reverse order. Participants were given a score of one for each correct set of numbers recalled with a possible total score of 24.

The Number Comparison (NC) task (Salthouse and Babcock, 1991) was used to measure processing speed because the pragmatic task involved testing the response time. Participants had to decide if pairs of numbers were the same or different. There were 3 sets of 12 pairs of three, six and nine digits making a total of 36 items. All participants were timed separately for each set of pairs beginning with the three-digit pairs followed by the six-digit pairs and then the nine-digit pairs. Processing speed was calculated by first dividing the time taken to complete each set by the total number of items in the set (i.e., 12), and then multiplying that by the number of items that were correctly identified as being either same or different. The total number of correct items for the entire task was then divided by the total time taken for correct identification to give the processing speed (number of correct items per second).

Experimental Pragmatic Tasks

Two pragmatic tasks were created to measure a range of non-literal language and literal language: an English (EPrag) and a Tamil (TPrag) task. Each task was made up of five sets of 10 short stories to cover non-conventional indirect requests, conversational implicatures, conventional metaphors, novel metaphors, and literal utterances. Standard Singapore English is based on Standard British English; while there is no variation in the grammar, lexical differences do exist (Gupta, 2010, 2012; Leimgruber, 2011). Vocabulary that may have different meanings in the two varieties of English were avoided in the stories. Similarly, all stories were created to be culturally neutral, that is, the situational contexts were applicable to both Singapore and the United Kingdom. The English conventional metaphors were selected from a familiarity rating list administered to nine healthy aging monolingual English speakers aged 60 years and above in the United Kingdom and six healthy aging bilingual English-Tamil speakers aged 60 years and above in Singapore. Similarly, the Tamil conventional metaphors were selected from a familiarity rating list administered to the same group of aging bilingual English-Tamil speakers. Participants completed three practice trials before starting on the actual task.

Each trial consisted of a short dialog by or between a male and a female character that were accompanied by a line drawing to create a story. Participants heard the target utterances at the end of these short dialogs. Each story started with the narrator providing the setting (e.g., "At a party") and background (e.g., "Jill is at a party.") and ended with a multiple-choice comprehension question in the format of "What will <story character's name or gender> say or do next?". Participants heard the narrator reading out the questions and the four options as

well as seeing the questions and options displayed on the screen below the line drawings. The questions and options for EPrag were typed onto the slide as text, whereas the questions and their answer options for Tamil had to be handwritten and uploaded as images because the experiment software did not support the Tamil script font. The complete story board for the EPrag task can be found in the **Supplementary Material**.

Each option can be categorized under one of four types: (a) inferred meaning, (b) literal meaning, (c) possible, but wrong reaction and (d) wrong answer. There were two “wrong answers” for the literal category as there are no inferred meanings for the literal target utterances. Participants pressed the corresponding key on the keyboard to record their answers, after which a new slide with the words “Next story?” appeared on the screen. Pressing the space bar then brought the participants to the next slide which had a fixation cross for 250 ms before a new story begun.

The dependent variables—accuracy scores and time taken to respond (TTR) (in seconds)—were recorded for each of the non-literal language types (i.e., non-conventional indirect requests, conversational implicatures, conventional metaphors, and novel metaphors) and literal utterances. The TTR measure was calculated only for correct responses for each non-literal and literal language type tested.

Procedure

The Pragmatic tasks were run using E-prime 2.0 Professional on an Acer Aspire 4820T laptop with an Intel® Core™ i5 processor 4.30 M and a 14.0-inch HD LED LCD screen. Participants were tested individually in separate sessions. The bilingual participants completed the English and Tamil tasks in separate sessions. The bilinguals' testing sessions were counterbalanced by language; the English and Tamil sessions were spaced apart by two to three weeks.

Data Analyses

The study has set out to answer two research questions: (1) “Is there an age effect on pragmatic inference-making?” and (2) “Is there a bilingual advantage in pragmatic inference-making?” Language Group (monolingual, bilingual) and Age (young, old) were the independent variables for this study.

The age of acquisition of English and Tamil and CIMS scores were analyzed with a Mann-Whitney test. Age, education and the variables arising from the background tests were analyzed with a two-way univariate analysis of variance (ANOVA) with Age and Language Group as factors. The MMSE was analyzed with a one-way ANOVA with Language Group as the independent variable. Variables arising from the Tamil background tests were analyzed with a one-way ANOVA with Age as the independent variable.

Each of the pragmatic tasks (the EPrag and TPrag tasks) had five dependent variables for the accuracy and five for the TTRs, corresponding to the five pragmatic conditions (non-conventional indirect requests, conversational implicatures, conventional metaphors, novel metaphors and literal utterances).

For the EPrag task, a two-way multivariate analysis of covariance (MANCOVA) was used to test the effects of Age and Language Group on the EPrag accuracy scores (i.e., arising

from the non-conventional indirect requests, conversational implicatures, conventional metaphors, novel metaphors and literal utterances) whilst controlling for potential effects of socioeconomic status, verbal IQ, education, inhibition, verbal short-term memory and working memory as well as age of acquisition of English that may affect the participants' inferential abilities. A similar analysis was conducted on the EPrag TTRs with Number Comparison as an additional covariate to control for the differing processing speed of the groups. Planned pairwise comparisons were conducted to compare differences between young and aging older adults, and monolinguals and bilinguals for each pragmatic condition separately.

For the TPrag task, a one-way MANCOVA was run to test for effects of Age on the TPrag accuracy scores (arising from the non-conventional indirect requests, conversational implicatures, conventional metaphors, novel metaphors, and literal utterances) with Education, Tamil Vocabulary List, Stroop Arrow, Block Design, Tamil Verbal Fluency, Age of Acquisition of Tamil and Digit Span as covariates. The covariates were included to control for socioeconomic status, verbal IQ, differing educational levels between groups, inhibition, verbal short-term memory, and working memory that can potentially affect inferential abilities, and to reduce error variances. Similarly, a one-way MANCOVA was conducted on the TPrag TTRs with Number Comparison as an additional covariate to control for differing processing speed of the groups. Finally, planned pairwise comparisons were conducted to compare differences between young and aging older bilingual adults for each pragmatic condition.

RESULTS

Demographics

There was no significant difference between the monolinguals and bilinguals for Age in Years [$F(1, 68) = 0.523, p = 0.472, d = 0.2, 1 - \beta = 0.12$]¹ and for Years of Education [$F(1, 68) = 0.037, p = 0.849, d = 0.06, 1 - \beta = 0.06$]. As expected, there was a significant difference in Age in Years between the young and older adults [$F(1, 68) = 2353.2, p < 0.001, d = 11.8, 1 - \beta = 1.0$] with a significant interaction between Age and Language Group [$F(1, 68) = 4.776, p = 0.032, d = 0.5, 1 - \beta = 0.6$]: Age in Years was different between young and aging older monolinguals [$F(1, 37) = 1036.4, p < 0.001, d = 10.7, 1 - \beta = 1.0$] and between young and aging older bilinguals [$F(1, 31) = 1724.3, p < 0.001, d = 14.8, 1 - \beta = 1.0$]. However, there was also a significant difference between young and older adults in Years of Education [$F(1, 68) = 6.14, p = 0.016, d = 0.6, 1 - \beta = 0.71$]. There was no significant interaction between Age and Language Group for Years of Education [$F(1, 68) = 2.443, p = 0.123, d = 0.4, 1 - \beta = 0.36$]. The difference in education between young and older adults is due to differences in years of education across generations, especially in Singapore, and was impossible to control for due to changes in the society.

¹Effect size and power for all analyses were calculated using G*Power (Version 3.1.9.2) and Lenhard and Lenhard (2016) (https://www.psychometrika.de/effect_size).

Hence, Years of Education was used as a covariate to address this confounding factor.

There was no significant difference on the MMSE between the monolingual and bilingual aging older adults [$F(1, 33) = 0.113$, $p = 0.739$, $d = 0.1$, $1 - \beta = 0.06$].

Mann-Whitney tests comparing the age of acquisition for English and Tamil between the groups showed a significant difference in the age of acquisition of English between the aging older monolinguals and bilinguals ($U = 0.000$, $p < 0.001$, $r = 0.9$, $1 - \beta = 1.0$), and the young and aging older bilinguals ($U = 19$, $p < 0.001$, $r = 0.8$, $1 - \beta = 1.0$). There was no significant difference between the young monolinguals and bilinguals ($U = 123.5$, $p = 0.15$, $r = 0.4$, $1 - \beta = 0.89$). As for the age of acquisition of Tamil, there was no significance difference between the young and aging older bilinguals ($U = 141$, $p = 0.973$, $r = 0.02$, $1 - \beta = 0.05$).

The Mann-Whitney test comparing the CIMS scores did not show any significant difference between the aging older monolinguals and bilinguals ($U = 125$, $p = 0.354$, $r = 0.17$, $1 - \beta = 0.23$).

Background Tests

Table 3 shows the results from the background tests.

Lexical and Semantic Measures

In terms of vocabulary knowledge in English (RVS), there was a significant main effect of Language Group [$F(1, 68) = 4.188$, $p < 0.05$, $d = 0.5$, $1 - \beta = 0.55$], but no significant main effect of Age [$F(1, 68) = 1.847$, $p > 0.05$, $d = 0.3$, $1 - \beta = 0.28$]. There was a significant interaction effect between Language Group and Age [$F(1, 68) = 4.141$, $p < 0.05$, $d = 0.5$, $1 - \beta = 0.54$]. Follow-up simple effects showed that aging older monolinguals had better vocabulary knowledge than young monolinguals [$F(1, 68) = 6.309$, $p < 0.05$, $d = 0.6$, $1 - \beta = 0.72$] and aging older bilinguals [$F(1, 68) = 8.026$, $p < 0.01$, $d = 0.7$, $1 - \beta = 0.82$]. There were no significant differences in the vocabulary knowledge of the young monolinguals and bilinguals [$F(1, 68) = 0.000$, $p > 0.05$, $d = 0.00$, $1 - \beta = 0.05$], and between young bilinguals and aging older bilinguals [$F(1, 68) = 0.210$, $p > 0.05$, $d = 0.1$, $1 - \beta = 0.074$]. In terms of vocabulary knowledge in Tamil (TVL), the young bilinguals and aging older bilinguals did not differ [$F(1, 32) = 0.696$, $p > 0.05$, $d = 0.3$, $1 - \beta = 0.13$].

The two-way ANOVA on the English Verbal Fluency test (EVF) showed a significant main effect of Language Group [$F(1, 68) = 5.266$, $p < 0.05$, $d = 0.6$, $1 - \beta = 0.64$], but no significant main effect of Age [$F(1, 68) = 1.852$, $p > 0.05$, $d = 0.3$, $1 - \beta = 0.29$]. There was a significant interaction effect between Language Group and Age [$F(1, 68) = 9.208$, $p < 0.01$, $d = 0.7$, $1 - \beta = 0.87$]. Both aging older monolinguals [$F(1, 68) = 13.685$, $p < 0.001$, $d = 0.9$, $1 - \beta = 0.96$] and young bilinguals [$F(1, 68) = 8.886$, $p < 0.01$, $d = 0.7$, $1 - \beta = 0.86$] had better verbal fluency than aging older bilinguals. There were no significant differences between the young monolinguals and aging older monolinguals [$F(1, 68) = 1.534$, $p > 0.05$, $d = 0.3$, $1 - \beta = 0.24$], and between the young monolinguals and young bilinguals [$F(1, 68) = 0.284$, $p > 0.05$, $d = 0.1$, $1 - \beta = 0.083$].

The young bilinguals and aging older bilinguals did not differ in the Tamil Verbal Fluency test (TVF) [$F(1, 32) = 0.055$, $p > 0.05$, $d = 0.09$, $1 - \beta = 0.057$].

Measures of Cognitive Abilities

A two-way ANOVA showed no significant main effect of Language Group on the Stroop Effect [$F(1, 68) = 0.116$, $p > 0.05$, $d = 0.09$, $1 - \beta = 0.07$] and no significant interaction of Language Group and Age [$F(1, 68) = 2.243$, $p > 0.05$, $d = 0.36$, $1 - \beta = 0.33$]. However, there was a highly significant main effect of Age on the Stroop Effect [$F(1, 68) = 24.15$, $p < 0.001$, $d = 1.2$, $1 - \beta = 0.999$] indicating that young adults had better inhibitory abilities than aging older adults.

The Kruskal-Wallis test showed a highly significant effect of Age on the Block Design [$H(1) = 17.985$, $p < 0.001$]. There was no significant effect of Language Group [$H(1) = 1.968$, $p > 0.05$]. Follow-up Mann-Whitney tests indicated that the young bilinguals had higher scores on the Block Design than the aging older bilinguals ($U = 2.0$, $p < 0.001$, $d = 2.1$). There was no difference between the young and aging older monolinguals ($U = 148.5$, $p > 0.025$, $d = 0.38$). (A Bonferroni correction was applied, and all effects are reported at a 0.025 level of significance).

There was a significant main effect of Language Group on the Digit Span [$F(1, 68) = 9.731$, $p < 0.01$, $d = 0.76$, $1 - \beta = 0.89$], but no significant main effect of Age [$F(1, 68) = 3.598$, $p > 0.05$, $d = 0.49$, $1 - \beta = 0.48$]. There was a significant interaction effect between Language Group and Age [$F(1, 68) = 14.001$, $p < 0.001$, $d = 0.91$, $1 - \beta = 0.97$]. Follow-up simple effects analyses showed the young bilinguals had a significantly better verbal short-term memory and working memory than young monolinguals [$F(1, 68) = 24.461$, $p < 0.001$, $d = 1.2$, $1 - \beta = 0.999$], and aging older bilinguals [$F(1, 68) = 14.623$, $p < 0.001$, $d = 0.93$, $1 - \beta = 0.97$]. There were no differences between young monolinguals and aging older monolinguals [$F(1, 68) = 1.864$, $p > 0.05$, $d = 0.33$, $1 - \beta = 0.29$], and between aging older monolinguals and bilinguals [$F(1, 68) = 0.187$, $p > 0.05$, $d = 0.11$, $1 - \beta = 0.08$].

There was no significant main effect of Language Group [$F(1, 68) = 2.173$, $p > 0.05$, $d = 0.36$, $1 - \beta = 0.32$] on the Number Comparison and no significant interaction effect between Language Group and Age [$F(1, 68) = 0.878$, $p > 0.05$, $d = 0.23$, $1 - \beta = 0.16$]. However, there was a highly significant main effect of Age [$F(1, 68) = 25.206$, $p < 0.001$, $d = 1.2$, $1 - \beta = 0.999$], indicating that the young adults had better processing speed than the older adults.

Pragmatic Tasks

EPrag Accuracy Scores and TTRs

Figure 1 shows the participants' accuracy scores for the English Pragmatic (EPrag) task.

The MANCOVA on the accuracy scores showed a significant effect of Age on the combined dependent variables (non-conventional indirect requests, conversational implicatures, conventional metaphors, novel metaphors and literal utterances) [$\lambda = 0.779$, $F(5, 57) = 3.225$, $p < 0.05$, $d = 1.1$], indicating differences between young and aging older participants. There

TABLE 3 | Untransformed mean scores (*SD*) of all participants for the background tests.

		Monolinguals		Bilinguals		Main effect of Age	Main effect of Language Group	Age x Language Group interaction
		Young (N = 19)	Old (N = 20)	Young (N = 18)	Old (N = 15)			
RVS	Mean (<i>SD</i>)	15.79 (3.05)	19.20 (4.20)	15.78 (2.88)	15.1 (6.43)	ns	*	*
	Min-Max	10–21	12–28	10–20	5–24			
TVL [#]	Mean (<i>SD</i>)	NA	NA	21.94 (2.91)	20.80 (5.03)	ns	NA	NA
	Min-Max	NA	NA	16–26	15–31			
ELF	Mean (<i>SD</i>)	48.53 (10.93)	57.90 (19.18)	53.39 (10.23)	42.53 (17.79)	ns	*	**
	Min-Max	30–74	25–90	34–69	11–83			
ESCF	Mean (<i>SD</i>)	24.3 (3.96)	24.45 (5.48)	24.11 (4.79)	19.13 (4.72)	ns	NA	NA
	Min-Max	18–32	15–36	16–35	11–27			
TLF [#]	Mean (<i>SD</i>)	NA	NA	46.05 (10.92)	45.27 (13.18)	ns	NA	NA
	Min-Max	NA	NA	29–74	21–69			
TSCF [#]	Mean (<i>SD</i>)	NA	NA	17.42 (2.85)	17.07 (5.78)	ns	NA	NA
	Min-Max	NA	NA	12–24	8–28			
SA	Mean (<i>SD</i>)	35.59 (33.44)	64.33 (40.37)	25.86 (25.52)	79.79 (40.89)	***	ns	ns
	Min-Max	–23.09–106.75	9.05–188.4	–4.35–94.01	31.1–180.66			
BD	Mean (<i>SD</i>)	38.84 (7.32)	34.95 (9.89)	41.56 (6.11)	22.20 (6.27)	***	ns	–
	Min-Max	23–50	16–50	29–49	11–32			
DS	Mean (<i>SD</i>)	15.32 (2.89)	16.65 (2.89)	20.28 (2.7)	16.20 (3.78)	ns	**	***
	Min-Max	11–21	11–21	14–24	10–22			
NC	Mean (<i>SD</i>)	0.28 (0.1)	0.21 (0.06)	0.26 (0.07)	0.19 (0.08)	***	ns	ns
	Min-Max	0.17–0.61	0.12–0.37	0.18–0.45	0.09–0.36			

RVS, Raven's Vocabulary Scale; TVL, Tamil Vocabulary List; ELF, English Letter Fluency; ESCF, English Semantic Category Fluency; TLF, Tamil Letter Fluency; TSCF, Tamil Semantic Category Fluency; SA, Stroop Arrow; BD, Block Design; DS, Digit Span; NC, Number comparison; EVF, English Verbal Fluency; TVF, Tamil Verbal Fluency; YM, Young Monolinguals; YB, Young Bilinguals; OM, Old Monolinguals; OB, Old Bilinguals. [#]The Tamil background tasks were analyzed with N = 19 for young bilinguals. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

was no significant effect of Language Group on the combined dependent variables [$\lambda = 0.948$, $F(5, 57) = 0.626$, $p > 0.05$, $d = 0.5$], indicating that monolinguals and bilinguals performed alike, and no significant interaction effect between Language Group and Age [$\lambda = 0.935$, $F(5, 57) = 0.793$, $p > 0.05$, $d = 0.5$], indicating that monolinguals and bilinguals show the same pattern of performance. The planned comparisons for each non-literal condition separately showed that young monolinguals were significantly better than aging older monolinguals at conventional metaphors [$F(1, 31) = 9.06$, $p = 0.005$, $d = 1.1$, $1 - \beta = 0.9$]. There was no significant difference between young bilinguals and aging older bilinguals for conventional metaphors [$F(1, 24) = 2.072$, $p > 0.05$, $d = 0.6$, $1 - \beta = 0.37$].

Figure 2 shows the participants' TTRs for the English Pragmatic (EPrag) task.

The MANCOVA on the TTRs showed a significant main effect of Age on the combined TTRs for the non-conventional indirect requests, conversational implicatures, conventional metaphors, novel metaphors and literal utterances [$\lambda = 0.746$, $F(5, 56) = 3.818$, $p < 0.01$, $d = 1.2$], indicating differences between young and aging older participants. There was no significant main effect of Language Group on the combined TTRs [$\lambda = 0.911$, $F(5, 56) = 1.096$, $p > 0.05$, $d = 0.6$], indicating that monolinguals and bilinguals performed alike. There was no significant interaction effect between Language Group and Age

[$\lambda = 0.963$, $F(5, 56) = 0.435$, $p > 0.05$, $d = 0.4$], indicating that monolinguals and bilinguals showed the same pattern of performance. The planned comparisons for each non-literal condition separately showed that young monolinguals were significantly faster than aging older monolinguals in inferring conventional metaphors [$F(1, 30) = 7.074$, $p = 0.012$, $d = 1.0$, $1 - \beta = 0.84$], whilst there was no significant difference between the young and aging older bilinguals [$F(1, 23) = 2.034$, $p > 0.05$, $d = 0.6$, $1 - \beta = 0.37$]. (A Bonferroni correction was applied, and the effects are reported at a 0.0125 level of significance). There were no significant differences between the young monolinguals and aging older monolinguals for the literal utterances TTR [$F(1, 30) = 1.401$, $p > 0.05$, $d = 0.4$, $1 - \beta = 0.26$], conversational implicatures TTR [$F(1, 30) = 5.112$, $p > 0.05$, $d = 0.8$, $1 - \beta = 0.7$] and novel metaphors TTR [$F(1, 30) = 6.195$, $p > 0.01$, $d = 0.9$, $1 - \beta = 0.78$]. Likewise, there were no significant differences between the young bilinguals and aging older bilinguals for literal utterances TTR [$F(1, 23) = 2.873$, $p > 0.05$, $d = 0.7$, $1 - \beta = 0.49$], conversational implicatures TTR [$F(1, 23) = 0.716$, $p > 0.05$, $d = 0.4$, $1 - \beta = 0.16$], and novel metaphors TTR [$F(1, 23) = 3.634$, $p > 0.05$, $d = 0.8$, $1 - \beta = 0.59$]. Planned comparison was not done for non-conventional indirect requests TTR because the independent one-way ANCOVA did not show a significant main effect of Age [$F(1, 60) = 4.755$, $p > 0.01$, $d = 0.6$, $1 - \beta = 0.65$].

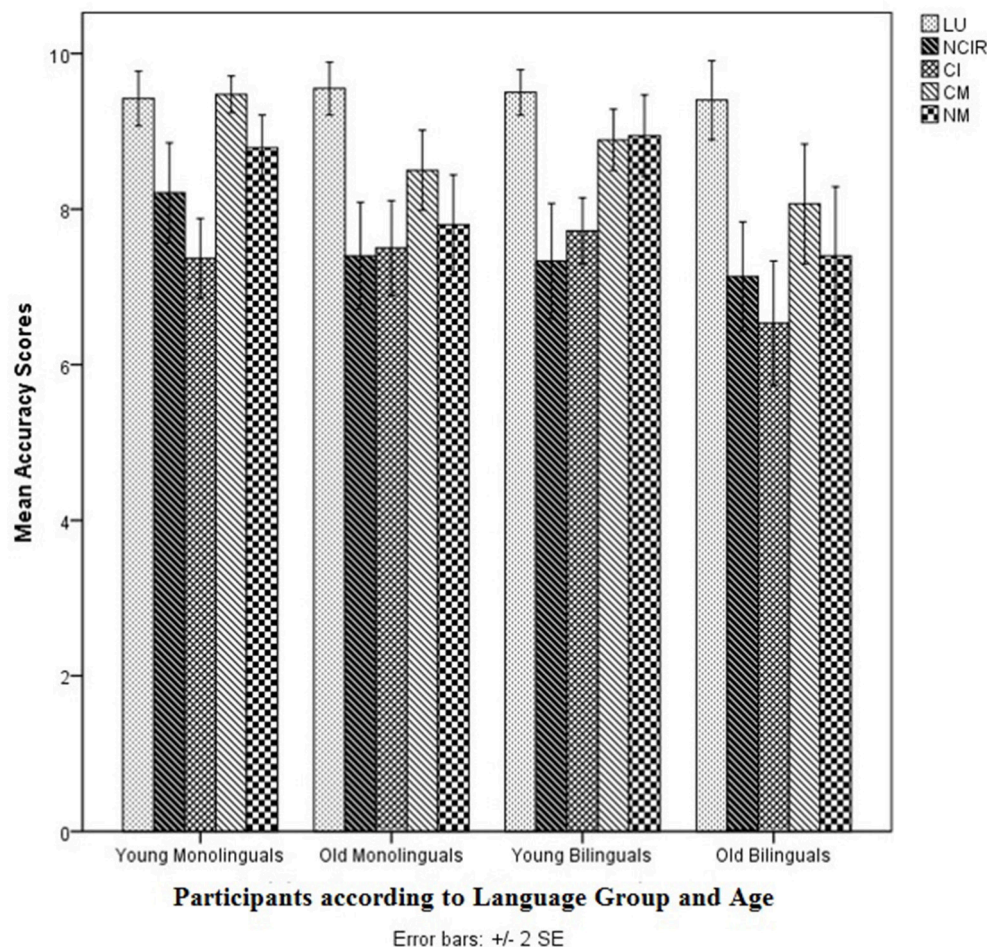


FIGURE 1 | Mean accuracy scores of all participants ($n = 72$) in the EPrag task. LU, Literal Utterances; NCIR, Non-conventional Indirect Requests; CI, Conversational Implicatures; CM, Conventional Metaphors; NM, Novel Metaphors.

TPrag Task Accuracy Scores and TTRs

Figures 3 and 4 show the accuracy scores and TTRs for the TPrag task.

The MANCOVA on the accuracy scores showed no significant main effect of Age on the combined accuracy scores [$\lambda = 0.873$, $F(5, 21) = 0.609$, $p > 0.05$, $d = 0.8$]. Likewise, the MANCOVA on the TTRs did not show a significant main effect of Age on the combined TTRs [$\lambda = 0.635$, $F(5, 20) = 2.3$, $p > 0.05$, $d = 1.5$].

DISCUSSION

Everyday communication comprises of an extensive use of non-literal language, such as idioms, proverbs, metaphors, indirect requests, and conversational implicatures. Although the developed world is facing a rapidly aging population, research on the comprehension of non-literal language in aging older adults is limited and is based mainly on monolingual speakers. Whilst some studies found that aging older adults are able to access the non-literal meanings of metaphors (Ulatowska et al., 1998; Newsome and Glucksberg, 2002; Qualls

and Harris, 2003; Morrone et al., 2010) and suggested that aging older adults are “as efficient” as younger adults when processing metaphors (Newsome and Glucksberg, 2002), some other studies demonstrated an age-related decline in non-literal language comprehension (Nippold et al., 1997; Uekermann et al., 2008). The differences in the findings of these studies could be related to the differences in the methodologies used, the variability in the participant populations, and the designs of the studies. Importantly, although context plays a key role in the comprehension of non-literal language, previous studies reviewed either did not present non-literal utterances within a situational context or presented them in texts that required connective inferences.

The current study aimed to fill the gap in the literature of aging older adults’ pragmatic inferential abilities using non-literal utterances embedded in situational contexts. It also sought to investigate if there was a bilingual advantage in pragmatic inference-making. Young and older monolinguals and bilinguals underwent a battery of background tests to measure their vocabulary knowledge, non-verbal IQ, verbal fluency,

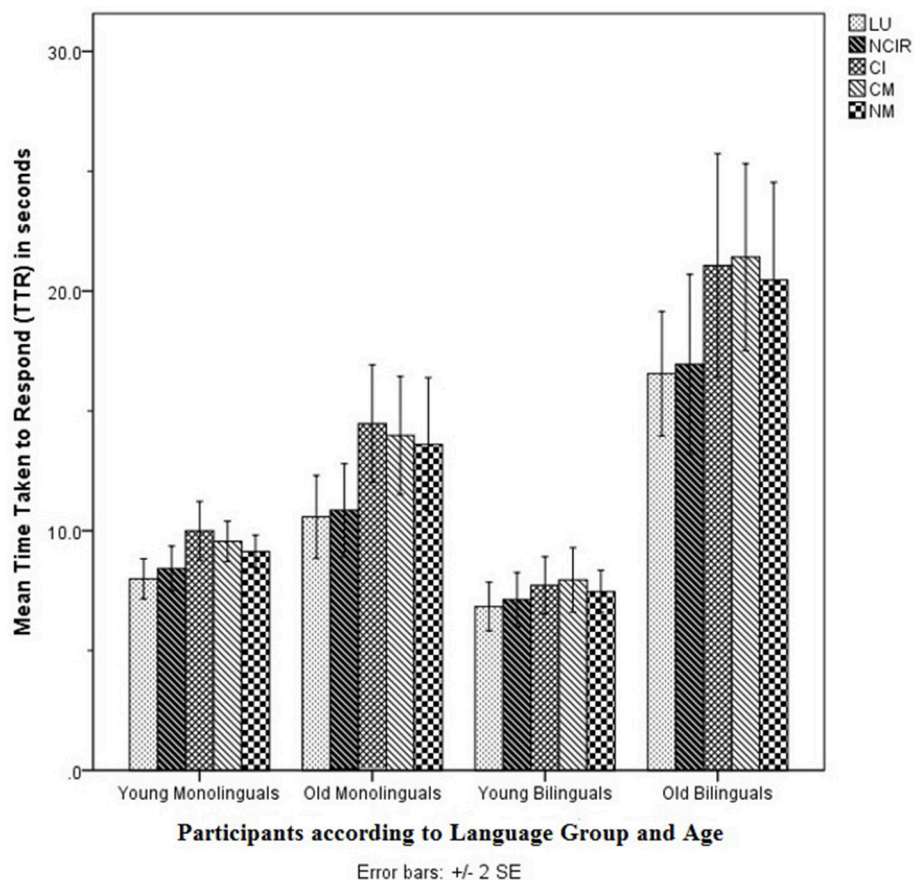


FIGURE 2 | Mean time taken to respond (TTR) of all participants ($n = 72$) in the EPrag task. LU, Literal Utterances; NCIR, Non-conventional Indirect Requests; CI, Conversational Implicatures; CM, Conventional Metaphors; NM, Novel Metaphors.

inhibition, verbal short-term memory and working memory, and processing speed as well as completed a language use and history questionnaire to provide information such as education, age of acquisition of English and language usage. To address their pragmatic inferential abilities, participants completed an English pragmatic task that had the target literal and non-literal utterances presented in context-based vignettes that were culturally neutral. The bilinguals were, in addition, tested with a Tamil pragmatic task. Participants were tested for both accuracy and response time. After controlling for education, vocabulary knowledge, non-verbal IQ, verbal fluency, inhibition, verbal short-term memory and working memory, age of acquisition of English and processing speed, a clear effect of age on the comprehension of English conventional metaphors emerged. Planned comparisons showed that aging older monolinguals were less accurate and slower than young monolinguals on the comprehension of English conventional metaphors. Aging older bilinguals, on the other hand, were as accurate and efficient as young bilinguals on the comprehension of English conventional metaphors. Moreover, although there was no effect of Language Group (i.e., bilingualism) for any of the non-literal language types tested, this effect of age found for the monolinguals was not found

for the bilinguals for any of the non-literal language types tested in the study, be it in English or Tamil.

Understanding Non-literal Language as We Age

In the present study, we found an age-related decline in conventional metaphor comprehension, but only for the monolinguals. Not only were the aging older monolinguals less accurate than the young monolinguals in comprehending conventional metaphors, they were also much slower when processing conventional metaphors. Past literature supports the present findings that monolingual aging older adults experience an age-related decline in non-literal language comprehension (Nippold et al., 1997; Uekermann et al., 2008). It is worth noting here that the conventional metaphors were selected based on the metaphor familiarity rating list completed by a sample of both monolingual and bilingual aging older adults, but not by the younger groups. Hence, older participants would have been guaranteed familiar with the conventional metaphors, more so than the young participants. In spite of this advantage, the aging older monolinguals were significantly less accurate and slower in inferring the metaphorical meaning of the utterances.

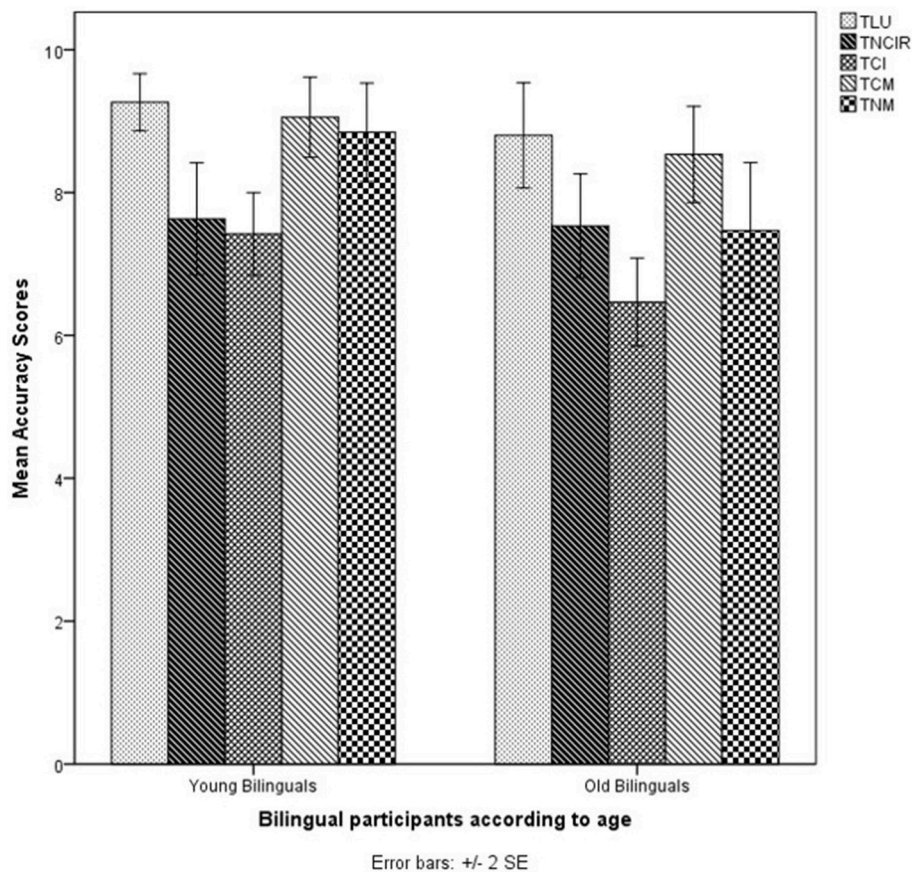


FIGURE 3 | Mean accuracy scores of all bilingual participants ($n = 34$) in the TPrag task. TLU, Tamil Literal Utterances; TNCIR, Tamil Non-conventional Indirect Requests; TCI, Tamil Conversational Implicatures; TCM, Tamil Conventional Metaphors; TNM, Tamil Novel Metaphors.

On the other hand, the aging older bilinguals were as accurate as the young bilinguals in terms of understanding English and Tamil metaphors (as well as the other non-literal language types tested); this is in line with studies showing that aging older adults are able to access the non-literal meanings of metaphors (Ulatowska et al., 1998; Newsome and Glucksberg, 2002; Qualls and Harris, 2003; Morrone et al., 2010). In addition, the aging older bilinguals were not significantly slower than the young bilinguals at arriving at the correct meaning of the English and Tamil metaphors. These findings suggest that aging older adults are “as efficient” as young adults when processing metaphors (Newsome and Glucksberg, 2002).

We now know that pragmatic inference-making does slow down with aging, even with processing speed attrition, cognition and other factors having been taken into account, but not for all non-literal language types and not for bilinguals.

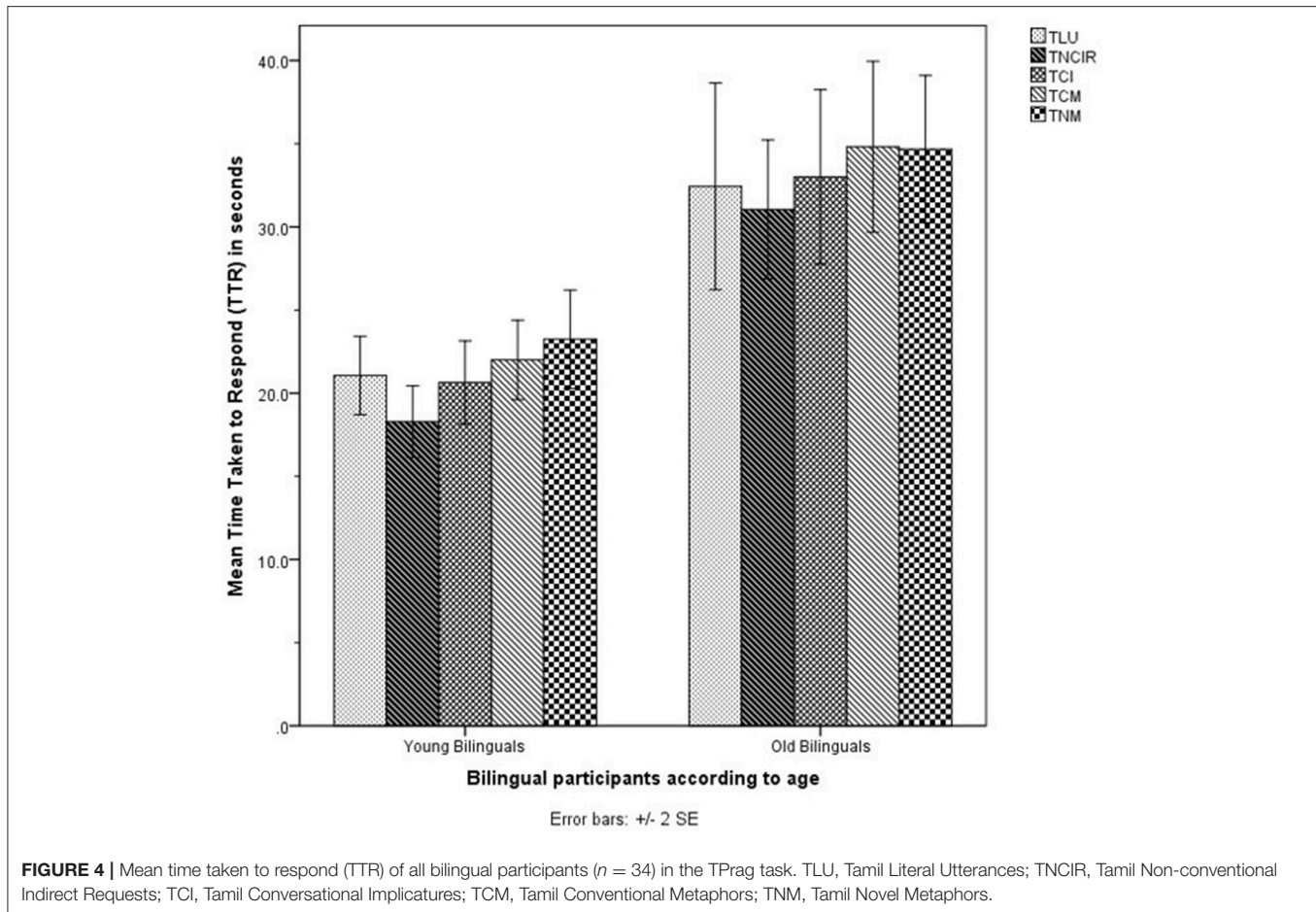
Bilinguals and Pragmatic Inference-Making

The present study did not find any significant differences between the monolinguals and bilinguals in terms of pragmatic inference-making. Of the very few studies that investigated the pragmatic inference-making abilities of bilinguals, one found no bilingualism effect on conversational implicatures for L2

learners and native speakers of English (Manowong, 2011), while another found a slightly higher correlation between linguistic comprehension and pragmatic comprehension of both indirect requests and conversational implicatures for L2 learners of English with higher English language proficiency than L2 learners with lower English language proficiency (Garcia, 2004).

In the present study, the bilinguals used the English language on a daily basis and had self-assessed their English language proficiency in speaking and listening as being between “Good” to “Native-like.” The bilinguals in the present study were not disadvantaged by their “non-native speaker” status unlike the L2 learners of English in Garcia’s (2004) study and did not display a significant disadvantage in discourse processing as seen by their performance in both the literal and non-literal language types tested in the pragmatic tasks.

Although there was no overall significant effect of bilingualism on pragmatic inference-making, the findings of the present study point to a bilingual advantage when it comes to comprehending English conventional metaphors; aging older bilinguals’ conventional metaphor processing was not affected by age unlike the aging older monolinguals’. As established earlier, pragmatic inferences require higher order cognitive skills (Champagne-Lavau and Joannette, 2009), and a number of studies



have shown bilingualism attenuating cognitive decline associated with aging (Luk et al., 2011) and bilinguals possessing superior cognitive abilities than monolinguals even as they get older (Bialystok et al., 2006). Thus, it should come as no surprise that aging older bilinguals were not affected by age whilst processing conventional metaphors unlike their monolingual counterparts.

The sample size of the present study was small, which is one of the limitations of the study. A second limitation is that the study focused only on comprehension and did not measure the participants' production of non-literal language. Future research can compare the comprehension with the production of non-literal language by a larger sample of aging older adults and examine the effects of Language Group. This would provide a complete picture of both comprehension and production of non-literal language.

CONCLUSION

The present study examined the effects of age(ing) and the effects of bilingualism on pragmatic inferences by monolingual and bilingual young and older adults. The present study has controlled for a large number of variables that can affect pragmatic inference-making. These variables include the participants' vocabulary knowledge, non-verbal IQ, education,

socioeconomic status, age of acquisition of English, inhibition, verbal short-term memory and working memory, verbal fluency, and processing speed. On top of this, the young and aging older bilinguals were tested in both their languages, English and Tamil. Regardless of language, aging older bilinguals were not affected by age whilst processing literal and non-literal language. This is in direct contrast to aging older monolinguals who displayed an age-related disadvantage when confronted with conventional metaphors. This suggests a bilingual advantage in pragmatic inferences of conventional metaphors.

ETHICS STATEMENT

This study was reviewed by the School of Psychology and Clinical Language Sciences' Ethics Committee and the University Research Ethics Committee (University of Reading) and was given a favorable ethical opinion for conduct.

DATASETS ARE AVAILABLE ON REQUEST

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

SS conceived and designed the study together with TM and AB. SS created the pragmatic tasks, collected, analysed, and interpreted the data. SS together with TM and AB wrote the paper.

ACKNOWLEDGMENTS

SS would like to thank Manicandane Annamalai for his voiceovers for the English and Tamil Pragmatic tasks and

other assistance rendered during the data collection. The study reported in this paper was part of the first author's Ph.D. thesis submitted to the University of Reading. SS would like to thank her supervisors for their guidance throughout the research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02230/full#supplementary-material>

REFERENCES

- Bak, T. H., Nissan, J. J., Allerhand, M. M., and Deary, I. J. (2014). Does bilingualism influence cognitive aging? *Ann. Neurol.* 75, 959–963. doi: 10.1002/ana.24158
- Baum, S., and Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Appl. Psycholinguist.* 35, 857–894. doi: 10.1017/S0142716414000174
- Bialystok, E., and Craik, F. I. M. (2010). Cognitive and linguistic processing in the bilingual mind. *Curr. Dir. Psychol. Sci.* 19, 19–23. doi: 10.1177/0963721409358571
- Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290
- Bialystok, E., Craik, F. I. M., and Luk, G. (2013). Bilingualism: consequences for mind and brain. *Trends Cogn. Sci.* 16, 240–250. doi: 10.1016/j.tics.2012.03.001
- Bialystok, E., Craik, F. I. M., and Ryan, J. (2006). Executive control in a modified antisaccade task: effects of aging and bilingualism. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1341–1354. doi: 10.1037/0278-7393.32.6.1341
- Bialystok, E., and Luk, G. (2012). Receptive vocabulary differences in monolingual and bilingual adults. *Biling. Lang. Cogn.* 15, 397–401. doi: 10.1017/S136672891100040X
- Blumenfeld, H. K., and Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition* 118, 245–257. doi: 10.1016/j.cognition.2010.10.012
- Bogulski, C. A., Rakoczy, M., Goodman, M., and Bialystok, E. (2015). Executive control in fluent and lapsed bilinguals. *Biling. Lang. Cogn.* 18, 561–567. doi: 10.1017/S1366728914000856
- Burke, D. M., and Shafto, M. A. (2008). “Language and aging,” in *The Handbook of Aging and Cognition*, 3rd Edn, eds F. I. M. Craik and T. A. Salthouse (New York, NY: Psychology Press), 373–443.
- Champagne-Lavau, M., and Joannette, Y. (2009). Pragmatics, theory of mind and executive functions after a right-hemisphere lesion: different patterns of deficits. *J. Neurolinguistics* 22, 413–426. doi: 10.1016/j.jneuroling.2009.02.002
- Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., et al. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 302–312. doi: 10.1037/0278-7393.34.2.302
- Craik, F. I. M., Bialystok, E., and Freedman, M. (2010). Delaying the onset of Alzheimer disease. *Neurology* 75, 1726–1729. doi: 10.1212/WNL.0b013e3181fc2a1c
- de Bruin, A., Bak, T. H., and Della Sala, S. (2015). Examining the effects of active versus inactive bilingualism on executive control in a carefully matched non-immigrant sample. *J. Mem. Lang.* 85, 15–26. doi: 10.1016/j.jml.2015.07.001
- de Oliveira, M. O., Nitrini, R., Yassuda, M. S., and Brucki, S. M. D. (2014). Vocabulary is an appropriate measure of premorbid intelligence in a sample with heterogeneous educational level in Brazil. *Behav. Neurol.* 2014:87596. doi: 10.1155/2014/875960
- Dong, Y., and Li, P. (2015). The cognitive science of bilingualism. *Lang. Linguist. Compass* 9, 1–13. doi: 10.1111/lnc3.12099
- Fernald, A., Marchman, V. A., and Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Dev. Sci.* 16, 234–248. doi: 10.1111/desc.12019
- Garcia, P. (2004). Pragmatic Comprehension of High and Low Level Language Learners. *The Electronic Journal for English as a Second Language* 8. Available online at: <https://files.eric.ed.gov/fulltext/EJ1068107.pdf>
- Gladsjo, J. A., Schuman, C. C., Evans, J. D., Peavy, G. M., Miller, S. W., and Heaton, R. K. (1999). Norms for letter and category fluency: demographic corrections for age, education, and ethnicity. *Assessment* 6, 147–178. doi: 10.1177/107319119900600204
- Glucksberg, S., Newsome, M. R., and Goldvarg, Y. (2001). Inhibition of the literal: filtering metaphor-irrelevant information during metaphor comprehension. *Metaphor Symbol* 16, 277–298. doi: 10.1080/10926488.2001.9678898
- Gollan, T., and Brown, A. S. (2006). From Tip-of-the-Tongue (TOT) data to theoretical implications in two steps: when more TOTs means better retrieval. *J. Exp. Psychol. Gen.* 135, 462–483. doi: 10.1037/0096-3445.135.3.462
- Grosjean, F. (2010). *Bilingual: Life and Reality*. Cambridge, MA: Harvard University Press. doi: 10.4159/9780674056459
- Gupta, A. F. (2010). “Singapore standard english revisited,” in *English in Singapore: Unity and Utility*, eds L. Lim, A. Pakir, and L. Wee (Hong Kong: Hong Kong University Press), 57–89. doi: 10.5790/hongkong/9789888028436.003.0003
- Gupta, A. F. (2012). “Grammar teaching and standards,” in *Principles and Practices for Teaching English As an International Language*, eds L. Alsagoff, Sandra Lee McKay, G. Hu, and W. A. Renandya (New York, NY; London: Routledge), 244–260.
- Hasher, L., and Zacks, R. T. (1988). Working memory, comprehension, and aging: a review and a new view. *Psychol. Learn. Motiv.* 22, 193–225. doi: 10.1016/S0079-7421(08)60041-9
- Hoff, E. (2003). The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. *Child Dev.* 74, 1368–1378. doi: 10.1111/1467-8624.00612
- Kavé, G., and Halamish, V. (2015). Doubly blessed: older adults know more vocabulary and know better what they know. *Psychol. Aging* 30, 68–73. doi: 10.1037/a0038669
- Kousaie, S., and Phillips, N. A. (2012). Ageing and bilingualism: absence of a “bilingual advantage” in Stroop interference in a nonimmigrant sample. *Q. J. Exp. Psychol.* 65, 356–369. doi: 10.1080/17470218.2011.604788
- Leimgruber, J. R. E. (2011). Singapore english. *Lang. Linguist. Compass* 5, 47–62. doi: 10.1111/j.1749-818X.2010.00262.x
- Lenhard, W., and Lenhard, A. (2016). *Calculation of Effect Sizes*. Dettelbach: Psychometrica. Available online at: https://www.psychometrica.de/effect_size.html
- Li, P., Sepanski, S., and Zhao, X. (2006). Language history questionnaire: a web-based interface for bilingual research. *Behav. Res. Methods* 38, 202–210. doi: 10.3758/BF03192770
- Luk, G., Bialystok, E., Craik, F. I. M., and Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *J. Neurosci.* 31, 16808–16813. doi: 10.1523/JNEUROSCI.4563-11.2011
- Manowong, S. (2011). “The study of ability to interpret conversational implicatures in English of Thai EFL learners,” in *The Asian Conference on Language Learning 2011 Official Proceedings* (Osaka).

- Mashal, N., Gavrieli, R., and Kavé, G. (2011). Age-related changes in the appreciation of novel metaphoric semantic relations. *Aging Neuropsychol. Cogn.* 18, 527–543. doi: 10.1080/13825585.2011.575202
- Mishra, R. K. (2015). Let's not forget about language proficiency and cultural variations while linking bilingualism to executive control. *Bilingualism* 18, 39–40. doi: 10.1017/S1366728914000455
- Morrone, I., Declercq, C., Novella, J., and Besche, C. (2010). Aging and inhibition processes: the case of metaphor treatment. *Psychol. Aging* 25, 697–701. doi: 10.1037/a0019578
- Newsome, M. R., and Glucksberg, S. (2002). Older adults filter irrelevant information during metaphor comprehension. *Exp. Aging Res.* 28, 253–267. doi: 10.1080/03610730290080317
- Nippold, M. A., Uhden, L. D., and Schwarz, I. E. (1997). Proverb explanation through the lifespan: a developmental study of adolescents and adults. *J. Speech Lang. Hear. Res.* 40, 245–253. doi: 10.1044/jslhr.4002.245
- O'Bryant, S. E., Humphreys, J. D., Smith, G. E., Ivnik, R. J., Graff-Radford, N. R., Petersen, R. C., et al. (2008). Detecting dementia with the Mini-Mental State Examination (MMSE) in highly educated individuals. *Arch. Neurol.* 65, 963–967. doi: 10.1001/archneur.65.7.963
- Olsen, R. K., Pangelinan, M. M., Bogulski, C., Chakravarty, M. M., Luk, G., Grady, C. L., et al. (2015). The effect of lifelong bilingualism on regional grey and white matter volume. *Brain Res.* 1612, 128–139. doi: 10.1016/j.brainres.2015.02.034
- Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., et al. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency. *Hum. Brain Mapp.* 19, 170–182. doi: 10.1002/hbm.10110
- Qualls, C. D., and Harris, J. L. (2003). Age, working memory, figurative language type, and reading ability: influencing factors in african american adults' comprehension of figurative language. *Am. J. Speech Lang. Pathol.* 12, 92–102. doi: 10.1044/1058-0360(2003/055)
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychol. Rev.* 103, 403–428. doi: 10.1037/0033-295X.103.3.403
- Salthouse, T. A., and Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Dev. Psychol.* 27, 763–776. doi: 10.1037/0012-1649.27.5.763
- Stewart, M. T., and Heredia, R. R. (2002). Comprehending spoken metaphoric reference: a real-time analysis. *Exp. Psychol.* 49, 34–44. doi: 10.1027/1618-3169.49.1.34
- Thornton, R., and Light, L. L. (2006). "Language comprehension and production in normal aging," in *Handbook of the Psychology of Aging*, 6th Edn, eds J. E. Birren and K. W. Schaie (London: Elsevier), 261–287. doi: 10.1016/B978-012101264-9/50015-X
- Titone, D., Pivneva, I., Sheikh, N. A., Webb, N., and Whitford, V. M. (2015). Doubling down on multifactorial approaches to the study of bilingualism & executive control. *Bilingualism* 18, 43–44. doi: 10.1017/S1366728914000595
- Tyler, L. K., Shafto, M. A., Randall, B., Wright, P., Marslen-Wilson, W. D., and Stamatakis, E. A. (2009). Preserving syntactic processing across the adult life span: the modulation of the frontotemporal language system in the context of age-related atrophy. *Cereb. Cortex* 20, 352–364. doi: 10.1093/cercor/bhp105
- Uekermann, J., Thoma, P., and Daum, I. (2008). Proverb interpretation changes in aging. *Brain Cogn.* 67, 51–57. doi: 10.1016/j.bandc.2007.11.003
- Ulatowska, H. K., Chapman, S. B., Highley, P., and Prince, J. (1998). Discourse in healthy old-elderly adults: a longitudinal study. *Aphasiology* 12, 619–633. doi: 10.1080/02687039808249562
- van Hell, J. G., and Poarch, G. (2014). How much bilingual experience is needed to affect executive control? *Appl. Psycholinguist.* 35, 925–28. doi: 10.1017/S0142716414000265
- Vega-Mendoza, M., West, H., Sorace, A., and Bak, T. H. (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition* 137, 40–46. doi: 10.1016/j.cognition.2014.12.008
- Westbury, C., and Titone, D. (2011). Idiom literality judgments in younger and older adults: age-related effects in resolving semantic interference. *Psychol. Aging* 26, 467–474. doi: 10.1037/a0022438
- Woods, D. L., Kishiyama, M. M., Yund, E. W., Herron, T. J., Edwards, B., Polivia, O., et al. (2011). Improving digit span assessment of short-term verbal memory. *J. Clin. Exp. Neuropsychol.* 33, 101–111. doi: 10.1080/13803395.2010.493149
- World Health Organisation (2002). *Proposed Working Definition of an Older Person in Africa for the MDS Project*. The World Health Organisation. Available online at: <http://www.who.int/healthinfo/survey/ageingdefnolder/en/>
- Zahodne, L. B., Schofield, P. W., Farrell, M. T., Stern, Y., and Manly, J. J. (2014). Bilingualism does not alter cognitive decline or dementia risk among Spanish-speaking immigrants. *Neuropsychology* 28, 238–246. doi: 10.1037/neu0000014

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Sundaray, Marinis and Bose. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Exploring Different Types of Inhibition During Bilingual Language Production

Maria Borraran¹, Clara D. Martin^{1,2}, Angela de Bruin¹ and Jon Andoni Duñabeitia^{1,3*}

¹ Basque Center on Cognition, Brain and Language, Donostia, Spain, ² IKERBASQUE, Basque Foundation for Science, Bilbao, Spain, ³ Facultad de Lenguas y Educación, Universidad Nebrija, Madrid, Spain

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Kenneth R. Paap,
San Francisco State University,
United States
Giovanna Bubbico,
G. d'Annunzio University of
Chieti-Pescara, Italy
J. Bruce Morton,
University of Western Ontario, Canada

*Correspondence:

Jon Andoni Duñabeitia
jdunabeitia@nebrja.es

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 02 July 2018

Accepted: 30 October 2018

Published: 20 November 2018

Citation:

Borraran M, Martin CD,
de Bruin A and Duñabeitia JA (2018)
Exploring Different Types of Inhibition
During Bilingual Language
Production. *Front. Psychol.* 9:2256.
doi: 10.3389/fpsyg.2018.02256

Multilinguals have to control their languages constantly to produce accurate verbal output. They have to inhibit possible lexical competitors not only from the target language, but also from non-target languages. Bilinguals' training in inhibiting incongruent or irrelevant information has been used to endorse the so-called bilingual advantage in executive functions, assuming a transfer effect from language inhibition to domain-general inhibitory skills. Recent studies have suggested that language control may rely on language-specific inhibitory control mechanisms. In the present study, unbalanced highly proficient bilinguals completed a rapid naming multi-inhibitory task in two languages. The task assessed three types of inhibitory processes: inhibition of the non-target language, inhibition of lexical competitors, and inhibition of erroneous auditory feedback. The results showed an interaction between lexical competition and erroneous auditory feedback, but no interactions with the inhibition of the non-target language. The results suggested that different subcomponents of language inhibition are involved during bilingual language production.

Keywords: inhibitory control, language production, bilingual experience, delayed auditory feedback, speech inhibition, lexical access

INTRODUCTION

Multilinguals have to manage different languages to control verbal speech on an everyday basis. They have to select the language that is needed at every specific moment and suppress interference from the situationally irrelevant languages. This mechanism is commonly referred to as language control and it has been associated with the use of a complex set of inhibitory control mechanisms (see Green, 1998; Kroll et al., 2008; Declerck and Philipp, 2015). Broadly speaking, inhibitory control refers to the suppression of interfering information or prepotent responses. In the influential framework published by Miyake et al. (2000), inhibition was proposed to be one of three separable components of executive functions (together with updating and shifting). However, a more recent framework (Miyake and Friedman, 2012) suggests that inhibition may not be a subcomponent but instead correlates perfectly with a "common executive function," defined as the ability to maintain and use task goals and goal-related information.

Research has shown that language inhibition may be required during bilingual speech production (e.g., Rodriguez-Fornells et al., 2005; de Bruin et al., 2014) and comprehension (e.g., Macizo et al., 2010; Durlak et al., 2016), with the results from the former group of studies being more consistent than those from the latter. However, the mechanisms underlying language inhibition

during bilingual speech production are not well understood yet, despite the importance of these inhibitory mechanisms in the debate on the so-called bilingual advantage in executive functions (see, among many others, Duñabeitia et al., 2014; Paap and Sawi, 2014; Bialystok, 2015; Duñabeitia and Carreiras, 2015; Sorge et al., 2017; see Lehtonen et al., 2018, for a recent review). In this respect, one important question is whether bilingual language inhibition is accomplished using the same mechanisms that are also used in non-linguistic inhibition tasks (i.e., a domain-general inhibitory mechanism; Jackson et al., 2001; Bialystok et al., 2008; Colzato et al., 2008; Linck et al., 2012; de Bruin et al., 2014) or whether mechanisms specific to linguistic inhibition are applied (Calabria et al., 2015; Branzi et al., 2016).

Furthermore, even within bilingual speech production, multiple types of linguistic conflict may be present that may be governed by different forms of inhibitory control mechanisms. The current study therefore assessed whether different types of interfering linguistic information are suppressed through a general inhibitory control mechanism or distinct mechanisms. To this end, we asked a group of highly proficient yet unbalanced Spanish-Basque bilinguals to complete a verbal production task either in their native language (Spanish) or in their non-native language (Basque), while parametrically manipulating other additional inhibitory demands (i.e., lexical inhibition and erroneous auditory feedback inhibition). While these three manipulations differ in many ways, they share one important component: the presence of interfering information that needs to be suppressed in order to correctly complete the task. In the current study, we explored the possible additive or interactive nature of the different types of interfering information in the context of a Rapid Automatized Naming (RAN) task (see Denckla and Rudel, 1974) in which we included several manipulations. The RAN task provides a unique opportunity to explore language-related interference at multiple levels, given that it taps into a fusion of linguistic, articulatory and attentional processes (see Cummine et al., 2014). In this line, it has been argued that over and above the obvious articulatory, motor and perceptual processes involved in the RAN, additional attentional, conceptual and phonological processes are required for successful performance (see Wolf and Bowers, 1999).

The first task manipulation concerned the use of the first or second language. It is widely assumed that multilingual speakers have to inhibit phonological and lexical competitors from the non-target language during speech production, so that speaking in one language requires *non-target language inhibition*. Green (1998) proposed an inhibitory control model in which multilinguals solve the conflict between languages through suppression of the representations from the non-target languages, while the representations from the target language are activated. Furthermore, the amount of inhibition needed to suppress the non-target language is argued to be related to language proficiency. In the case of a strong first language (L1) and weaker second language (L2), a relatively high level of inhibition of L1 is needed when speaking in L2. In contrast, when speaking in the stronger L1, less inhibition of the weaker L2 may be needed (although even

in these circumstances, non-target language inhibition may be needed).

But over and above inhibiting the non-target language, both monolingual and bilingual speech production require a series of processes related with *lexical inhibition*. According to Levelt's model of word production (Levelt et al., 1999), a series of automatic steps have to take place before a speaker voluntarily generates any word. First, she must identify a concept in the imagery system and activate the associated lexical representation(s). Then, she must select a suitable lexical item and inhibit the ones that share semantic, lexical, and syntactic properties with the target word. Finally, she must inhibit morphological and phonological competitors in order to retrieve the articulatory representation of the intended word. Hence, speakers have to inhibit possible lexical competitors in order to correctly produce the intended word (see Grainger and Jacobs, 1993; Abutalebi and Green, 2008; Philipp and Koch, 2009; Righi et al., 2010), and these lexical inhibitory mechanisms are qualitatively different from the non-target language inhibitory mechanisms insofar that the latter focus on the whole language system, while the former concentrate on the neighboring lexical representations.

But speech production does not exclusively rely on these two types of inhibitory mechanisms. During speech production, a speaker not only has to inhibit competitors at different levels of processing within the target and non-target languages, but she also has to trust her own auditory feedback to online monitor and control the articulatory output (Lee, 1950). Auditory feedback is a mechanism that helps to verify whether the current speech production is in agreement with the intention. In cases in which a mismatch is perceived, a correction mechanism operates at the level of production (see Burnett et al., 1998).

One interesting manipulation regarding speech monitoring is delayed auditory feedback (DAF), a technique that was initially developed to explore the importance of auditory feedback in speech production. The DAF is a technique in which speakers hear their own speech production through headphones, but with a short and artificially inserted lag between the actual production and its reproduction. Speech is normally inhibited using auditory feedback inhibition, which occurs online, with a very short delay. When the perception of speech is delayed – in this case artificially by playing back the sound with a delay – this auditory feedback inhibition becomes more costly and less efficient. Therefore, the auditory perceptual lag disturbs speech production, leading to disfluent utterances. The DAF technique facilitates understanding how production is achieved by exploring erroneous *auditory feedback inhibition* when such feedback is delayed and thus unreliable and even disturbing. In order to efficiently continue producing speech under DAF, speakers need to monitor the auditory feedback and inhibit the incorrectly timed input, while adjusting their utterances to the circumstances.

Some previous studies have suggested a relationship between auditory feedback and domain-general control processes. Adaptation to altered auditory feedback can be modulated by attentional load (e.g., Tumber et al., 2014; Scheerer et al., 2016)

and networks mediating domain-general cognitive control may also be involved in feedback monitoring (Schiffer et al., 2015). Other studies, however, did not observe such a link between altered auditory feedback and domain-general inhibitory control (Martin et al., 2018) and have suggested that feedback control may rely on perceptual acuity to compensate for this perturbation during speech production (Villacorta et al., 2005; Martin et al., 2018). Yet, as auditory feedback is linguistic, a relationship may exist with language control, and we tested this idea in the current study.

The precise way in which language proficiency impacts speech production under DAF is still a matter of debate. It is assumed that until a certain proficiency level is acquired in a non-native language (L2), the impact of the DAF technique is larger in that language than in the native one (L1). Several studies have shown an interaction between language dominance and erroneous auditory feedback inhibition, reporting longer speech latencies in L2 than in L1 (e.g., Lee, 1950; Mackay, 1970; Van Borsel et al., 2005). These results are consistent with the idea that bilinguals need more inhibitory resources when using their weaker L2 because they have to suppress the dominant L1 (Green, 1998). However, once multilingual speakers acquire a higher level of proficiency in L2, erroneous auditory feedback inhibition seems to occur similarly for native and non-native languages, suggesting the control of incorrect auditory feedback is not exclusively related to nativeness in a language (e.g., Siegel et al., 1984; Kvavik et al., 1991; Fabbro and Darò, 1995). The participants in our study were highly proficient in both languages, but still unbalanced with a higher proficiency level in L1 than in L2. As such, our participants could follow the pattern of previous studies showing similar erroneous auditory feedback inhibition for L1 and L2 in bilinguals with a high proficiency level. Alternatively, the unbalanced proficiency levels may still lead to an interaction between auditory feedback and language.

In the current study, DAF was used to assess whether the demands of inhibiting the delayed feedback could cause a processing bottleneck for other inhibitory demands during speech production in the context of a RAN task. The RAN task was originally designed to assess reading competence by naming pictures as fast as possible (see Denckla and Cutting, 1999, for a review). This task requires not only lexical access, but also the inhibition of the competitors flanking the target image (i.e., the neighboring representations sharing some properties with the target item). In cases when the matrices are made of pictorial elements referring to the same semantic category (e.g., a picture of an animal flanked by other animals of different species), the inhibitory demands increase, making lexical access slower and costlier (see Oppenheim et al., 2010; Mahon and Caramazza, 2011; Runnqvist et al., 2012). Thus, the RAN task seems to be a perfect test scenario to explore how multiple levels of linguistic inhibitory demands could interact with each other during language production.

We created a multilingual RAN-like picture naming task where three types of language inhibition mechanisms could be required across conditions: non-target language inhibition, lexical inhibition of the preponderant responses and the competitors, and erroneous auditory feedback inhibition. We

conceived an experimental design that allowed for observing how the system performs as a whole and whether these three variants of inhibitory processes at play during multilinguals' speech production interact with each other. Firstly, highly proficient bilingual participants were asked to name the pictures of the RAN either in their native language (Spanish) or in their non-native language (Basque), thus requiring non-target language inhibition to complete the different trials of the RAN scenario.

Furthermore, an additional artificial inhibitory demand was included in the experimental design, aimed at mimicking some of the lexical inhibitory processes that need to be carried out by multilinguals while producing speech. Multilinguals and language learners have to inhibit preponderant words from the native language that may interfere with the correct utterance in the non-native language (e.g., a Spanish-English bilingual would have to inhibit the translation equivalent "casa" to produce the word "house"). This is an everyday, constant demand for bilinguals, and in the context of the current experiment, we artificially created a similar demand with the aim of recreating a natural aspect of bilinguals' day-to-day interactions. We asked participants to substitute the name of certain pictures for that of some digits (e.g., say the word "two" when seeing the picture of a frog) in increasing order of difficulty, parametrically varying the number of to-be-replaced elements. Finally, auditory feedback demands were manipulated by including trials in which participants perceived their own speech without or with an artificial delay (DAF).

Thus, the aim of this study was to investigate how language inhibition works in multilingual speakers, particularly assessing whether distinct linguistic inhibitory processes that are at play during speech production interact with each other. By means of our multi-layered picture naming RAN-like task, we intended to tax the system and to evoke the use of large amounts of inhibitory resources, highlighting their independent effects and the interdependent interactions between them. The results of the current study will help us elucidating the extent to which language control mechanisms in multilingual speech production rely on a general mechanism of language control, or alternatively, on different subcomponents of language control. If the three types of inhibitory mechanisms interact with each other, this would suggest that language control relies at least partly on a shared inhibitory control mechanism. On the other hand, and in line with Sternberg's Additive Factors Method (Sternberg's (1969, 1998, 2011)), we argue that if the three types of interference manipulations show main effects that do not interact with each other, this would support the existence of independently operating processes. We would interpret a lack of interaction along the claims of the Additive Factors Method that has been successfully applied to visual object naming (see Sternberg, 1998), which endorses a view of additivity for functionally distinct processes that are separately modifiable (Sternberg, 2013). Thus, if we observe no interaction between non-target language inhibition (namely, the effect of naming items in L1 vs. in L2), erroneous auditory feedback inhibition (immediate feedback vs. delayed feedback), and lexical inhibition of competing representations, this would speak for a relative independence of the inhibitory

components, in line with the idea of different inhibitory mechanisms underpinning language control (e.g., Calabria et al., 2012; Branzi et al., 2016). If it is the case that the systems work separately, our results should also shed new light on specific inhibitory processes applied within the language-related inhibitory system.

MATERIALS AND METHODS

Participants

Fifty-six unbalanced Spanish-Basque bilingual young adults from the University of the Basque Country took part in this experiment (highest degree obtained was high school for 12 participants, professional training for 8, university degree for 31, and postgraduate degree for 2 participants). Three participants were excluded from analyses due to a high error rate (more than 35% errors in each matrix). All participants (M age = 23 years, SD = 3 years; 33 females) were native Spanish speakers, who acquired Basque early in life (see **Table 1**) and were more exposed to Spanish than to Basque in their daily life (see **Table 1**). Their language proficiency was assessed using two tests (see de Bruin et al., 2017, for further details): a picture naming task in which they were asked to name 65 common objects in each of the two languages (see **Table 1**), and a personal interview with a native bilingual linguist who rated them on a 1-to-5 scale (5: native-like competence; 1: basic/no knowledge; see **Table 1**). In addition, participants were asked to rate their competence (in terms of reading, speaking, writing, and understanding) on a scale from 0 to 10 (see **Table 1**). All participants were right-handed and none were diagnosed with language disorders, learning disabilities, or auditory impairments. After the experiment, they were reimbursed for their time. This study was carried out in accordance with the recommendations of the international ethical guidelines approved by the BCBL Ethics Committee with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the BCBL Ethics Committee.

Materials and Design

Pictures of concepts from two different semantic categories (animals and body parts) were used to create different matrices for the rapid naming task. The images were taken from the MultiPic database (Duñabeitia et al., 2018). While non-target language activation has often been studied with words that are similar between two languages (e.g., cognates), studies using non-cognates have also observed activation and inhibition of the non-target language (e.g., de Bruin et al., 2014). To avoid effects of cognate status, we explicitly avoided the use of concepts associated with cognate words, and all the items selected for the two categories had names that were non-cognates between Spanish and Basque, lacking substantial orthographic or phonological overlap between languages. The picture names were matched on syllable length, number of phonemes, and frequency of use between languages (see **Table 2**).

TABLE 1 | Table showing the participants' language profile.

	Spanish		Basque	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age of acquisition	0.5	1	3	1
Exposure	68.8	13	22.6	11
Picture naming test	64.6	0.8	52.1	6.6
Personal interview	5	0	3.9	0.3
Reading competence	9.6	0.7	7.4	1.2
Speaking competence	9.6	0.7	8.6	1.1
Writing competence	9.2	1.1	7.7	1.4
Understanding competence	9.6	0.7	8.6	1.3

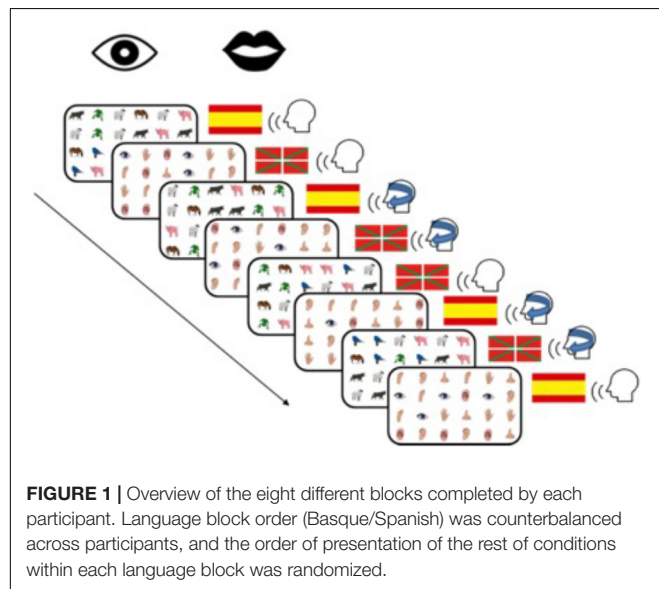
Mean and standard deviation are provided per language for age of acquisition (years), exposure (percentage of time exposed), picture naming test (number of correct items named in a scale from 0 to 65), personal interview (score in an ascending scale from 1 to 5), self-rated reading, speaking, writing, and understanding competence (scale 0–10).

The structure of the experiment and the order of conditions were as follows. Each participant completed eight blocks. Four blocks were completed in Spanish and four in Basque. Furthermore, two of the four blocks were completed without DAF and the other two with DAF (i.e., two of the four blocks) for each language. Each semantic category occurred once in each of these conditions, so that each participant completed the same semantic category twice in each language, once with and once without DAF (see **Figure 1**). The blocks were distributed over the experiment so that in the first half of the experiment, some participants completed all body part blocks in Spanish (both with DAF and without DAF) and all animal blocks in Basque. In the second half of the experiment, these participants would then complete all animal blocks in Spanish and all body part blocks in Basque. For other participants, this order was reversed so that the language in which each semantic category was named first was counterbalanced. Within each half of the experiment, the order of languages and DAF condition was randomized so that some participants started in Spanish and some in Basque.

Furthermore, each block was composed of 4 matrices that were always completed in the language of that block. Each matrix included 24 pictures (i.e., each of the 6 individual items was repeated four times; see **Appendix 1**) aligned in columns and rows and arranged at random. Participants were asked to name the pictures of the first matrices of each block normally. Then, in the subsequent 3 matrices, participants were asked to replace the name of certain items with the name of some numbers. In the second matrix of each block, the name of one item (i.e., one animal or one body part) had to be consistently replaced throughout the completion of the matrix with the name of the first cardinal number (e.g., replace the word “frog” with the word “one”). In the third matrix, the names of two pictures had to be substituted by the names of two digits. Finally, in the fourth matrix of each block, participants were asked to replace the names of three different items with those of three digits. The specific rules for the replacements in each matrix were presented to the participants at the beginning of each trial. The blocks were counterbalanced for language and the presence or absence of

TABLE 2 | Means (M) and standard deviations (SD) of the number of syllables, phonemes and of the frequency of use (in number of appearances per million) of the common names used in the task, as obtained from the E-Hitz (Perea et al., 2006) and B-Pal (Davis and Perea, 2005) databases.

	Animals				Body parts			
	Spanish		Basque		Spanish		Basque	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Syllables	2.5	0.5	2	0	2	0.6	2	0.6
Phonemes	5	0.9	4.3	0.5	4	0.8	4	1.4
Word frequency	20.6	21.5	33.6	24.1	136.8	132.3	116.3	106.1



DAF, but lexical substitutions were always presented in the same order following an increasing order of difficulty from 0 to 3.

Procedure

Participants were tested individually in a soundproof room. They were seated at a distance of about 45 cm from a laptop with a 13-inch screen, where all the stimuli were presented using Experiment Builder (SR-Research, Ontario, Canada). The same software was also used to collect the verbal responses, which were recorded from the onset of the presentation of each matrix to the moment in which the participant pressed the space bar to indicate that she had finished naming the items. Participants wore a headset throughout the experiment and they were instructed to name the items of each matrix as if they were reading a text (left to right, top to bottom), as fast as they could and trying not to make errors. Delayed auditory feedback time was set to 200 ms for the DAF condition in accordance with previous studies (Stuart et al., 2002; Hashimoto and Sakai, 2003) and to 0 ms for the immediate (no-DAF) condition. To this end, a SmallTalk device (Casa Futura Technologies, Colorado, United States) set at 80 dB was used.

Each matrix in each block was preceded by a screen specifying the instructions and conditions about the auditory feedback (delayed vs. immediate), the assigned language (L1 vs. L2), and

the number of lexical substitutions that they had to do (0, 1, 2, or 3). Participants were instructed whether there would be a delay in auditory feedback to avoid the disruptive effect of the delay being larger at the beginning of the task due to it being a surprise. Participants were familiarized with the pictures' names before performing the rapid naming task in the two languages and they practiced the lexical substitution task with fruit matrices before the experiment started.

Data Analysis

Data were analyzed in two ways. Firstly, we performed a three way ANOVA testing the effects of Language (L1| L2), Auditory Feedback (immediate| delayed), and Lexical Replacements (namely, the number of substitutions: 0| 1| 2| 3) on the naming latencies. As we aimed to examine whether different types of linguistic conflict interacted or not, we furthermore analyzed the data using Bayesian analysis. In the case of a null effect, a *p*-value can only say that there was no evidence for an effect, but it does not support the absence of an effect. By reporting Bayes Factors (BF), we show the ratio of the probability that the data were observed under the null hypothesis (e.g., “no interaction between auditory feedback and language”) vs. the probability of observing the data under the alternative hypothesis (e.g., “an interaction between auditory feedback and language”). For instance, a BF_{01} of 5 indicates that the observed data were five times more likely to have occurred under the null than alternative hypothesis. Bayesian analyses were conducted with JASP 0.8.5 using Bayesian repeated measures ANOVA with 100,000 samples. As we were interested in the interactions between the three different manipulations, we compared the model with the three main effects of Language, Auditory Feedback, and Lexical Replacements to a model including those three main effects plus, (a) the interaction between Language and Auditory Feedback; (b) the interaction between Language and Lexical replacements; and (c) the interaction between Auditory Feedback and Lexical Replacements.

RESULTS

We exclusively focused on the naming latencies given that the average number of errors was fairly low ($M = 1.05$ errors per matrix, $SD = 1.30$; range: 0.17–2.17). Besides, it is likely that any effects of production errors are also observable in the naming

latencies since participants corrected themselves when making a mistake, thus requiring more time to complete the matrix.

The three main effects were significant. For the effect of Language, blocks that were named in Basque (L2, non-native language) yielded longer reaction times ($M = 23.61$ s, $SD = 5.20$) than blocks completed in Spanish (L1, native language; $M = 22.67$ s, $SD = 4.93$), $F(1, 52) = 9.30$, $p < 0.004$, $\eta_p^2 = 0.152^1$. Regarding the effect of Auditory Feedback, blocks that had to be named under the DAF condition required more time ($M = 24.18$ s, $SD = 5.19$) than blocks that did not include any DAF ($M = 22.10$ s, $SD = 4.95$, $F(1, 52) = 57.26$, $p < 0.001$, $\eta_p^2 = 0.524$). Finally, the main effect of Lexical Replacements showed an increase in the naming latencies as a function of the number of words that had to be substituted, $F(3, 156) = 38.73$, $p < 0.001$, $\eta_p^2 = 0.427$, ranging from trials requiring no replacements ($M = 21.88$ s, $SD = 4.60$) to trials requiring 3 substitutions ($M = 24.87$ s, $SD = 5.91$) (see **Table 3** and **Figure 2** for details). There was a significant interaction between Auditory Feedback and Lexical Replacements, $F(3, 156) = 3.67$, $p = 0.014$, $\eta_p^2 = 0.066$, such that the effect of the DAF diminished as the number of replacements increased (see **Figure 2**). Nonetheless, and in spite of the decreasing magnitudes of the effect of the auditory feedback with the increased lexical replacement demands, this effect was always significant [No substitutions: $t(52) = 5.28$, $p < 0.001$, Cohen's $d = 0.726$; 1 substitution: $t(52) = 6.95$, $p < 0.001$, Cohen's $d = 0.956$; 2 substitutions: $t(52) = 6.76$, $p < 0.001$, Cohen's $d = 0.929$; 3 substitutions: $t(52) = 2.83$, $p < 0.006$, Cohen's $d = 0.390$]. Importantly, there was no interaction between Language and Auditory Feedback, $F(1, 52) = 0.20$, $p = 0.653$, $\eta_p^2 = 0.004$ (see **Figure 3**), or between Language and Lexical Replacements, $F(3, 156) = 1.32$, $p = 0.270$, $\eta_p^2 = 0.025$ (see **Figure 4**), nor was there a three-way interaction between all the factors, $F(3, 156) = 0.41$, $p = 0.750$, $\eta_p^2 = 0.008$.

We then conducted Bayesian analyses in which we compared models including the main effects only to the model including the main effects plus the interaction of interest. For the model including the interaction between Language and Lexical Replacements, the BF_{01} was 37.85 ($\pm 2.65\%$), suggesting that model without an interaction fits the data around 38 times better than a model with this interaction term included. Similarly, for the model including the Language \times Auditory Feedback interaction, the BF_{01} was 8.76 ($\pm 1.71\%$), suggesting that the model without such interaction accounts for the data nearly 9 times better than the model with the interaction. Thus, both analyses suggested that there was no interaction between Language and Auditory Feedback or Lexical Replacements. Regarding the interaction between Lexical Replacements and

Auditory Feedback, while the p -value showed a significant interaction, the Bayes Factor analysis showed some evidence against an interaction with a BF_{01} of 4.25 ($\pm 4.18\%$).

DISCUSSION

The aim of the study was to explore how distinct types of language inhibition that are typically needed by multilingual speakers to efficiently produce speech (namely, non-target language inhibition, inhibition of lexical competitors, and inhibition of erroneous auditory feedback) interact with each other during completion of a rapid naming task. Our main interest was to examine whether these different processes rely on the same linguistic inhibitory system or whether, alternatively, several types of different and independent linguistic inhibitory mechanisms underlie each of the distinct processes. To this end, we designed a highly demanding rapid naming task to allow us to observe how the inhibitory system(s) work(s) while multitasking.

In accordance with previous findings and as predicted, our results revealed main effects in all the three variables of interest. Firstly, participants exhibited longer naming times overall in L2 (Basque) as compared to L1 (Spanish), in line with the bulk of preceding evidence at this regard (e.g., Meuter and Allport, 1999; Costa and Santesteban, 2004). Second, longer naming latencies were also observed under DAF conditions as compared to immediate feedback conditions (see Lee, 1950; Mackay, 1970; Van Borsel et al., 2005). Finally, regarding the effect of replacing a preponderant response with a different lexical label according to newly learned rules, we found that naming times increased as a function of the number of replacements that were required. These effects suggest that the current test scenario readily tapped into a set of inhibitory mechanisms whose role and degree of implication were more prominent as the task demands increased (see Wickens, 2002).

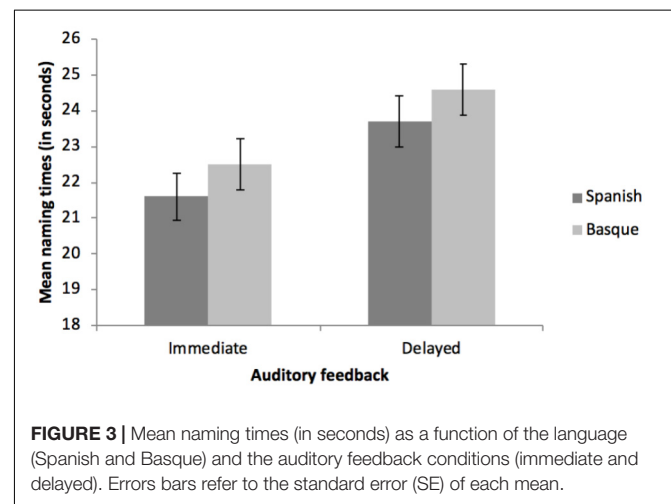
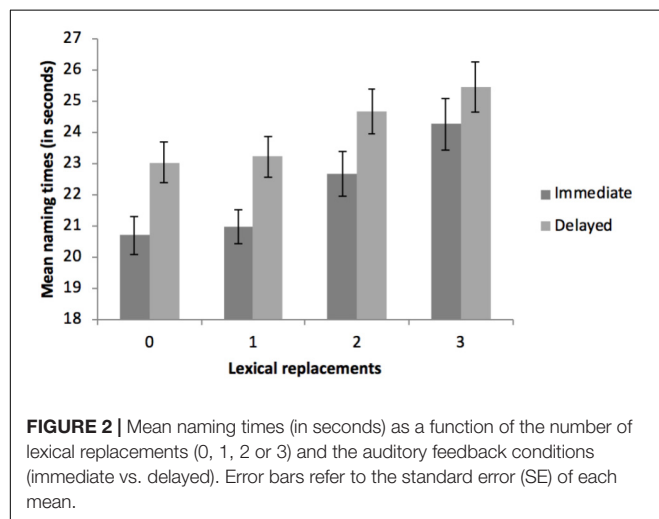
Interestingly, only two of the constructs associated with different inhibitory demands interacted with each other in the classic factorial analysis of variance (although the Bayes Factor actually provided some evidence *against* this interaction). The negative impact of the DAF partially decreased as the lexical competition increased (namely, as the need for controlling and inhibiting a preponderant response increased), potentially suggesting that both may tap into similar inhibitory resources. This interaction could also be understood in terms of a plateau effect in the three-replacement condition when the maximal taxation of cognitive resources is reached. However, none of these effects significantly interacted with the language at use (native vs. non-native), and the relative independence of this effect speaks for a certain degree of separation or autonomy of the different types of inhibitory mechanisms that multilinguals may use and require.

Crucially, the language (L1 or L2) did not interact with either lexical competition or auditory feedback alteration. This suggests that inhibitory mechanisms used to suppress

¹To further examine the role of non-target language inhibition in L1 vs. L2, we examined effects of language order. If L1 is inhibited strongly during L2 production, L1 responses should be relatively slow after an L2 task compared to L2 after L1. Participants always named one semantic category in one language in the first half of the experiment and in the other language in the other half. There was a significant effect of task half [$F(1, 52) = 8.77$, $p = 0.005$, $\eta_p^2 = 0.144$], reflecting that responses were faster in the first half ($M = 22.70$ s, $SD = 4.30$) than in the second half ($M = 23.58$ s, $SD = 4.50$) of the experiment. However, while the difference between task half was numerically larger for Spanish ($M = 1.23$; $SD = 2.62$) than Basque ($M = 0.53$; $SD = 3.95$), task half did not interact significantly with language [$F(1, 52) = 1.00$, $p = 0.321$, $\eta_p^2 = 0.019$].

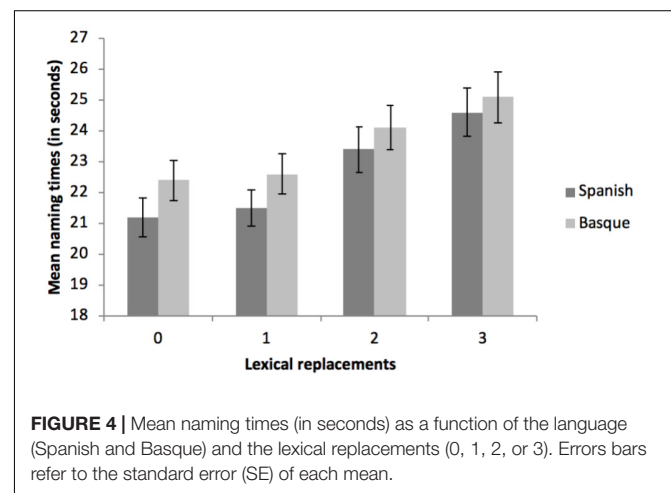
TABLE 3 | Descriptive statistics of the mean naming latencies (in seconds) across participants in each condition.

	Lexical replacements							
	0		1		2		3	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
NATIVE LANGUAGE (SPANISH)								
Immediate	20.10	3.91	20.11	3.92	22.37	5.28	24.11	5.65
Feedback								
Delayed	22.41	5.02	22.60	4.58	24.50	5.28	25.18	5.86
Feedback								
NON-NATIVE LANGUAGE (BASQUE)								
Immediate	21.33	4.94	21.32	4.24	22.99	5.33	24.47	6.35
feedback								
Delayed	23.68	4.53	23.85	5.21	24.83	5.19	25.72	5.81
Feedback								



the non-target language are not identical to the inhibition mechanisms used to accomplish suppression of erroneous auditory feedback or lexical competitors. The latter manipulation may also have introduced an increase in working memory load, which could explain the main effect of lexical competitors that was observed. However, beyond the working memory component related to remembering which replacement to use, the task also required inhibition of the word that could not be used (or resistance to the interference created by this salient representation). Thus, if the inhibitory component of lexical competition is related to the type of control used in language inhibition, we should have observed an interaction. Instead, and in line with Sternberg's Additive Factors Method Sternberg's (1998; 2013), the additive nature of these effects and the demonstration that they are independently modifiable, support the idea of a functional difference between them.

Previous studies have already suggested that language inhibition may be based on its own specific and independent resources (see Abutalebi and Green, 2008; Calabria et al., 2012;



Blanco-Elorrieta and Pykkänen, 2016; Branzi et al., 2016). Our results are in line with this view, suggesting that the inhibition of the non-target language is likely to be managed through a

set of inhibitory resources that are not shared or required by other tasks. These results follow what Shell et al. (2015) showed in a similar study manipulating inhibitory control demands in a picture-word interference task that also involved a language-switching paradigm. No interaction between the effects of the lexical competitors and the effects of the language at use was found in this study either, suggesting that the underlying processes may not require overlapping or shared inhibitory mechanisms, ultimately suggesting that multilingual language control could use a highly specific and independent inhibitory mechanism for non-target language inhibition.

On the other hand, our data are less compatible with previous studies suggesting that language inhibition is at least partly related to the inhibition mechanisms used in other (non-verbal) tasks (e.g., Linck et al., 2012; de Bruin et al., 2014). However, these studies have typically used a language switching paradigm, which may place additional demands on language inhibition and/or may require a different form of inhibition. For instance, while our task required more global, proactive inhibition of the non-target language, language switching may make additional use of local, reactive inhibition mechanisms (Green, 1998). Considering the design of the current experiment, it was not possible to examine more short-lived effects of inhibition at the level of individual items. Future experiments will need to examine whether different tasks eliciting stronger and/or more local effects of language inhibition show connections between different types of language inhibition.

The absence of an interaction between inhibition of the delayed, erroneous auditory feedback and inhibition of the non-target language is line with a previous study (Martin et al., 2018) showing that the type of inhibition applied to compensate for the presence of altered feedback does not correlate with other types of inhibitory processes. It is furthermore worth noting that while some preceding studies have shown an interaction between the effects of DAF and those of language dominance (see Lee, 1950; Mackay, 1970; Van Borsel et al., 2005), such an interaction has not been found in samples of highly proficient bilinguals. For instance, Fabbro and Darò (1995) conducted a study with highly proficient interpreters performing under DAF conditions and found no interaction between the critical variables of interest. The participants in our study were highly proficient in both languages. In line with preceding evidence showing that highly proficient bilinguals rely on different language-selection mechanisms than low proficient bilinguals (e.g., Costa and Santesteban, 2004; Costa et al., 2006), increased proficiency in the non-native language could make the reliance on erroneous auditory feedback more similar to that of the native language. If so, the set of inhibitory mechanisms that are used to partial out the negative effect of incorrect (altered or delayed) auditory feedback could be similar in L1 and L2. This is precisely what we found in the current study, suggesting that in highly proficient bilinguals, the underlying processes responsible to monitor erroneous auditory feedback and inhibit the potentially disturbing incorrect feedback may not be linguistic in essence, and that it may instead correspond to a specific type of mechanisms that are linked to perceptual acuity.

This study suggests that multilinguals do not rely on one unique inhibitory system to produce speech in one of the known languages, but rather that they rely on a set of different mechanisms that operate separately. As said, the interpretation that additive factors can reflect independent processes is in line with Sternberg's Additive Factors Method (Sternberg's (1998)). This approach has been used frequently to study the independence of and similarities between different processes involved in inhibitory control (e.g., Los, 2004; van den Wildenberg and van der Molen, 2004). However, other studies have questioned the reliability of inferring underlying processes from patterns seen in response time data. Stafford and Gurney (2011) showed that both models with discrete stages and continuous models with simultaneously run, interacting processes could mimic additive factors. Any inference with respect to distinct vs. interacting processes based on additive factors should thus be interpreted with caution. In line with Sternberg's (2013) response to Stafford and Gurney, we therefore interpret our data as *supporting* rather than unequivocally implying distinct inhibitory control processes.

The three manipulations used in our study were not only tapping into inhibitory control, but are also likely to recruit other forms of cognitive control (e.g., working memory or rule learning, in the case of the manipulation involving lexical replacements). While it is perfectly possible that the three manipulations were different enough not to recruit a shared inhibitory control mechanism, it should be considered that the three of them required the inhibition of interfering information (in the form of another language, lexical items, or erroneous auditory feedback). We believe that if different types of inhibitory control are governed by a common inhibitory control mechanism, these three specific forms of resistance to linguistic interference would be expected to tap into this general mechanism.

Our results open up the possibility to think about how the inhibitory system is further divided into domain-specific sub-mechanisms that do not necessarily work at par. These distinct mechanisms should be explored in more detail in order to better understand how they work and relate to each other in an interactive fashion, and specifically, how they are used by multilinguals to efficiently face the highly demanding communicative scenarios they encounter in their daily life. Furthermore, brain imaging techniques could be used to shed more light on the possible differences between or overlap in the spatial and temporal characteristics of different types of inhibitory control. We tentatively propose that the debate on the generality or specificity of the language-related inhibitory mechanisms should be moved to a new arena, leaving aside the simplistic dichotomy between language-specific inhibition and domain-general inhibition. The current study suggests that the different inhibitory processes that mediate multilingual speech production are somewhat independent from each other, probably referring to diverse inhibitory modules. Some of these mechanisms, such as the inhibitory control of erroneous auditory feedback and the inhibition of competing lexical representations, could tap into similar inhibitory resources. However, other mechanisms, such as the set of inhibitory processes applied to

control for the interference from the non-target language, seem to be independent. Together, these results make us think of a series of inhibitory modules that go beyond a unitary conception of language-specific inhibitory mechanisms.

In sum, these findings demonstrate that some of the mechanisms related to language control require allocating particular and independent inhibitory resources that remain unaffected by the concurrent requirement of other inhibitory mechanisms. This suggests that multilingual language control builds on a set of specific inhibitory mechanisms that are not shared by other cognitive or even other language processes. Future studies will help us elucidating the precise nature and role of these seemingly independent inhibitory processes, and the way in which they are acquired, developed and trained in multilingual contexts.

REFERENCES

- Abutalebi, J., and Green, D. W. (2008). Control mechanisms in bilingual language production: neural evidence from language switching studies. *Lang. Cogn. Process.* 23, 557–582. doi: 10.1080/01690960801920602
- Bialystok, E. (2015). Bilingualism and the development of executive function: the role of attention. *Child Dev. Perspect.* 9, 117–121. doi: 10.1111/cdep.12116
- Bialystok, E., Craik, F., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol.* 34, 859–873. doi: 10.1037/0278-7393.34.4.859
- Blanco-Elorrieta, E., and Pykkänen, L. (2016). Bilingual language control in perception versus action: meg. reveals comprehension control mechanisms in anterior cingulate cortex and domain-general control of production in dorsolateral prefrontal cortex. *J. Neurosci.* 36, 290–301. doi: 10.1523/JNEUROSCI.2597-15.2016
- Branzi, F. M., Calabria, M., Boscarino, M. L., and Costa, A. (2016). On the overlap between bilingual language control and domain-general executive control. *Acta Psychol.* 166, 21–30. doi: 10.1016/j.actpsy.2016.03.001
- Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *J. Acoust. Soc. Am.* 103, 3153–3161. doi: 10.1121/1.423073
- Calabria, M., Branzi, F. M., Marne, P., Hernández, M., and Costa, A. (2015). Age-related effects over bilingual language control and executive control. *Bilingualism* 18, 65–78. doi: 10.1017/S1366728913000138
- Calabria, M., Hernández, M., Branzi, F. M., and Costa, A. (2012). Qualitative differences between bilingual language control and executive control: evidence from tasks switching. *Front. Psychol.* 2:399. doi: 10.3389/fpsyg.2011.00399
- Colzato, L. S., Bajo, M. T., Van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., et al. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *J. Exp. Psychol.* 34, 302–312. doi: 10.1037/0278-7393.34.2.302
- Costa, A., and Santesteban, M. (2004). Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Mem. Lang.* 50, 491–511. doi: 10.1016/j.jml.2004.02.002
- Costa, A., Santesteban, M., and Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *J. Exp. Psychol.* 32, 1057–1074. doi: 10.1037/0278-7393.32.5.1057
- Cummine, J., Szepesvari, E., Chouinard, B., Hanif, W., and Georgiou, G. K. (2014). A functional investigation of RAN letters, digits, and objects: how similar are they? *Behav. Brain Res.* 275, 157–165. doi: 10.1016/j.bbr.2014.08.038
- Davis, C. J., and Perea, M. (2005). BuscaPalabras: a program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behav. Res. Methods* 37, 665–671. doi: 10.3758/BF03192738
- de Bruin, A., Carreiras, M., and Duñabeitia, J. A. (2017). The BEST dataset of language proficiency. *Front. Psychol.* 8:522. doi: 10.3389/fpsyg.2017.00522
- de Bruin, A., Roelofs, A., Dijkstra, T., and Fitzpatrick, I. (2014). Domain-general inhibition areas of the brain are involved in language switching: fMRI evidence from trilingual speakers. *Neuroimage* 90, 348–359. doi: 10.1016/j.neuroimage.2013.12.049
- Declerck, M., and Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychon. Bull. Rev.* 22, 1630–1645. doi: 10.3758/s13423-015-0836-1
- Denckla, M. B., and Cutting, L. E. (1999). History and significance of rapid automatized naming. *Ann. Dyslexia* 49, 29–42. doi: 10.3389/fpsyg.2016.00021
- Denckla, M. B., and Rudel, R. (1974). Rapid “Automatized” Naming of Pictured Objects, Colors, Letters and Numbers by Normal Children. *Cortex* 10, 186–202. doi: 10.1016/S0010-9452(74)80009-2
- Duñabeitia, J. A., and Carreiras, M. (2015). The bilingual advantage: acta est fabula? *Cortex* 73, 371–372. doi: 10.1016/j.cortex.2015.06.009
- Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., et al. (2018). MultiPic: a standardized set of 750 drawings with norms for six European languages. *Q. J. Exp. Psychol.* 71, 808–816. doi: 10.1080/17470218.2017.1310261
- Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., et al. (2014). The inhibitory advantage in bilingual children revisited: myth or reality? *Exp. Psychol.* 61, 234–251. doi: 10.1027/1618-3169/a000243
- Durlak, J., Szewczyk, J., Muszyński, M., and Wodniecka, Z. (2016). Interference and inhibition in bilingual language comprehension: evidence from Polish-English interlingual homographs. *PLoS One* 11:e0151430. doi: 10.1371/journal.pone.0151430
- Fabbro, F., and Darò, V. (1995). Delayed auditory feedback in polyglot simultaneous interpreters. *Brain Lang.* 48, 309–319. doi: 10.1006/brln.1995.1013
- Grainger, J., and Jacobs, A. M. (1993). Masked partial-word priming in visual word recognition: effects of positional letter frequency. *J. Exp. Psychol.* 19, 951–964. doi: 10.1037/0096-1523.19.5.951
- Green, D. W. (1998). Mental control of the bilingual lexicosemantic system. *Bilingualism* 1, 67–81. doi: 10.1017/S1366728998000133
- Hashimoto, Y., and Sakai, K. L. (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: an fMRI study. *Hum. Brain Mapp.* 20, 22–28. doi: 10.1002/hbm.10119
- Jackson, G. M., Swainson, R., Cunningham, R., and Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism* 4, 169–178. doi: 10.1017/S1366728901000268
- Kroll, J. F., Bobb, S. C., Misra, M., and Guo, T. (2008). Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychol.* 128, 416–430. doi: 10.1016/j.actpsy.2008.02.001
- Kvavik, K. H., Katsuki-Nakamura, J., Siegel, G. M., and Pick, H. L. Jr. (1991). Delayed auditory feedback effects on learners of Japanese and Spanish. *Folia Phoniatr. Logop.* 43, 282–290. doi: 10.1159/000266139
- Lee, B. S. (1950). Some effects of side-tone delay. *J. Acoust. Soc. Am.* 22, 639–640. doi: 10.1121/1.1906665

AUTHOR CONTRIBUTIONS

MB, CM, and JD designed the study. MB and JD created the materials and experimental protocols. MB collected the data, analyzed the data under the supervision of all the authors, and drafted the manuscript. All authors provided critical comments on the manuscript before submission.

FUNDING

This research has been partially funded by grants PSI2015-65689-P, PSI2014-54500, PSI2017-82941-P, and SEV-2015-0490 from the Spanish Government, PI-2015-1-25 from the Basque Government and ATHEME-613465 from the European Union.

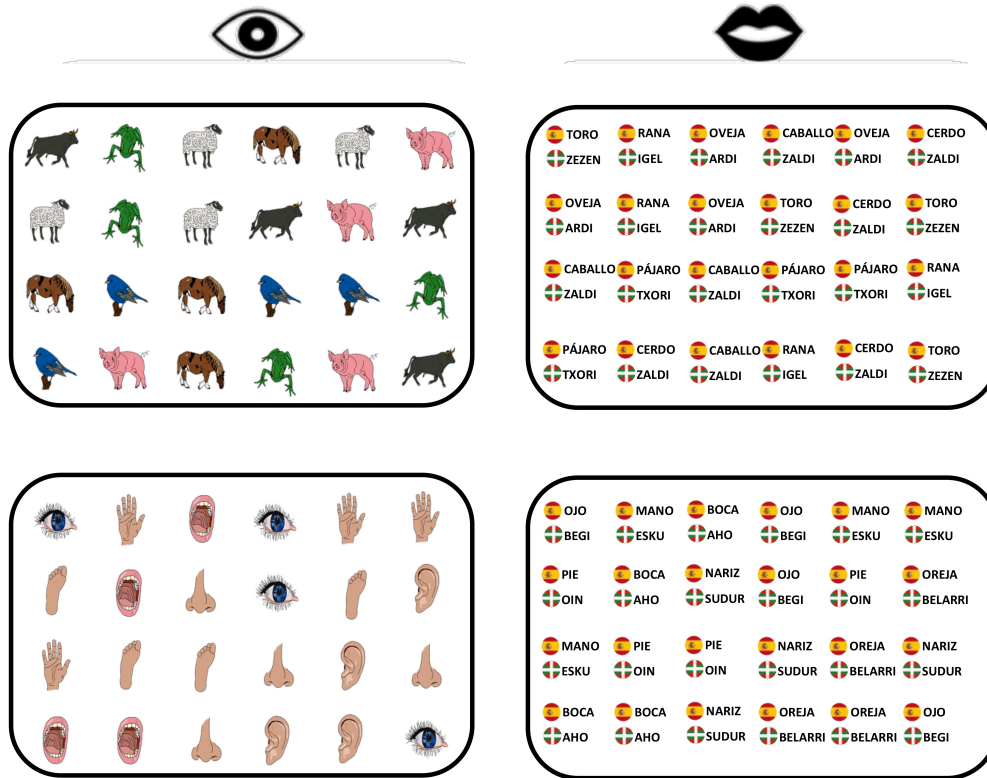
- Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., and Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? *Psychol. Bull.* 144, 394–425. doi: 10.1037/bul0000142
- Levelt, W. J., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–38. doi: 10.1017/S0140525X99001776
- Linck, J. A., Schwieter, J. W., and Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism* 15, 651–662. doi: 10.1017/S136672891100054X
- Los, S. A. (2004). Inhibition of return and nonspecific preparation: separable inhibitory control mechanisms in space and time. *Percept. Psychophys.* 66, 119–130. doi: 10.3758/BF03194866
- Macizo, P., Bajo, T., and Martín, M. C. (2010). Inhibitory processes in bilingual language comprehension: evidence from Spanish–English interlexical homographs. *J. Mem. Lang.* 63, 232–244. doi: 10.1016/j.jml.2010.04.002
- Mackay, D. G. (1970). How does language familiarity influence stuttering under delayed auditory feedback? *Percept. Motor Skills* 30, 655–669. doi: 10.2466/pms.1970.30.2.655
- Mahon, B. Z., and Caramazza, A. (2011). What drives the organization of object knowledge in the brain? *Trends Cogn. Sci.* 15, 97–103. doi: 10.1016/j.tics.2011.01.004
- Martin, C. D., Niziolek, C. A., Duñabeitia, J. A., Pérez, A., Hernández, D., Carreiras, M., et al. (2018). Online adaptation to altered auditory feedback is predicted by auditory acuity and not by domain-general executive control resources. *Front. Hum. Neurosci.* 12:91. doi: 10.3389/fnhum.2018.00091
- Meuter, R. F., and Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *J. Mem. Lang.* 40, 25–40. doi: 10.1037/a0034060
- Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734
- Oppenheim, G. M., Dell, G. S., and Schwartz, M. F. (2010). The dark side of incremental learning: a model of cumulative semantic interference during lexical access in speech production. *Cognition* 114, 227–252. doi: 10.1016/j.cognition.2009.09.007
- Paap, K. R., and Sawi, O. M. (2014). Bilingual advantages in executive functioning: problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Front. Psychol.* 5:962. doi: 10.3389/fpsyg.2014.00962
- Perea, M., Urkia, M., Davis, C. J., Agirre, A., Laseka, E., and Carreiras, M. (2006). E-Hitz: a word-frequency list and a program for deriving psycholinguistic statistics in an agglutinative language (Basque). *Behav. Res. Methods* 38, 610–615. doi: 10.3758/BF03193893
- Philipp, A. M., and Koch, I. (2009). Inhibition in language switching: what is inhibited when switching between languages in naming tasks? *J. Exp. Psychol. Learn. Mem. Cogn.* 35:1187. doi: 10.1037/a0016376
- Righi, G., Blumstein, S. E., Mertus, J., and Worden, M. S. (2010). Neural systems underlying lexical competition: an eye tracking and fMRI study. *J. Cogn. Neurosci.* 22, 213–224. doi: 10.1162/jocn.2009.21200
- Rodriguez-Fornells, A., Van Der Lugt, A., Rotte, M., Britti, B., Heinze, H. J., and Münte, T. F. (2005). Second language interferes with word production in fluent bilinguals: brain potential and functional imaging evidence. *J. Cogn. Neurosci.* 17, 422–433. doi: 10.1162/0898929053279559
- Runnqvist, E., Strijkers, K., Alario, F. X., and Costa, A. (2012). Cumulative semantic interference is blind to language: implications for models of bilingual speech production. *J. Mem. Lang.* 66, 850–869. doi: 10.1016/j.jml.2012.02.007
- Scheerer, N. E., Tumber, A. K., and Jones, J. A. (2016). Attentional demands modulate sensorimotor learning induced by persistent exposure to changes in auditory feedback. *J. Neurophysiol.* 115, 826–832. doi: 10.1152/jn.00799.2015
- Schiffer, A. M., Waszak, F., and Yeung, N. (2015). The role of prediction and outcomes in adaptive cognitive control. *J. Physiol.* 109, 38–52. doi: 10.1016/j.jphysparis.2015.02.001
- Shell, A., Linck, J., and Slevc, L. R. (2015). “Examining the role of inhibitory control in bilingual language switching,” in *Proceedings of the 37th Annual Conference of the Cognitive Science Society*, ed. P. P. Magli (Austin, TX: Cognitive Science Society).
- Siegel, G. M., Pick, H. L., and Garber, S. R. (1984). Auditory feedback and speech development. *Adv. Child Dev. Behav.* 18, 49–79. doi: 10.1016/S0065-2407(08)60369-9
- Sorge, G., Toplak, M., and Bialystok, E. (2017). Interactions between levels of attention ability and levels of bilingualism in children’s executive functioning. *Dev. Sci.* 20:e12408. doi: 10.1111/desc.12408
- Stafford, T., and Gurney, K. N. (2011). Additive factors do not imply discrete processing stages: a worked example using models of the Stroop task. *Front. Psychol.* 2:287. doi: 10.3389/fpsyg.2011.00287
- Sternberg, S. (1969). The discovery of processing stages: extensions of Donders’ method. *Acta Psychol.* 30, 276–315. doi: 10.1016/0001-6918(69)90055-9
- Sternberg, S. (1998). “Discovering mental processing stages: The method of additive factors,” in *An Invitation to Cognitive Science. Methods, Models, and Conceptual Issues: An Invitation to Cognitive Science*, Vol. 4, eds D. Scarborough and S. Sternberg (Cambridge, MA: The MIT Press), 703–863.
- Sternberg, S. (2011). Modular processes in mind and brain. *Cogn. Neuropsychol.* 28, 156–208. doi: 10.1080/02643294.2011.557231
- Sternberg, S. (2013). The meaning of additive reaction-time effects: some misconceptions. *Front. Psychol.* 4:744. doi: 10.3389/fpsyg.2013.00744
- Stuart, A., Kalinowski, J., Rastatter, M. P., and Lynch, K. (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *J. Acoust. Soc. Am.* 111:2241. doi: 10.1121/1.1466868
- Tumber, A. K., Scheerer, N. E., and Jones, J. A. (2014). Attentional demands influence vocal compensations to pitch errors heard in auditory feedback. *PLoS One* 9:e109968. doi: 10.1371/journal.pone.0109968
- Van Borsel, J., Sunaert, R., and Engelen, S. (2005). Speech disruption under delayed auditory feedback in multilingual speakers. *J. Fluency Disord.* 30, 201–217. doi: 10.1016/j.jfludis.2005.05.001
- van den Wildenberg, W. P., and van der Molen, M. W. (2004). Additive factors analysis of inhibitory processing in the stop-signal paradigm. *Brain Cogn.* 56, 253–266. doi: 10.1016/j.bandc.2004.06.006
- Villacorta, V., Perkell, J., and Guenther, F. (2005). Relations between speech sensorimotor adaptation and perceptual acuity. *J. Acoust. Soc. Am.* 117, 2618–2619. doi: 10.1121/1.4778207
- Wickens, C. D. (2002). Multiple resources and performance prediction. *Theor. Issues Ergon. Sci.* 3, 159–177. doi: 10.1080/14639220210123806
- Wolf, M., and Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *J. Educ. Psychol.* 91, 415–438. doi: 10.1037/0022-0663.91.3.415

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Borrigan, Martin, de Bruin and Duñabeitia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX 1

Examples of matrices from the two different categories (animals and body parts) used in the experiment, together with the expected names to be said in Spanish and Basque (the English translations are as follows: toro, bull; rana, frog; oveja, sheep; caballo, horse; cerdo, pig; pájaro, bird; ojo, eye; mano, hand; boca, mouth; pie, foot; nariz, nose; oreja, ear) under the no replacement condition.





Symbiosis, Parasitism and Bilingual Cognitive Control: A Neuroemergentist Perspective

Arturo E. Hernandez*, Hannah L. Claussenius-Kalman, Juliana Ronderos and Kelly A. Vaughn

Department of Psychology, University of Houston, Houston, TX, United States

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Claudia C. von Bastian,
The University of Sheffield,
United Kingdom
Mark Antoniou,
Western Sydney University, Australia

*Correspondence:

Arturo E. Hernandez
aehernandez@uh.edu

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 29 June 2018

Accepted: 22 October 2018

Published: 19 November 2018

Citation:

Hernandez AE,
Claussenius-Kalman HL, Ronderos J
and Vaughn KA (2018) Symbiosis,
Parasitism and Bilingual Cognitive
Control: A Neuroemergentist
Perspective. *Front. Psychol.* 9:2171.
doi: 10.3389/fpsyg.2018.02171

Interest in the intersection between bilingualism and cognitive control and accessibility to neuroimaging methods has resulted in numerous studies with a variety of interpretations of the bilingual cognitive advantage. Neurocomputational Emergentism (or Neuroemergentism for short) is a new framework for understanding this relationship between bilingualism and cognitive control. This framework considers Emergence, in which two small elements are recombined in an interactive manner, yielding a non-linear effect. Added to this is the notion that Emergence can be captured in neural systems using computationally inspired models. This review poses that bilingualism and cognitive control, as examined through the Neuroemergentist framework, are interwoven through development and involve the non-linear growth of cognitive processing encompassing brain areas that combine and recombine, in symbiotic and parasitic ways, in order to handle more complex types of processing. The models that have sought to explain the neural substrates of bilingual cognitive differences will be discussed with a reinterpretation of the entire bilingual cognitive advantage within a Neuroemergentist framework incorporating its neural bases. It will conclude by discussing how this new Neuroemergentist approach alters our view of the effects of language experience on cognitive control. Avenues to move beyond the simple notion of a bilingual advantage or lack thereof will be proposed.

Keywords: bilingualism, cognitive control, development, computational models, language acquisition

INTRODUCTION

Bilinguals vary tremendously in the ways in which they learn two languages. Some learn a second language in childhood, adolescence, or adulthood while others may learn two languages in infancy. They may also vary in the ways in which these two languages are used with some involving formal schooling and others being mostly spoken languages. This naturally brings up the question of how a bilingual manages these two languages. To account for this, the notion of a language switch was proposed, adapting the notion of a switch first proposed by Penfield (1965) (Penfield and Roberts, 1959). The notion of a language switch was the inspiration for a greater number of studies across at least four decades (Hernandez, 2013). More recently, this debate has taken on a more modern nomenclature by considering the nature of cognitive control and its role in managing two languages.

Despite the long history of linking cognitive control to bilingualism, recent discussion has become contentious (Hilchey and Klein, 2011; Abutalebi et al., 2012; Hernández et al., 2013; Paap and Greenberg, 2013; Anton et al., 2014; Paap et al., 2014; de Bruin et al., 2015; Duñabeitia and Carreiras, 2015). Seminal studies in several labs have found that being exposed to two languages is associated with better performance on non-verbal cognitive control (Bialystok, 2007). Moving beyond the notion of language switching and its neural bases, work in the neuroimaging literature supports the view that bilingualism has the potential to strengthen frontal-striatal pathways due to the constant use of more than one language (Abutalebi et al., 2007, 2012; Green and Abutalebi, 2013; Stocco and Prat, 2014; Stocco et al., 2014) leading to advantages in cognitive control (Bialystok et al., 2004, 2007; Engel de Abreu et al., 2012; Bradley et al., 2013; Buchweitz and Prat, 2013; Marian et al., 2014). However, there is still considerable debate about whether the learning and use of two languages leads to a non-verbal cognitive control advantage in bilinguals (Hilchey and Klein, 2011; Abutalebi et al., 2012; Hernández et al., 2013; Paap and Greenberg, 2013; Anton et al., 2014; Paap et al., 2014; de Bruin et al., 2015; Duñabeitia and Carreiras, 2015). This has led many in the field to propose that we stop investigating the role that bilingualism may play in non-verbal cognitive control.

Despite the calls for completely abandoning the notion of a “bilingual advantage,” there is ample evidence that cognitive control and bilingualism are intimately linked. For example, work with older adults has found a posterior-to-anterior shift in normal aging (Davis et al., 2008; Dennis et al., 2008). The posterior-to-anterior shift is consistent with the view that older adults engage brain systems involved in cognitive control to a greater extent than young adults when performing the same task. Recent work in the bilingual literature finds the opposite pattern showing an anterior-to-posterior shift in bilinguals relative to monolinguals (Grant et al., 2014). The finding that bilingualism leads to the opposite, anterior-to-posterior (and subcortical) effect, has led researchers to consider whether bilingualism may serve as a cognitive protective factor (Grundy et al., 2017). The notion of a simple bilingual advantage has also led researchers to consider the variety of experiences including the age of acquisition, amount of language use, amount of switching, individual differences in language ability, and individual differences in flexibility and working memory that might influence the control mechanisms that are used in various non-verbal tasks (Green and Abutalebi, 2016; Yang et al., 2016; Birdsong, 2018). With all these different influences collectively pushing and pulling on the development of a bilingual’s two languages, a number of researchers have proposed that bilingualism has the characteristics of a non-linear dynamical system (Hernandez et al., 2005; Hernandez and Li, 2007; De Bot, 2008; Hernandez, 2013; Li et al., 2014).

In this piece, we build on this work using the notion of Neurocomputational Emergentism (Neuroemergentism) to provide a framework within which we can better understand the neural bases of cognitive control in bilinguals within a developmental framework. This approach seeks to consider Emergence, in which two small elements are recombined in an interactive manner, yielding a non-linear effect. Added to

this will be the notion that this Emergence can be captured in neural systems using computationally inspired models. While Neuroemergentism has been applied more directly to the content of language (i.e., grammar, semantics, etc.) it is likely that these diverging and converging influences affect cognitive control as well. We will begin by reviewing the field of Emergentism before proceeding to consider the nature of this within a computational neuroscience view that takes into account the developmental processes associated with the learning of two languages.

EMERGENTISM

The notion of emergent function has its roots in the work of Mill et al. (1974) who used examples from chemistry to argue that combining two simple things can lead to a much more complex form. For example, combining hydrogen and oxygen results in a new compound, water, due to the transformation of two gaseous elements into a liquid. The nature of dynamical changes seen in emergent forms was also considered by Bates et al. (1979); Bates (1999), who, inspired by the work of Thompson (1917) on the mathematics of physical transformation in the animal kingdom, proposed that language itself involved building “a new machine out of old parts.” The fact that old things may recombine into a newer whole can also be seen in other analogies used by Bates (1999). For example, one analogy included the giraffe’s neck, an adaptation to having to eat leaves high up in a tree that led to a number of cascading changes in the cardiovascular system and even the size and distribution of the hind and forelegs in order to conserve its balance. Thus, a small change could lead to changes across multiple different anatomical systems. This Emergentist view was further developed in the Competition Model, which sought to explain how grammar emerged from coalitions and rivalries between cues at one level that were then recombined to create a complex grammar that signaled who did what to whom at a much higher level (Bates and MacWhinney, 1981).

Work within the framework of the competition model considered the case of bilingualism. Bates and MacWhinney (1981) argued that a second language could potentially interact in a number of ways, including the bidirectional influences of first and second language grammatical processing termed forward and backward transfer, respectively. Amalgamation in which the grammars of both languages are fused together and differentiation in which each grammar is kept apart were also viable possibilities. Results from a series of studies on bilingual grammatical processing provided support for both forms of transfer as well as for amalgamation and differentiation.

In 2005, Hernandez et al. (2005) built on the Competition Model to propose an Emergentist framework that describes the acquisition of two languages. In that paper, the authors argued that through competitive interplay, children would learn to map form to meaning in each language and adjust it according the situation. For example, a child might learn to use a word form “taza” or “cup” depending on the language of a speaker. In a similar fashion, a bilingual model of vocabulary learning based on a model of vocabulary acquisition in a single language revealed that exposure to English and Mandarin at an early stage leads

to two independent representations (Zhao and Li, 2006). When the network was exposed to a second language (either Mandarin or English) at a later point in time, the second language became parasitic on the first. This competitive process is also reminiscent of Hernandez's (2013) view that a bilingual's two language are like two species in a single ecosystem that can compete for or share resources depending on the situation.

To date, the Emergentist view has mostly focused on the outcome of language processing, including the learning of new vocabulary as well as grammar, using a cognitive framework. However, this view can offer a framework with which to try and make sense of a newer question: whether bilinguals have some cognitive advantage relative to monolinguals. Although Emergentism does not make particular predictions about cognitive control, it can be extended to consider how the competitive process plays out across time resulting in the use of control to dampen it down. Furthermore, this framework also suggests that the nature of the competitive interplay across languages will differ depending on the age of acquisition of a second language.

Hernandez and Li (2007) were the first to propose a neurocomputational approach for the dynamic nature of change across time to explain age of acquisition effects across multiple domains. They suggested that learning earlier and later in life would rely on different neural and cognitive systems. Whereas learning early in life relied to a greater degree on neural systems involved in sensorimotor processing, later learning would rely on association areas which were used to bind sensory information or to combine it with motor processing. Computationally, early AoA effects led to distinct representations for each language whereas later learning resulted in a parasitic relationship in which the second language was built around the first (Zhao and Li, 2006).

One limitation of Hernandez and Li's (2007) Sensorimotor Hypothesis is that it might be mistakenly interpreted as suggesting that early learning only has an effect early in life. To overcome this static interpretation of age of acquisition, a newer framework, Neuroemergentism has been introduced. This framework is based on the idea that development involves the non-linear growth in brain areas that combine and recombine information in order to handle more complex types of processing. As such effects early in life can carry repercussions later in life. In the case of bilingualism, it is the case that early experiences can leave the neural substrate open to new experiences later in life. We will return to this point later after discussing the types of new questions that existing frameworks of cognitive control in bilingualism are opening up for researchers.

In the following sections, a summary and discussion of the bilingual advantage will be presented. A discussion of the models that have sought to elucidate the neural substrates of bilingual cognitive differences will also be discussed. The piece will end with a reinterpretation of the entire bilingual cognitive advantage within a Neuroemergentist framework incorporating its neural bases. It will conclude by discussing how this new Neuroemergentist approach alters our view of the effects of language experience on cognitive control. Avenues to move

beyond the simple notion of a bilingual advantage or lack thereof will be proposed.

THE BILINGUAL COGNITIVE ADVANTAGE: FINDINGS AND MODELS

The notion of a bilingual cognitive advantage began most recently in the mid-2000's when a number of research studies with different age groups found that bilinguals outperformed monolinguals across a variety of cognitive control tasks (Bialystok, 2006; Bialystok et al., 2004, 2007). Based on these findings, researchers concluded that the bilingual experience leads to an improvement in the ability to use executive control. In recent years, this claim has met considerable skepticism due to evidence both at the behavioral and neural levels (Morton and Harper, 2007; Paap, 2012; Kousaie et al., 2014; de Bruin et al., 2015). Given the variability in findings across studies, one is left wondering whether the advantage exists or not (for further discussion see Bialystok et al., 2015).

Research investigating the bilingual advantage is based to a greater extent on behavior; the brain mechanisms underlying potential differences in the use of cognitive control in bilinguals relative to monolinguals have not received nearly as much attention. Whether language experience results in behavioral differences or not, it is entirely possible that learning two languages results in differences at the neural level (Vaughn et al., 2015). Furthermore, researchers have begun to offer theoretical models that account for the effects that bilingual experience has on the cognitive control system, including the Bilingual Adaptation Model, the Adaptive Control Hypothesis (ACH) and the "Brain Training" model. Examining these three models in terms of their focus on cognitive changes or neural changes associated with bilingualism: the Bilingual Adaptation Model is based mostly on cognitive changes, the ACH is based on both cognitive and neural changes, and the Brain Training Model is based mostly on neural changes.

The Bilingual Adaptation Model, proposed by Bialystok (2017) suggests that being raised in a bilingual environment enables the development of a more flexible attention system associated with frontal brain regions. This model focuses on behavioral findings, such as gaze direction in infants and task performance in children, and suggests that, in general, the neural adaptations associated with executive attention in bilinguals overlap with language processing and selection. The Bilingual Adaptation Model includes aspects of Engle (2002) model of working memory capacity, as well as Posner and Petersen's (1990) sustained, selective, and executive attention networks.

The ACH, posited by Green and Abutalebi (2013), focuses on different bilingual environments and neurological and cognitive adaptations related to those environments. For example, the executive functions needed for interacting in a single language context are different from those needed to interact in a dual-language context and those needed to interact in a dense code-switching context. Three potential neural mechanisms may account for the improvements in executive function as a result of the language contexts: "through a change in structural resources

or capacity (e.g., gray matter density), through a change in regional efficiency (e.g., through tuning neuronal populations or changing the responsiveness of neuronal populations) or through a change in the connectivity of the network (e.g., white matter connectivity)” (p. 517). The specific networks altered by bilingualism depend on the cognitive functions being altered, which depend on the language context. Finally, the ACH highlights the importance of the frontal-striatal tract in learning two languages, and also considers other areas including the SMA, ACC, and inferior parietal lobule.

The importance of the striatum is the topic of the “brain-training” model (Stocco and Prat, 2014). Stocco et al. (2014) suggest that the basal ganglia is particularly well-equipped to handle bilingual language switching linking this phenomenon to earlier work with the conditional routing model (Stocco et al., 2012). During the learning process or in situations where automatized cognitive routines cannot accomplish a task (for example, in task-switching), the basal ganglia can amplify signals from a selected source region to enhance the likelihood that this signal will influence behavior despite weaker cortico-cortical network connections. This model also fits with more recent work that has begun to consider the role of dopamine and dopamine-related genes in cognitive flexibility and stability.

DEVELOPMENT: THE MISSING PIECE

Despite calls for researchers to take into account the variability in the bilingual experience, there has been a paucity of work on the developmental mechanism or mechanisms that might contribute to cognitive control. Hernandez and Li (2007), although not focused on cognitive control *per se*, provide a foundation for considering this topic. In their review of the literature, they noted that human development is characterized by successive waves of change that begin in areas that are devoted to sensorimotor processing, proceed to multisensory integration areas and end with the development of the prefrontal cortex which binds the senses and motor responses, paving the way for complex forms of cognition. One important prediction is that early learning involves greater potential for structuring the building blocks of a much simpler system which carries with it effects at a higher level. For example, age of acquisition effects, which have been linked to brain areas involved in phonological and prosodic processing, could then “spill over” to the processing of grammar. Age of acquisition effects in grammatical processing could be the outcome of the warping of sound space to particular combinations of both speech and prosodic patterns that contain grammatical information. Thus, the loss of plasticity in areas involved in sensorimotor processing can have pervasive effects beyond that level into other higher-level domains. As neural connections are solidified, plasticity is lost (for a more extensive discussion see Hernandez and Li, 2007).

Interestingly, this developmental view of sensorimotor integration during early ages also sheds light on the possible effect of early and later experience with a second language on cognitive control. Here we will offer two examples that elucidate the ways in which development can contribute to our understanding of

the relationship between cognitive control and the learning of two languages. We will base our discussion around the notion of an anterior to posterior/subcortical shift and its roots in a developmental model of bilingualism.

SYMBIOSIS, PARASITISM AND COGNITIVE CONTROL

Simultaneous bilinguals provide one of the most interesting test cases for theories of language development. Learners of two languages in infancy reveal utterances that are well organized and toddlers are able to adapt to the language output of an adult speaker (Genesee et al., 2008). At the same time, they can code-switch and produce utterances with words from both languages. Catch them early enough and infants exposed to two languages will even babble in each language, producing utterances that sound like one or the other language (Andruski et al., 2013). The brain must adapt to these two systems from the beginning of life.

How might these early language effects influence cognitive control? One mechanism proposed by Bialystok (2017) is that of executive attention, which bilingual infants use to focus on the mouths of speakers. Another mechanism would be at the speech categorization level. For example, Krizman and Marian (2015) suggest that the auditory system is intimately tied in with the executive control system via the basal ganglia up into cortical areas involved in cognitive control including the anterior cingulate cortex (ACC) and the dorsolateral prefrontal cortex (DLPFC). There are also feedback loops from the ACC back down toward the basal ganglia. If one considers an individual who learned two languages from a young age, it is likely that adaptations at the level of the basal ganglia would have happened due to the exposure to two distinct phonological and prosodic patterns for each language. Thus, simultaneous bilingualism would lead to a distinct set of neural signatures in subcortical structures and in the basal ganglia in particular. Early childhood bilingualism might involve attentional areas needed to distinguish speakers of each language and to map space onto language use. Later second language acquisition would involve the prefrontal cortex to a greater extent, in order to overcome the preponderant responses of the first language. In this view, the pre-frontal cortex would become involved in order to overcome an entrenched second language that might be parasitic on the first language. This competitive and cooperative view of language development at various ages takes us back to the notion of development as being characterized by cascading interactions at various neural levels, the hallmark of a Neuroemergentist approach.

One final piece of this puzzle is to consider neurocomputational models and how these might fit in with a Neuroemergentist approach. Only one model of bilingual cognitive control, Stocco’s “brain training” model has a neurocomputational implementation (Stocco et al., 2014). One interesting aspect is that Stocco’s model is not framed around cognitive control *per se* but rather around reinforcement learning. In this vein, the basal ganglia are recruited during the learning process or in when automatized cognitive routines are

unable to successfully complete a task (for example, in task-switching). In these cases, the basal ganglia can amplify signals from a selected source region to enhance the likelihood that this signal will influence behavior despite weaker cortico-cortical network connections.

That suggests that language history plays a role in learning was conducted by Bradley et al. (2013). In that study, a group of Spanish-English bilinguals and monolinguals were asked to learn a set of “new” German words via translation. After reaching 90% correct, they were placed in the scanner and asked to make a living/non-living judgment. The results revealed better performance in the bilinguals in that they had lower reaction times relative to monolinguals. In addition, bilinguals showed increased activity in the putamen whereas monolinguals showed relatively greater activity in the caudate nucleus and cortical cognitive control areas. All of our bilingual subjects had learned English relatively early in life. Thus, these results are consistent with the view that an early age of second language acquisition leads to neural adaptations when learning a new task. Together with Stocco’s neurocomputational model they would suggest that age of acquisition is not only affecting the content of cognition, it may be also affecting the way in which the brain handles new learning. Future studies could use a combination of brain and computational science to further examine this question.

A FINAL NOTE ON NON-LINEAR DYNAMICS

So far, our Neuroemergentist view, which has focused on the sensorimotor aspect of early learning which differs from later learning, suggests that signatures would include a dynamic interaction between an individual and his or her environment. The nature of this dynamic Neuroemergentist view can be seen in the surprising result from a study on biomarkers and behavioral responses in a group of middle-age monolingual and bilingual participants in which early bilingualism was associated with better performance on tasks of executive function (Estanga et al., 2017). The early bilingual group was also found to have a lower presence of t-tau levels in their cerebrospinal fluid (CSF). The presence of a biomarker in the CSF is an unlikely place to search for an effect of early bilingualism. However, this biomarker is one that has received considerable attention in the literature. One area that is subject to tau-pathology (Grudzien et al., 2007), the locus coeruleus, has been posited to play a role in cognitive reserve via noradrenergic stimulation (Mather and Harley, 2016). Furthermore, the locus coeruleus, tyrosine-hydroxylase-expressing (THb) neurons have been found to mediate post-encoding memory enhancement, possibly through the release of dopamine in the hippocampus (Takeuchi et al., 2016), another structure of crucial importance for Alzheimer’s disease.

The fact that changes in the CSF are associated with language learning history opens up two questions. The first revolves around age of acquisition, a key topic that was discussed at length in previous Neuroemergentist pieces (Hernandez et al., 2005; Hernandez, 2013; Hernandez and Li, 2007;

Hernandez et al., In press). If early acquisition of two languages involves earlier-developing neural systems, then one is left wondering whether cognitive reserve in older adults is due to the continued use of two languages across a long period of time or due to the age at which second language exposure began. The fact that this effect differed between two groups of middle-age individuals with significant length of exposure and use of two languages is consistent with the view that that age of acquisition and not use of both languages may be leading to greater cognitive reserve. This fits in with the view that learning early in life is much more embodied, a view that has been found in a number of domains both between, within, and outside of language (Hernandez and Li, 2007; Hernandez et al., 2011).

Recently, we have proposed the term Neuroemergentism to suggest that these non-linear interactions occur at the level of the brain. To be clear, other researchers have suggested similar approaches. This Neuroemergentist approach does share features with Neuroconstructivism (Karmiloff-Smith, 2009, 2015) and the Interactive Specialization (Johnson, 2011) approaches that have been proposed by Karmiloff-Smith, Johnson and colleagues. However, Neuroemergentism emphasizes the appearance of effects from the combination of much smaller systems. This fits with the examples given earlier. Namely, that low level interactions in the auditory system of early bilinguals as well as possible effects of arousal, attention and/or memory handled by the locus coeruleus in concert with a wide range of other brain areas, would lead to the presence of a pervasive effect that lasted into middle age and beyond. The question remains of what late bilingualism brings to the table. Here, our contention is that these effects are more likely to appear in cortical systems and work their way down to subcortical systems, leading to significant rewiring (Li et al., 2014).

One question that remains with regard to Neuroemergentism is what tangible alternatives it might offer compared to other frameworks. First, because of its emphasis on the non-linear dynamics of change, it takes into account developmental variables that have not been traditionally considered in the ACH of Abutalebi and Green as well as the Executive Attention proposal of Bialystok. These developmental effects are likely the product of varying forms of bilingual experience that could be studied as well (Antoniou, 2018). Both of these generally do not consider how development might contribute to cognitive control. Second, our framework seeks to look at how cognitive control is built up across time. This non-linear process has potential weakness in that it could be seen as explaining any particular non-linearity that might appear using the veil of Emergentism an *aposteriori* explanation. To overcome this limitation, Neurocomputational Emergentism, actually seeks to find specific causes for emergent behavior due to neural reorganization. One specific example would involve the use of the “brain training” model of the basal ganglia as proposed by Stocco and colleagues. It could be adapted to handle linguistic input during development. Additional experiments with children could then be used to look at the effects of second or dual-language learning on cognitive control. This would in turn generate new predictions. This cycle of experimental observation of emergent behavior, modeling of that behavior and new predictions would eventually allow a more

complete view of what the causes of any particular advantage in cognitive control might be. Although this framework is in an early phase of development at this time, the issues discussed in this piece point to the benefits of a Neuroemergentist approach to this question.

However, rather than seeing bilingualism as a unique contributor, within this Neuroemergentist view we would argue that it serves as a window within which researchers can observe the interconnectedness of neural systems that are thought to be dissociated. In short, experience with two languages does more than alter the neural substrate responsible for language. Rather, the effects of language experiences are likely pervasive, and not easily reduced to a simple reaction time advantage or a single pattern of brain activity. In addition, the case of adoptees goes beyond a simple bilingual/monolingual dichotomy. For example, neural traces have been observed in adults even when there is no conscious knowledge of a language that was discontinued very early in childhood (Pierce et al., 2015). Future endeavors should continue to use traditional methods in behavioral and brain science along with computational models that can handle the type of non-linear interactions seen in development. The remaining question, thus, should not center on whether language experience at different points in life affects or does not affect cognition. The question is how it does. A question that with the help of dynamic developmental frameworks such

as Neuroemergentism is likely to keep researchers occupied for years to come.

AUTHOR CONTRIBUTIONS

AH contributed to the introduction and conclusion as well as with the overall organization of the entire manuscript. HC-K contributed to discussion of bilingualism and its relationship with attention and manuscript formatting. JR contributed to the discussion of Neuroemergentism in the context of bilingual language development and manuscript formatting. KV contributed to the overview of the bilingual cognitive advantage and assisted with formatting the manuscript according to Frontiers standards.

FUNDING

This research was supported in part by grant # 5R03HD079873-02, Effects of genetic differences and bilingual status on cognitive control, to AH and P50 HD052117, Texas Center for Learning Disabilities, from the Eunice Kennedy Shriver National Institute of Child Health and Human Development to the University of Houston.

REFERENCES

- Abutalebi, J., Annoni, J. M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., et al. (2007). Language control and lexical competition in bilinguals: an event-related fMRI study. *Cereb. Cortex* 18, 1496–1505. doi: 10.1093/cercor/bhm182
- Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., et al. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cereb. Cortex* 22, 2076–2086. doi: 10.1093/cercor/bhr287
- Andruski, J. E., Caselles, E., and Nathan, G. (2013). Is bilingual babbling language-specific? Some evidence from a case study of Spanish–English dual acquisition. *Biling. Lang. Cogn.* 17, 660–672. doi: 10.1017/S1366728913000655
- Anton, E., Dunabeitia, J. A., Estevez, A., Hernandez, J. A., Castillo, A., Fuentes, L. J., et al. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Front. Psychol.* 5:398. doi: 10.3389/fpsyg.2014.00398
- Antoniou, M. (2018). The advantages of bilingualism debate. *Annu. Rev. Linguist.* 5. doi: 10.1146/annurev-linguistics-011718-011820
- Bates, E. A. (1999). “Plasticity, localization, and language development,” in *The Changing Nervous System: Neurobehavioral Consequences of Early Brain Disorders*, eds S. H. Broman, J. M. Fletcher, S. H. Broman, and J. M. Fletcher (New York, NY: Oxford University Press), 214–253.
- Bates, E. A., Benigni, L., Bretherton, I., Camaioni, L., and Volterra, V. (1979). *The Emergence of Symbols: Cognition and Communication in Infancy*. New York, NY: Academic Press.
- Bates, E. A., and MacWhinney, B. (1981). “Second language acquisition from a functionalist perspective: pragmatic, semantic, and perceptual strategies,” in *Annals of the New York Academy of Sciences Conference on Native and Foreign Language Acquisition*, ed. H. Winitz (New York, NY: New York Academy of Sciences).
- Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Can. J. Exp. Psychol.* 60, 68–79. doi: 10.1037/cjep.2006008
- Bialystok, E. (2007). Cognitive effects of bilingualism: how linguistic experience leads to cognitive change. *Int. J. Biling. Educ. Biling.* 10, 210–223. doi: 10.2167/beb441.0
- Bialystok, E. (2017). The bilingual adaptation: how minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099
- Bialystok, E., Craik, F. I. M., and Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia* 45, 459–464. doi: 10.1016/j.neuropsychologia.2006.10.009
- Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290
- Bialystok, E., Kroll, J. F., Green, D. W., MacWhinney, B., and Craik, F. I. M. (2015). Publication bias and the validity of evidence: what’s the connection? *Psychol. Sci.* 26, 944–946. doi: 10.1177/0956797615573759
- Birdsong, D. (2018). Plasticity, variability and age in second language acquisition and bilingualism. *Front. Psychol.* 9:81. doi: 10.3389/fpsyg.2018.00081
- Bradley, K. A., King, K. E., and Hernandez, A. E. (2013). Language experience differentiates prefrontal and subcortical activation of the cognitive control network in novel word learning. *Neuroimage* 67, 101–110. doi: 10.1016/j.neuroimage.2012.11.018
- Buchweitz, A., and Prat, C. (2013). The bilingual brain: flexibility and control in the human cortex. *Phys. Life Rev.* 10, 428–443. doi: 10.1016/j.plrev.2013.07.020
- Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., and Cabeza, R. (2008). Que PASA? The posterior-anterior shift in aging. *Cereb. Cortex* 18, 1201–1209. doi: 10.1093/cercor/bhm155
- De Bot, K. (2008). Introduction: second language development as a dynamic process. *Mod. Lang. J.* 92, 166–178. doi: 10.1111/j.1540-4781.2008.00712.x
- de Bruin, A., Treccani, B., and Della Sala, S. (2015). Cognitive advantage in bilingualism: an example of publication bias? *Psychol. Sci.* 26, 99–107. doi: 10.1177/0956797614557866
- Dennis, N. A., Hayes, S. M., Prince, S. E., Madden, D. J., Huettel, S. A., and Cabeza, R. (2008). Effects of aging on the neural correlates of successful item and source memory encoding. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 791–808. doi: 10.1037/0278-7393.34.4.791
- Dunabeitia, J. A., and Carreiras, M. (2015). The bilingual advantage: acta est fabula? *Cortex* 73, 371–372. doi: 10.1016/j.cortex.2015.06.009
- Engel de Abreu, P. M. J., Cruz-Santos, A., Tourinho, C. J., Martin, R., and Bialystok, E. (2012). Bilingualism enriches the poor: enhanced cognitive control in low-income minority children. *Psychol. Sci.* 23, 1364–1371. doi: 10.1177/0956797612443836

- Engle, R. W. (2002). Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11, 19–23. doi: 10.1111/1467-8721.00160
- Estanga, A., Ecay-Torres, M., Ibañez, A., Izaguirre, A., Villanua, J., Garcia-Sebastian, M., et al. (2017). Beneficial effect of bilingualism on Alzheimer's disease CSF biomarkers and cognition. *Neurobiol. Aging* 50, 144–151. doi: 10.1016/j.neurobiolaging.2016.10.013
- Genesee, F., Nicoladis, E., and Paradis, J. (2008). Language differentiation in early bilingual development. *J. Child Lang.* 22, 611–631. doi: 10.1017/S0305000900009971
- Grant, A., Dennis, N. A., and Li, P. (2014). Cognitive control, cognitive reserve, and memory in the aging bilingual brain. *Front. Psychol.* 5:1401. doi: 10.3389/fpsyg.2014.01401
- Green, D. W., and Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377
- Green, D. W., and Abutalebi, J. (2016). Language control and the neuroanatomy of bilingualism: in praise of variety. *Lang. Cogn. Neurosci.* 31, 340–344. doi: 10.1080/23273798.2015.1084428
- Grudzien, A., Shaw, P., Weintraub, S., Bigio, E., Mash, D. C., and Mesulam, M. M. (2007). Locus coeruleus neurofibrillary degeneration in aging, mild cognitive impairment and early Alzheimer's disease. *Neurobiol. Aging* 28, 327–335. doi: 10.1016/j.neurobiolaging.2006.02.007
- Grund, J. G., Anderson, J. A. E., and Bialystok, E. (2017). Bilinguals have more complex EEG brain signals in occipital regions than monolinguals. *Neuroimage* 159, 280–288. doi: 10.1016/j.neuroimage.2017.07.063
- Hernandez, A. E. (2013). *The Bilingual Brain*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780199828111.001.0001
- Hernandez, A. E., Claussenius-Kalman, H. L., Ronderos, J., Castilla-Earls, A. P., Sun, L., Weiss, S. D., et al. (in press). Neuro emergentism: a framework for studying cognition and the brain. *J. Neurolinguist.*
- Hernandez, A. E., and Li, P. (2007). Age of acquisition: its neural and computational mechanisms. *Psychol. Bull.* 133, 638–650. doi: 10.1037/0033-2909.133.4.638
- Hernandez, A. E., Li, P., and MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends Cogn. Sci.* 9, 220–225. doi: 10.1016/j.tics.2005.03.003
- Hernandez, A. E., Mattarella-Micke, A., Redding, R. W. T., Woods, E. A., and Beilock, S. (2011). Age of acquisition in sport: starting early matters. *Am. J. Psychol.* 124, 253–260. doi: 10.5406/amerjpsyc.124.3.0253
- Hernández, M., Martin, C. D., Barceló, F., and Costa, A. (2013). Where is the bilingual advantage in task-switching? *J. Mem. Lang.* 69, 257–276. doi: 10.1016/j.jml.2013.06.004
- Hilchey, M., and Klein, R. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychon. Bull. Rev.* 18, 625–658. doi: 10.3758/s13423-011-0116-7
- Johnson, M. H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Dev. Cogn. Neurosci.* 1, 7–21. doi: 10.1016/j.dcn.2010.07.003
- Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: rethinking the study of developmental disorders. *Dev. Psychol.* 45, 56–63. doi: 10.1037/a0014506
- Karmiloff-Smith, A. (2015). An alternative to domain-general or domain-specific frameworks for theorizing about human evolution and ontogenesis. *AIMS Neurosci.* 2, 91–104. doi: 10.3934/Neuroscience.2015.2.91
- Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., and Taler, V. (2014). Executive function and bilingualism in young and older adults. *Front. Behav. Neurosci.* 8:250. doi: 10.3389/fnbeh.2014.00250
- Krizman, J., and Marian, V. (2015). “Neural consequences of bilingualism for cortical and subcortical function,” in *The Cambridge Handbook of Bilingual Processing*, ed. J. W. Schwieter (Cambridge: Cambridge University Press).
- Li, P., Legault, J., and Litcosky, K. A. (2014). Neuroplasticity as a function of second language learning: anatomical changes in the human brain. *Cortex* 58, 301–324. doi: 10.1016/j.cortex.2014.05.001
- Marian, V., Chabal, S., Bartolotti, J., Bradley, K., and Hernandez, A. E. (2014). Differential recruitment of executive control regions during phonological competition in monolinguals and bilinguals. *Brain Lang.* 139, 108–117. doi: 10.1016/j.bandl.2014.10.005
- Mather, M., and Harley, C. W. (2016). The locus coeruleus: essential for maintaining cognitive function and the aging brain. *Trends Cogn. Sci.* 20, 214–226. doi: 10.1016/j.tics.2016.01.001
- Mill, J. S., Robson, J. M., McRae, R. F., Routledge, Kegan, P., and University of Toronto press (1974). *Collected Works of John Stuart Mill*, Vol. 7. Toronto: University of Toronto press, 7.
- Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x
- Paap, K. R. (2012). No coherent evidence for a bilingual advantage in executive processing. *Can. J. Exp. Psychol.* 66, 232–258.
- Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Paap, K. R., Johnson, H. A., and Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *J. Cogn. Psychol.* 26, 615–639. doi: 10.1080/20445911.2014.944914
- Penfield, W. (1965). Conditioning the uncommitted cortex for language learning. *Brain* 88, 787–798. doi: 10.1093/brain/88.4.787
- Penfield, W., and Roberts, L. (1959). *Speech and Brain Mechanisms*. Princeton, NJ: Princeton University Press.
- Pierce, L. J., Chen, J.-K., Delcenserie, A., Genesee, F., and Klein, D. (2015). Past experience shapes ongoing neural patterns for language. *Nat. Commun.* 6:10073. doi: 10.1038/ncomms10073
- Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42. doi: 10.1146/annurev.ne.13.030190.000325
- Stocco, A., Lebiere, C., O'Reilly, R. C., and Anderson, J. R. (2012). Distinct contributions of the caudate nucleus, rostral prefrontal cortex, and parietal cortex to the execution of instructed tasks. *Cogn. Affect. Behav. Neurosci.* 12, 611–628. doi: 10.3758/s13415-012-0117-7
- Stocco, A., and Prat, C. S. (2014). Bilingualism trains specific brain circuits involved in flexible rule selection and application. *Brain Lang.* 137, 50–61. doi: 10.1016/j.bandl.2014.07.005
- Stocco, A., Yamasaki, B., Natalenko, R., and Prat, C. S. (2014). Bilingual brain training: a neurobiological framework of how bilingual experience improves executive function. *Int. J. Biling.* 18, 67–92. doi: 10.1177/1367006912456617
- Takeuchi, T., Duzkiewicz, A. J., Sonneborn, A., Spooner, P. A., Yamasaki, M., Watanabe, M., et al. (2016). Locus coeruleus and dopaminergic consolidation of everyday memory. *Nature* 537, 357–362. doi: 10.1038/nature19325
- Thompson, D. A. W. (1917). *On Growth and Form*. Cambridge: Cambridge University press.
- Vaughn, K. A., Greene, M. R., Ramos Nuñez, A. I., and Hernandez, A. E. (2015). The importance of neuroscience in understanding bilingual cognitive control. *Cortex* 73, 373–374. doi: 10.1016/j.cortex.2015.06.010
- Yang, H., Hartanto, A., and Yang, S. (2016). The complex nature of bilinguals' language usage modulates task-switching outcomes. *Front. Psychol.* 7:560. doi: 10.3389/fpsyg.2016.00560
- Zhao, X., and Li, P. (2006). “A self-organizing connectionist model of bilingual lexical development,” in *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, (Mahwah, NJ: Lawrence Erlbaum).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Hernandez, Claussenius-Kalman, Ronderos and Vaughn. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Multilingual Language Control and Executive Function: A Replication Study

Gregory J. Poarch*

Department of English Linguistics, University of Münster, Münster, Germany

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Arturo Hernandez,
University of Houston, United States
Mathieu Declerck,
Aix-Marseille Université, France

*Correspondence:

Gregory J. Poarch
g.poarch@gmx.net

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 31 July 2018

Accepted: 27 September 2018

Published: 26 October 2018

Citation:

Poarch GJ (2018) Multilingual
Language Control and Executive
Function: A Replication Study.
Front. Commun. 3:46.
doi: 10.3389/fcomm.2018.00046

Recent discussion has called into question whether navigating and controlling multiple languages in daily life influences the development of executive function. Given the dearth in replications of studies that have documented differences in executive function between multilingual and monolingual children, the present study replicates a study on executive function in children (Poarch and Van Hell, 2012a) with a child population from the same educational and socio-economic background. Two executive function tasks (Simon and Flanker) were administered to 163 children aged 5–13 years who were either monolingual second language (L2) learners of English or multilinguals [German-English bilinguals or German-Language X bilingual third language (L3) learners of English]. While the Simon task yielded no differences between groups, the Flanker task differed significantly across groups with multilinguals showing enhanced conflict resolution over L2 learners. While the children's performance on the two tasks yielded diverging results, the outcome is partially in line with the view that enhanced executive function in multilingual children arises from their permanent need to monitor, control, and shift between multiple languages. These findings are discussed against the backdrop of varying inhibitory processes invoked by the specific nature of the two tasks and of developmental trajectories of executive function.

Keywords: executive function, simon task, flanker task, second language learners, bilinguals, third language learners

INTRODUCTION

There is a growing body of research documenting that children who grow up with and regularly use multiple languages exhibit differential non-verbal executive function compared to children who only grow up and use one language. Such differences between multilingual and monolingual children are assumed to be linked to the lifelong multilingual experience of having to control and use multiple languages in daily life (for reviews, see Bialystok et al., 2012; Kroll and Bialystok, 2013; Baum and Titone, 2014; Valian, 2015; Bialystok, 2017; Poarch and Van Hell, 2017; Poarch, 2018). While there is ample experimental evidence in support of the notion that sustained and long-term multilingual experience positively affects executive function development in children (e.g., Carlson and Meltzoff, 2008; Poarch and Van Hell, 2012a; Poarch and Bialystok, 2015), there are now also studies that have yielded no executive function differences between multilingual and monolingual children (Duñabeitia et al., 2013; Antón et al., 2014; Gathercole et al., 2014).

Given these mixed findings, and in order to move forward in addressing the question of whether speaking more than one language on a regular basis indeed impacts the development of executive function, there is a need to replicate previous findings of executive function differences between groups in similar populations of children, and in doing so to possibly identify more specifically under which conditions multilingual children profit from their language control experience, and to assess whether the specific experimental measures used so far in the research field are unequivocally appropriate to adequately tap executive function processes. The present study attempts to address these issues by closely replicating a published study (Poarch and Van Hell, 2012a) with a very similar population from the same environment (extended to a larger age range) and using the same types of experimental measures. Such an approach is also warranted in light of the limited reproducibility of research in psychological science (Open Science Collaboration, 2015).

Executive Function and Multilingualism

Our cognitive system is geared toward making choices in daily life between alternative and competing responses (cf. Keye et al., 2009). The mechanism responsible for detecting situations in which such conflicting information is present, needs to be processed, and subsequently resolved is subsumed under the so-called executive function system. This system incorporates cognitive functions such as selective attention, updating information, shifting between sets of information, and monitoring for and resolving conflict (see, e.g., Botvinick et al., 2001; Engle, 2002; Miyake and Friedman, 2012; Diamond, 2013) and develops from early childhood until it reaches maturity during adolescence (Anderson, 2002). The theoretical basis of multilingualism affecting domain-general non-verbal cognitive processing is grounded in the finding that the processes subserving multilingual language control and non-verbal cognitive control show extensive overlap (Declerck et al., 2017; but see Calabria et al., 2015; Branzi et al., 2016, for evidence of less overlap) and that multilinguals need to cognitively control multiple competing languages and are exposed to nearly constant cross-language activation and interaction (e.g., during lexical processing; Thierry and Wu, 2007; Poarch and Van Hell, 2012b). Such control processes, which are also drawn on during bilingual language processing (e.g., Filippi et al., 2015) or when switching from one language to another (e.g., Anderson et al., 2018a), induce repetitive cognitive load that over time impacts the neural networks responsible for and subserving executive function (e.g., Calabria et al., 2018). These processes are also assumed to influence the development and efficacy of executive function (see Green and Abutalebi, 2013; for comprehensive reviews, see Bialystok, 2017; Antoniou, 2019).

There are numerous studies that have reported executive function differences between groups of multilingual and monolingual children matched on a variety of language and social background variables (e.g., Carlson and Meltzoff, 2008; Engel de Abreu et al., 2012; Morales et al., 2013; Blom et al., 2014, 2017; Ladas et al., 2015; Poarch and Bialystok,

2015; Crivello et al., 2016; De Cat et al., 2018; Thomas-Sunesson et al., 2018; for a review of research with children, see Poarch and Van Hell, 2017). The study most relevant to the present study, and the one described in detail at this point, is that by Poarch and Van Hell (2012a) who administered two executive function tasks (the Simon task and a variant of the Flanker task) to four groups of children (monolinguals, L2 learners, bilinguals, and trilinguals) aged 5–8. Bilinguals and trilinguals were defined as children who regularly used multiple languages (see Surraín and Luk, 2017, for how bilinguals are characterized in the literature). The study aimed to extend previous research that had compared only monolingual and bilingual children and to investigate executive function in children who were matched on proficiency in their first language (L1), on socio-economic status, while differing on language backgrounds and proficiency in their second language (L2). In the Simon task (Experiment 1), bilinguals and trilinguals showed significantly faster conflict resolution than monolinguals, and marginally so than L2 learners. Furthermore, bilinguals and trilinguals did not differ in their performance, and L2 learners and monolinguals did not differ either. The performance in the Flanker-type task yielded similar results, with bilinguals and trilinguals outperforming L2 learners in resolving conflict induced by the incongruent condition (see Description of tasks and measures below). Note that there was no monolingual participant group in Experiment 2. These findings were interpreted as indicating enhanced inhibitory control for bilinguals and trilinguals over L2 learners (and monolinguals in Experiment 1) stemming from the necessity for multilingual children to control their developing and interacting languages. Training language control processes regularly and repeatedly may boost the multilingual children's shifting of attention, task monitoring, and conflict resolution in these tasks. Alternatively, it may also modulate the impact of distracting information during task performance.

However, as indicated above, there are studies reporting no differences between multilingual and monolingual children in executive function task performance (Duñabeitia et al., 2013; Antón et al., 2014; Gathercole et al., 2014; Ross and Melinger, 2017). As such, the latter studies can be seen to challenge the assumption that multilingualism has an effect on the development of executive function and have fuelled the discussion on whether and how multilingual language experience can impact executive function (see, e.g., Poarch and Van Hell, 2017), and whether the executive function tasks used are ideally equipped to measure the efficacy of the executive function system (see Valian, 2015; Poarch and Van Hell, in press).

Description of Tasks and Measures

There are a number of experimental paradigms that tap non-verbal cognitive processes, two of which have been used ubiquitously in the field of research on multilingualism and executive function: the Eriksen Flanker task (1974) and the Simon task (Simon and Rudell, 1967). Both tasks are thought to induce cognitive conflict during task performance, requiring selective attention to identify conflict and subsequent cognitive resources for conflict resolution (see, e.g., Hommel, 2011;

Wöstmann et al., 2013), albeit in slightly different manners. While the Flanker task uses arrays of arrows that are either congruent or incongruent to measure resistance to the interference of flanking distractors (Friedman and Miyake, 2004), the Simon task uses colored squares to induce conflict by a spatial stimulus-response mismatch in incongruent trials compared to an absence of a mismatch in congruent trials.

Note that in Poarch and Van Hell (2012a) a modified and more elaborate version of the Flanker task was used, the Attentional Networks Task (ANT; Fan et al., 2002; Rueda et al., 2004). In essence, the ANT is a Flanker task (with the customary inhibitory control component that requires inhibiting distractors) with added executive function components, namely alerting and orienting. However, these additional components are disregarded in the present study in order to focus on the main question at hand as to whether multilinguals and monolinguals differ in conflict monitoring and inhibitory control.

In both tasks, beyond inspecting overall reaction times in the congruent and incongruent conditions, a difference score as an index of inhibitory control is calculated (the congruent condition reaction time subtracted from the incongruent condition reaction time). The difference score magnitude indicates how strongly distracted individuals are in the incongruent condition compared to the congruent condition. A larger magnitude indexes poorer interference control (for a more detailed account of how performance in these tasks can be modeled, see Botvinick et al., 2001; Keye et al., 2009).

Finally, since the Simon task and the Flanker task are used to tap participants' conflict monitoring and inhibitory control, performance on both tasks can be expected to correlate positively. For conflict monitoring, overall processing speed across the tasks should correlate, for inhibitory control, performance on incongruent trials and the difference score should correlate across tasks (see, e.g., Keye et al., 2009; Wöstmann et al., 2013). In contrast, if there is no correlation across task performance, the tasks may not entirely tap the same executive function components (see Fan et al., 2003; Valian, 2015).

To the author's knowledge, there are only two studies with children that have correlated performance across the two tasks of interest (Ross and Melinger, 2017; Poarch and Van Hell, in press). Ross and Melinger (2017) found child bilinguals, bidialectals, and monolinguals to not differ on overall performance and the calculated difference scores in the Simon and Flanker tasks. Critically, the congruent and incongruent reaction times correlated significantly across tasks indicating convergent validity. The difference scores indexing inhibitory control were, however, not analyzed separately. In contrast, Poarch and Van Hell (in press) re-analyzed the data from their original study (Poarch and Van Hell, 2012a) and found that neither congruent and incongruent conditions nor the difference score correlated across tasks, which in turn calls the convergent validity across tasks into question (see also Paap and Greenberg, 2013). The inconsistent convergent validity for these executive function tasks (see also Keye et al., 2009) indicates that one (or both) of the tasks may not fully or only partially measure the efficacy of the conflict monitoring mechanism.

The Present Study

The main objective of the present study is to replicate previous work by Poarch and Van Hell (2012a), using the same task types and experimental set-up (Simon and Flanker), with a very similar population from the same environment (second language learners, bilinguals and bilingual third language learners), and in an extension, to also focus on children from a wider age range. Bilinguals and bilingual third-language learners are defined here as regular users of either two or three languages (Surrain and Luk, 2017). Both groups of children have been found to exhibit similar effects on executive function development compared to monolingual children, irrespective of the number of languages controlled on a daily basis (Poarch and Van Hell, 2012a; Poarch and Bialystok, 2015). Accordingly, similar executive function task performance by bilinguals and bilingual third language learners was expected. Hence, for the purpose of the present study, in the initial analyses the two groups were collapsed into a single group of multilinguals (subsequently, the two groups were also analyzed separately to confirm that their performance was indeed similar) and the following predictions were made:

- 1) If the cognitive effort of constantly controlling multiple languages has an effect on executive function development, then executive function task performance should differ between monolingual second language learners with little language control experience and multilinguals with more extensive language control experience.
- 2) The difference in performance between multilinguals and second language learners could be displayed by: (a) better overall performance by multilingual compared to second-language learners, which would amount to more efficient selective attention and task monitoring in executive function tasks in which participants are faced with congruent and incongruent stimuli. Such a performance difference has been interpreted as enhanced general-domain executive function in multilinguals (e.g., Martin-Rhee and Bialystok, 2008; Yang et al., 2011; Kapa and Colombo, 2013); and/or (b) a smaller difference between performance on congruent and incongruent stimuli, yielding a smaller difference score magnitude and better inhibitory control for multilinguals, which would amount to enhanced domain-specific executive function (e.g., Engel de Abreu et al., 2012; Poarch and Van Hell, 2012a; Poarch and Bialystok, 2015; Yang and Yang, 2016).

Furthermore, employing two tasks ubiquitously used in past research to tap cognitive processing in bilingual and monolingual children allowed for correlational analyses of the children's performance on the two tasks. Note that only very few studies so far have used these executive function tasks in children and, critically, have subsequently correlated task performance (Ross and Melinger, 2017; Poarch and Van Hell, in press).

MATERIALS AND METHODS

Participants

Participants were 163 children, 5- to 13-years old, who attended private primary and secondary German-English immersion

schools in Frankfurt, Germany. Four children were excluded due to incomplete data sets and/or background information. Thus, of the remaining 159 children, 77 children were German monolingual second language learners of English (henceforth L2 learners; 43 girls), 34 German-English bilinguals (12 girls), and 48 German-Language X third-language learners of English (henceforth L3 learners; 30 girls). The children's mean age was 9.7 years ($SD = 2.3$; range = 5.2–13.3 years).

Signed consent was provided by the children's parents¹, who also completed an earlier version of the Language and Social Background Questionnaire (LSBQ; Anderson et al., 2018b), in which the home language environment and proficiency in each language is assessed. The L2 learners were all native speakers of German and had been learning English for an average of 1.8 years ($SD = 1.5$). The bilingual children lived in homes in which German and English were the primary languages, with German being the main language outside the home, and German and English used at school. They had been learning English in educational contexts for an average of 3.0 years ($SD = 1.6$). The L3 learners spoke two languages at home (one of which being German), German and English at school, and had been learning English for an average of 1.8 years ($SD = 1.5$). The home languages spoken apart from German included Arabic (5), Croatian (2), Danish (2), Dutch (3), Eritrean (1), Greek (2), Hebrew (2), Hindi (1), Italian (4), Japanese (3), Lithuanian (1), Polish (3), Portuguese (2), Russian (4), Serbian (1), Spanish (5), Swedish (2), Turkish (3), Urdu (1), Vietnamese (1).

Parents were asked to rate their children's daily language usage on a set of 5-point scales that extended from "All German" (0) to "Only other language" (4). An average score of 2 indicates that home communication was divided equally between German and other languages. The mean score across these scales for L2 learners was 0.7 ($SD = 0.5$), for bilinguals it was 2.1 ($SD = 1.1$), and for L3 learners it was 1.9 ($SD = 0.9$), indicating that the monolinguals' homes functioned primarily in German, while those of the bilinguals and L3 learners showed a more balance use of German and English or German and another language, $F(2, 156) = 55.75$, $p < 0.001$, with subsequent Tukey *post-hoc* analyses confirming the assumption that bilinguals and L3 learners did not differ significantly, $p = 0.48$, whereas both differed significantly from L2 learners, $ps < 0.001$. Parents' highest levels of education (on a 5-point scale: 1 = not completed high school to 5 = graduate or professional degree) were collapsed across both parents and used to index socio-economic status (SES). There were no differences between groups, $F < 1$, $p > 0.80$. Background measures are reported in **Table 1**.

As mentioned above, bilinguals and L3 learners were expected to perform similarly on the executive function tasks, and were thus subsumed under the label *multilinguals*. This resulted in subsequent comparisons of two instead of three groups.

TABLE 1 | Mean scores (and standard deviations) for background measures by language group.

	L2 learners (n = 77)	Bilinguals (n = 34)	L3 learners (n = 48)
Sex (female/male)	43/34	12/22	30/18
Age (years)	9.78 (2.25)	9.14 (2.29)	9.98 (2.28)
Length of Immersion (English)	1.83 (1.61)	3.04 (1.63)	1.79 (1.46)
Home language environment ^a	0.68 (0.51)	2.07 (1.08)	1.87 (0.89)
Parents' education ^b	3.46 (0.72)	3.52 (0.65)	3.53 (0.75)
PC usage (hours per day)	0.68 (0.54)	0.76 (0.54)	0.69 (0.50)
TROG-D standard score ^c	105.12 (12.31)	104.44 (17.09)	103.77 (16.78)
TROG-2 standard score	73.42 (14.69)	99.88 (11.71)	81.15 (16.20)
Raven's standard score	98.42 (11.42)	103.82 (12.50)	100.00 (10.21)

^aHome language environment was quantified using a 5-point scale where 0, "All German"; 2, "Half German; half other language(s)"; 4, "Only other language."

^bEducation was quantified using a 5-point scale where 1, no high school diploma; 2, high school graduate; 3, some college or college diploma; 4, bachelor's degree; 5, graduate degree; score collapsed across parents.

^cTROG-D German standard score calculated by transforming TROG-D T-score (range 20 to 80; $M = 50$, $SD = 10$) to TROG-2 English standard score (range of 55–145; $M = 100$, $SD = 15$). Formula for converting T-scores into standard scores: $b = ((a - \text{mean}) / a \text{ SD}) \times b \text{ SD} + b \text{ mean}$, where $a = \text{T-score}$ and $b = \text{standard score}$.

MATERIALS AND PROCEDURE

The background measures and experimental tasks were completed by the children in one session of approximately 45 min. First, one of the language proficiency tasks was administered, followed by one of the executive function tasks, the Raven's test, the other executive function task, and finally the other language proficiency task. The order of the language proficiency tasks and executive function tasks was counterbalanced. The children were informed before the experiment session began that they could choose to discontinue being tested at any time during the testing session. Each child was tested individually in a quiet room of their schools by a trained experimenter. Once the session was completed, the children received a small gift for the participation.

Background Measures

Test for Reception of Grammar

The Test for Reception of Grammar measures the receptive language proficiency of children. It was originally created by Bishop for English (TROG-2; Bishop, 2003), and is also available in revised and amended version for German (TROG-D; Fox, 2006). While the materials used in both test versions have some overlap, half the items are different. To counteract any spillover effects, the two tests were administered in a counterbalanced manner at the beginning and at the end of the test battery.

Raven's Colored Progressive Matrices

The Raven's CPM test (Raven et al., 1998) is a measure of non-verbal visuospatial reasoning. Participants are shown two arrays of colored pictures: one picture forms a pattern and a second one depicts potential components of the pattern. Participants must indicate the picture in the second array that best matches and

¹There is no ethics committee available for experimental studies conducted with human participants at the Faculty of Philology, University of Münster. The present study is in accordance with local legislation and the institutional requirements and follows the Code of Ethics "Rules of Good Scientific Practice" of the University of Münster (2002) and The European Code of Conduct for Research Integrity (European Federation of Academies of Sciences and Humanities, 2017).

fits into the picture in the first array. Results are calculated as standard scores corrected for age.

Executive Functions Tasks

The executive function tasks were the Simon task (Simon and Rudell, 1967) and the Flanker task (Eriksen and Eriksen, 1974).

Simon Task

In the Simon task, the children see single colored squares on the computer screen and need to press a left or right button to indicate the color of the square. The position of each square on the screen renders a condition either congruent (e.g., a red-color square on the *left* calls for a left button press) or incongruent (e.g., a red-color square on the *right* calls for a left button press). Incongruent trials induce response conflict through a spatial stimulus-response mismatch, the resolution of which requires participants to draw on inhibitory processes for conflict resolution. In contrast, congruent trials with a spatial stimulus-response match induce no conflict. Each trial was initiated with a fixation cross at screen center 350 ms prior to stimulus onset, followed by a blank screen for 150 ms, after which the stimulus was displayed. Each stimulus remained on screen until a participant response or for a maximum of 3,000 ms. Before each next trial, an inter-trial interval of 850 ms ensued. All trials were counterbalanced with left/right responses. The experiment was presented in four blocks. First, there was a block of 12 practice trials to make participants familiar with the experiment. After this, there were three mixed blocks of 42 trials (14 central, congruent, and incongruent trials each), presented in a randomly generated order by the E-prime program.

Flanker Task

In the Flanker task, the children need to indicate the direction of a target arrow (pointing left or right) in the middle of an array of five arrows, using two buttons on a serial response box. Depending on which variant of the task is used, there are up to four types of trials. Baseline trials display a single arrow in the middle of the screen, while in neutral trials, two diamonds each flank the central arrow. These trial types were not used in the present study since they are sometimes reported in research but rarely analyzed (similarly to the central condition in the Simon task). Congruent trials show the flanking arrows pointing in the same direction as the target arrow, while incongruent trials have target and flanking arrows pointing in opposite directions. Each trial was initiated with a fixation cross at screen center 350 ms prior to stimulus onset, followed by a 150 ms blank, and then immediately by a stimulus. Each stimulus remained on screen until a participant response or for a maximum of 3,000 ms. The experiment was presented in five blocks. First, there was a block of 12 random congruent and incongruent practice trials to familiarize participants with the experiment. Then, there were four mixed blocks of 32 trials (16 congruent and 16 incongruent) presented in a randomly generated order by E-prime. Prior to each next trial, an inter-trial interval of 850 ms ensued. Only RTs of correct responses were included in the analysis.

By subtracting the performance in the congruent condition from that of the incongruent condition, a difference score

indexing inhibitory control is calculated in both the Simon and the Flanker task. The magnitude of each difference score indicates the distraction by the induced conflict experienced by individuals. Larger difference scores indicate less efficient conflict resolution and interference control.

RESULTS

Results from the demographic background, German and English receptive grammar, and non-verbal intelligence measures are presented in **Table 1**.

T-tests comparing the two groups' scores for German and English receptive grammar and non-verbal intelligence showed no difference in either German receptive grammar, $p = 0.65$, or in non-verbal intelligence, $p = 0.11$, while the children did differ in English receptive grammar, $p < 0.001$. One-way ANOVAs comparing the three original groups confirmed these results: German receptive grammar, $p > 0.50$, non-verbal intelligence, $p > 0.10$, English receptive grammar, $p < 0.001$ (Tukey *post-hoc* comparisons, all $ps < 0.01$), with the bilinguals showing the highest scores, followed by the L3 learners, and the lowest scores by L2 learners.

Mean response times (RT) and mean accuracy rates were calculated for each condition of the two executive function tasks. Central trials in the Simon task were part of the experimental set-up; however, they are conventionally not compared in subsequent analyses and are thus not reported here.

Data Trimming Procedure

Incorrect responses (Simon: 3.9% for the congruent condition, 9.6% for the incongruent condition; Flanker: 1.8% for the congruent condition, 5.1% for the incongruent condition) were excluded from the RT analysis, as were outliers with RTs shorter than 200 ms (Simon: 0.6% for the congruent condition, 0.6% for the incongruent condition; Flanker: 0.7% for the congruent condition, 1.1% for the incongruent condition). Contrary to Poarch and Bialystok (2015), RTs above 2,000 ms were not considered outliers (see Zhou and Krott, 2016, for rationale; see also De Cat et al., 2018). RT and accuracy data for both tasks are presented in **Table 2**.

Simon Task Results

RTs and accuracies on the two critical trial types in the Simon task, the congruent and incongruent trials, were analyzed using repeated measures mixed ANOVAs with trial type (congruent and incongruent) as within-group variable and language group (L2 learners, multilinguals) as between-group variable, and given the substantial age range, age was entered as a covariate. The RT analysis yielded a significant main effect of trial type, $F_{(1, 156)} = 68.24$, $\eta^2 = 0.28$, $p < 0.001$, no significant effect of language group, $F_{(1, 156)} < 1.5$, $\eta^2 < 0.01$, $p = 0.23$, and a significant effect of age, $F_{(1, 156)} = 143.52$, $\eta^2 = 0.48$, $p < 0.001$. Furthermore, there was no significant interaction between trial type and language group, $F_{(1, 156)} < 1.1$, $\eta^2 < 0.01$, $p = 0.34$, and a significant interaction between trial type and age, $F_{(1, 156)} = 15.29$, $\eta^2 = 0.06$, $p < 0.001$, with the children in the middle age range showing less performance

TABLE 2 | Mean RT and accuracy scores (and standard deviations) in Simon and Flanker task by language group.

Condition/Task	Language Group			
	L2 learners	Multilinguals	L2 learners	Multilinguals
	Simon		Flanker	
RT				
(a) Congruent	592 (122)	577 (128)	713 (246)	670 (197)
(b) Incongruent	649 (129)	641 (142)	798 (269)	725 (198)
(c) Conflict (b–a)	57	64	85	55
ACCURACY				
(a) Congruent	0.965 (0.038)	0.959 (0.050)	0.981 (0.029)	0.980 (0.034)
(b) Incongruent	0.906 (0.075)	0.902 (0.071)	0.942 (0.076)	0.952 (0.055)
(c) Conflict (b–a)	0.059	0.057	0.039	0.028

overlap across groups than the younger and older children, who showed a large performance overlap in both conditions. The non-significant interaction between trial type and language group was confirmed by the similar conflict magnitudes (incongruent condition RTs—congruent condition RTs) for L2 learners (57 ms) and multilinguals (64 ms).

The accuracy analysis similarly yielded a significant main effect of trial type, $F_{(1, 156)} = 11.40$, $\eta^2 = 0.07$, $p < 0.001$, none of language group, $F_{(1, 156)} < 1$, $\eta^2 < 0.01$, $p > 0.50$, and a significant main effect of age, $F_{(1, 156)} = 10.37$, $\eta^2 = 0.06$, $p = 0.002$. Furthermore, there were no significant interactions, $F_{(1, 156)} < 1.3$, $\eta^2 < 0.01$, $ps > 0.26$. The results indicate that the groups performed similarly overall (no domain-general executive function difference) and displayed similar effect magnitudes (no domain-specific inhibitory difference), and as such did not differ in resolving conflict in the Simon task.

Flanker Task Results

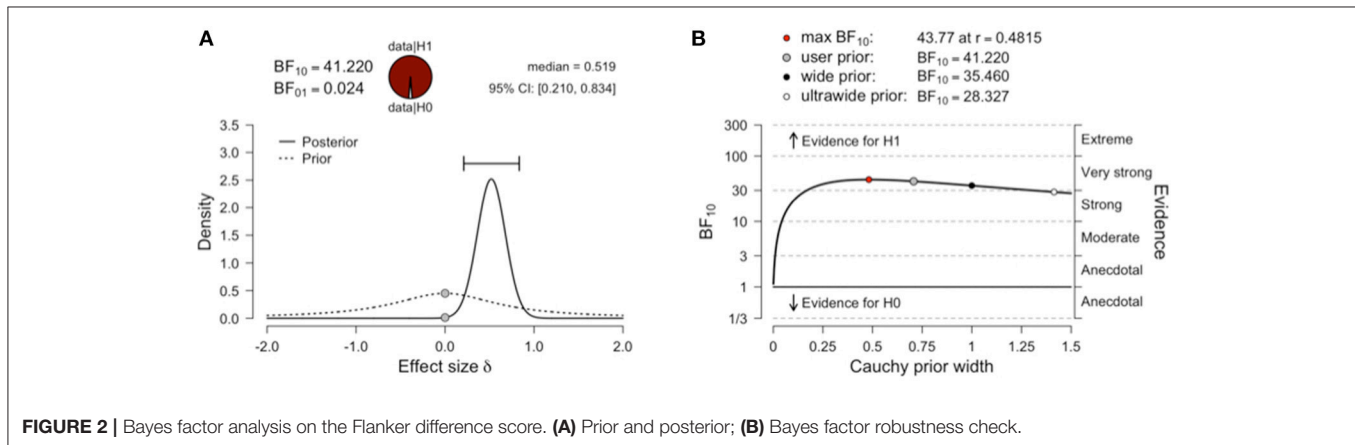
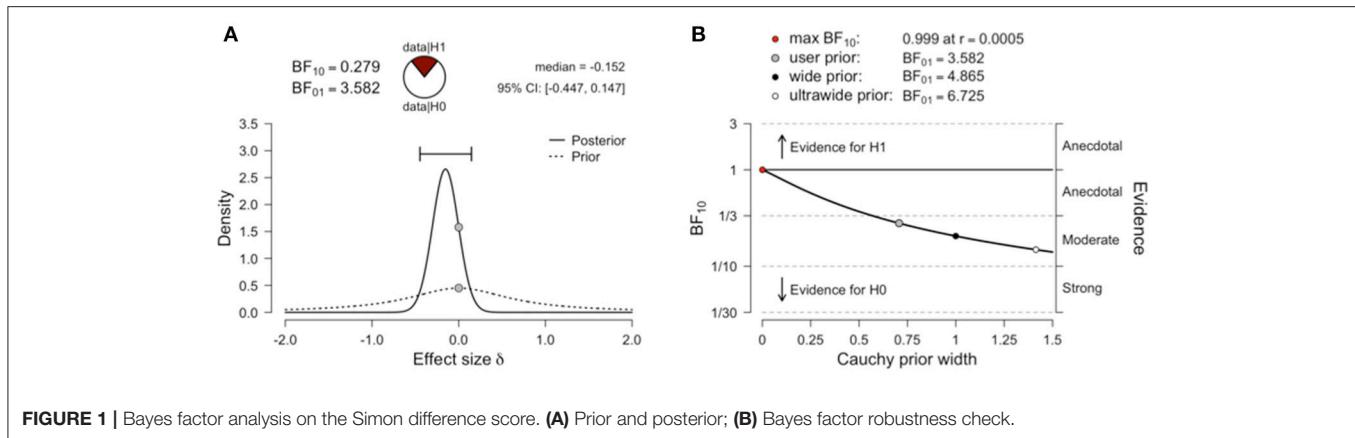
Subsequently, performance on the congruent and incongruent trials of the Flanker task was analyzed in the same way as for the Simon task. The RT analysis yielded a significant main effect of trial type, $F_{(1, 156)} = 24.96$, $\eta^2 = 0.13$, $p < 0.001$, a main effect of language group, $F_{(1, 156)} = 7.77$, $\eta^2 = 0.02$, $p = 0.006$, and a main effect of age, $F_{(1, 156)} = 187.34$, $\eta^2 = 0.53$, $p < 0.001$. Furthermore, there was a significant interaction between trial type and language group, $F_{(1, 156)} = 12.59$, $\eta^2 = 0.06$, $p < 0.001$, but none between trial type and age, $F_{(1, 156)} < 1.7$, $\eta^2 < 0.01$, $p = 0.20$. The interaction between trial type and language group was further investigated through a separate one-way ANOVA on the conflict magnitudes, $F_{(1, 157)} = 12.19$, $\eta^2 = 0.07$, $p < 0.001$, showing a larger conflict for L2 learners (85 ms) than for multilinguals (55 ms). This result was further confirmed by comparisons of performance on congruent and incongruent conditions separately. While the groups did not differ significantly in the congruent condition, $F_{(1, 157)} = 1.44$, $\eta^2 < 0.01$, $p = 0.23$, they did so marginally in the incongruent condition, $F_{(1, 157)} = 3.76$, $\eta^2 = 0.02$, $p = 0.054$, which is assumed to have driven the significant main effect of language group.

The accuracies were at ceiling performance and the analysis yielded no main effect of trial type, $F_{(1, 156)} = 1.6$, $\eta^2 = 0.01$, $p = 0.21$, none of language group, $F_{(1, 156)} < 1$, $\eta^2 < 0.01$, $p = 0.56$, but a significant main effect of age, $F_{(1, 156)} = 11.82$, $\eta^2 = 0.07$, $p < 0.001$. Furthermore, there were no significant interactions, $F_{(1, 156)} < 1.9$, $\eta^2 < 0.02$, $p > 0.18$. As such, the results show no overall faster performance for multilinguals compared to L2 learners. However, they do indicate that multilinguals exhibit enhanced conflict resolution over L2 learners in the Flanker task and thus better domain-specific inhibitory control.

To tease apart whether the collapsed group of multilinguals also differed from the L2 learners when separated into the original two groups of bilinguals and L3 learners, a repeated measures mixed ANOVAs with trial type (congruent and incongruent) as within-group variable, language group (L2 learners, bilinguals, L3 learners) as between-group variable, and age as a covariate was conducted on the Flanker RTs only. The RT analysis yielded significant main effects of trial type, $F_{(1, 156)} = 21.93$, $\eta^2 = 0.12$, $p < 0.001$, of language group, $F_{(1, 156)} = 4.90$, $\eta^2 = 0.03$, $p = 0.009$, and of age, $F_{(1, 156)} = 190.32$, $\eta^2 = 0.54$, $p < 0.001$. Furthermore, there was a significant interaction between trial type and language group, $F_{(1, 156)} = 6.26$, $\eta^2 = 0.07$, $p = 0.002$, but none between trial type and age, $F_{(1, 156)} < 1.7$, $\eta^2 < 0.01$, $p = 0.21$. The interaction between trial type and language group was further investigated through a separate one-way ANOVA on the conflict magnitudes, yielding significant differences between groups, $F_{(1, 157)} = 6.09$, $\eta^2 = 0.07$, $p = 0.003$. Tukey *post-hoc* comparisons showed that bilinguals (57 ms) and L3 learners (54 ms) both resolved conflict significantly faster than L2 learners (85 ms), $p = 0.034$ and $p = 0.006$, respectively. Critically, bilinguals and L3 learners did not differ significantly, $p = 0.96$. The results confirm the two-group comparison above and indicate that bilinguals and L3 learners showed smaller effect magnitudes and were thus better at resolving conflict than L2 learners in the Flanker task.

Bayes Analyses

Finally, in an attempt to confirm the results obtained from the repeated measures ANOVA and to better adjudicate between the null hypothesis (H_0), which means that the groups of children did not differ significantly in their performance, and the alternative hypothesis (H_1), namely that the groups did indeed differ, Bayes factor analyses were performed (Wagenmakers et al., 2016) using JASP (JASP Team, 2018). Bayes factors indicate the weighted evidence either for or against specific effects of interest, which is displayed using BF_{01} for evidence in favor of the null hypothesis (H_0) vs. BF_{10} for evidence in favor of the alternative hypothesis (H_1) (for more detailed information on Bayesian inference, see Wagenmakers et al., 2018). For example, Bayes factors below 1 provide little evidence for the effects of interest, whereas Bayes factors above 30 provide very strong evidence for such effects (see Figures 1, 2). For the difference score obtained in the Simon task, the Bayes factor with a BF_{10} value of 0.28 indicated moderate evidence for the null hypothesis (see Figure 1), which means the difference scores across groups were similar. In contrast, for the Flanker task difference scores,



there was strong to very strong evidence for the alternative hypothesis, indicating that the language groups differed, with a BF_{10} value of 41.22 (see **Figure 2**). The latter Bayes factor indicates that the data are 41.22 times more likely under H_1 than under H_0 .

Correlational Analyses Simon Task and Flanker Task

The present study employed two executive function tasks that are customarily used to tap individuals' inhibitory control. Hence, one could hypothesize that performance on one task should correlate with that on the other (see Poarch and Van Hell, in press, for a more detailed rationale). To test this hypothesis, RT performance from both tasks on the congruent condition, the incongruent condition, and the resulting difference score (i.e., the conflict magnitudes also referred to as the Simon effect and the Flanker effect) were entered into a correlational analysis (see **Table 3**).

Within-Task Correlations

The Simon task congruent and incongruent conditions correlated significantly, $r = 0.95$, $p < 0.001$, as did the incongruent condition and the Simon effect, $r = 0.40$, $p < 0.001$, while the congruent condition and the Simon effect did not, $r = 0.10$,

$p = 0.20$. For the Flanker task, the congruent and incongruent conditions correlated significantly, $r = 0.97$, $p < 0.001$, as did the incongruent condition and the Flanker effect, $r = 0.37$, $p < 0.001$, and the congruent condition and the Flanker effect correlated marginally, $r = 0.14$, $p = 0.07$.

Cross-Task Correlations

The Simon and Flanker congruent conditions, $r = 0.71$, $p < 0.001$, and the incongruent conditions, $r = 0.70$, $p < 0.001$, correlated significantly. Critically, however, the Simon and Flanker effects did not correlate, $r = 0.07$, $p = 0.37$.

Finally, the Simon and Flanker effects were entered into a correlational analysis with the Home Language Environment score as an index for how multilingual the language environment of the children was outside of their educational context. While the Home Language Environment score and the Simon effect showed no significant correlation, $r = -0.06$, $p = 0.47$, the Home Language Environment score and the Flanker effect correlated significantly, $r = -0.16$, $p = 0.05$. Evidently, the more multilingual the children's environment was, the better they were at resolving conflict in the Flanker task, but not in the Simon task. Hence, the tasks may be tapping different components of executive function and inhibitory control (see Discussion for a more detailed interpretation). As Keye et al. (2009) have also

TABLE 3 | Pearson correlations between performance in task conditions and measures of inhibitory control.

Conditions and congruency effects	Simon			Flanker			
	1	2	3	4	5	6	7
1. Simon congruent	—	0.954***	0.102	0.713***	0.688***	0.077	−0.051
2. Simon incongruent		—	0.396***	0.726***	0.703***	0.092	−0.065
3. Simon effect			—	0.225**	0.228**	0.071	−0.061
4. Flanker congruent				—	0.972***	0.142°	−0.107
5. Flanker incongruent					—	0.370**	−0.137°
6. Flanker effect						—	−0.156*
7. Home lang. environment							—

Cross-correlations significant at $p < 0.05^*$; at $p < 0.01^{**}$; at $p < 0.001^{***}$; and marginally significant at $p < 0.10^\circ$; all in bold.

pointed out, the conflicts induced in both tasks are likely caused by more than one source of variance, which may make it is less likely to find a correlation of the conflicts across tasks.

DISCUSSION

The rationale for the present study was to explore whether the sustained cognitive control exerted on a daily basis by multilingual children in order to control their languages affects the development of their non-verbal executive function differently than that of monolingual children, and, in doing so, to replicate an earlier study by Poarch and Van Hell (2012a) with a very similar population living in the same language environment but extended to a wider age range. While the original study had focused on children aged 5–8, the present study tested 5- to 13-year old children. For this purpose, two executive function tasks were administered to the children to investigate whether their performance would differ across groups in their task monitoring (i.e., overall speed) and in their resolution of conflict (i.e., the difference score).

The Simon task data yielded no difference between groups, with multilinguals and monolinguals performing similarly both in overall speed and accuracy and in the obtained difference score. In contrast, the Flanker task showed that multilinguals and monolinguals differed significantly in their efficacy to resolve conflict, notably, and critically driven by differing performance in the incongruent condition, in that multilinguals displayed significantly smaller difference scores than monolinguals. While the Simon tasks results are not in line with those of the previous study, the Flanker results corroborate the earlier findings.

In light of these mixed findings, two issues will be highlighted and discussed in the following: (1) the nature of the population tested and matching of groups, and (2) the type of tasks used to tap executive function.

First, previous mixed findings have, amongst other explanations, been attributed to various factors inherent in

comparing groups experimentally, such as whether or not multilingual and monolingual children had been adequately matched on first language proficiency and socio-economic status (Paap et al., 2015). However, as Poarch and Van Hell (2017) have pointed out, the matching of children groups has not been overtly systematically different—in both research documenting differences between groups and that reporting null-results—to serve as a sufficient explanation for the mixed results (see also Baum and Titone, 2014; Bialystok, 2017). In the present study, the groups of children all attended private immersion schools, were meticulously matched on age, socio-economic status, fluid intelligence, PC usage, and L1 proficiency. The groups did differ, however, on the background variables L2 proficiency and home language environment, which are exactly those that could be assumed to differentiate multilinguals from monolinguals. Additional information on multilingual language usage patterns following the Adaptive Control Hypothesis by Green and Abutalebi (2013) may, in the future, offer a more fine-grained assessment of multilingual individuals and offer insight into within-group differences based on distinct contexts of multilingual interaction. According to Green and Abutalebi, single language, dual language, and dense code-switching contexts in a multilingual's life require differing degrees of cognitive control and thus also pose varying demands on the executive function system (see also Yang et al., 2016). However, for researchers to utilize such information, multilingual participants would need to be able to validly indicate which of these contexts pervade their lives. Moreover, a caveat to most research conducted in the field so far is that there are other lifestyle variables that have an effect on the development of executive function and may thus also influence performance on executive function tasks. Musical expertise (Peretz and Zatorre, 2005; Zuk et al., 2014; Schroeder et al., 2016) has been shown to be one of these variables, as has physical exercise (Best, 2010), dietary intake (Kim and Wang, 2017), circadian rhythm (Hahn et al., 2012), and sleep quality (Kuula et al., 2015). Future research could take all these additional variables into account and possibly an array of others (see Bak and Robertson, 2017), although the measurement of all of these may prove rather cumbersome in the scope of experimental research conducted in the field. What is striking, however, is that the effects on executive function of these diverse lifestyle variables seem to be less controversial than those of using multiple languages in daily life (cf. Bak, 2016).

Second, two prominent tasks in multilingualism research, the Simon and the Flanker task, have in the past been interchangeably and ubiquitously used to investigate executive function, and more specifically, conflict resolution, inhibitory control, and task monitoring. However, on closer inspection, the two tasks display differences in task demands that may inadvertently draw on both overlapping and non-overlapping subcomponents of executive function during task performance. According to Botvinick et al. (2001), both tasks can be described using the conflict monitoring and control theory, in which a conflict detector in the brain's ACC is triggered by a conflict signal. In the prefrontal cortex, control processes are then engaged to focus on relevant stimulus features in the task, which is stimulus location in the Simon task and feature dimension in

the Flanker task. Subsequently, stimulus-response compatibility is determined, upon which initiation of the correct response follows. While in both tasks, performance depends on whether the condition is compatible or incompatible, performance is modulated differently: in the Simon task through stimulus-response compatibility and bi-dimensional perceptual and motor conflicts, whereas in the Flanker task through stimulus-stimulus compatibility and uni-dimensional perceptual conflict (e.g., Keye et al., 2009; Ambrosi et al., 2016; see also Posner, 1980; Abrahamse and Van der Lubbe, 2008; Snyder et al., 2015). As such, these differences in how conflict is elicited may engage partially differing cognitive processes and induce varying cognitive loads during task performance, possibly also modulated depending on the age of the participant. Given the disparate developmental trajectories of the various executive function subcomponents in children (Anderson, 2002), one may adduce that children at varying ages may be differentially cognitively taxed during task performance of the Simon and the Flanker. The findings by Poarch and Van Hell (2012a), who found differences between groups in both the Simon and the Flanker for 5- to 8-year-old children, and the results of the present study with 5- to 13-year-old children, who differed only in the Flanker task, speaks to the effect of age on task performance and its development.

The correlational analyses conducted in the present study are thus informative as they indicate significant correlations across task conditions, corroborating the results reported by Ross and Melinger (2017), who found the performance of their groups of children to correlate across tasks (see also Poarch et al., 2018, for adults). However, in the present study, similarly to Poarch and Van Hell (in press), the difference score did not correlate significantly across tasks (see Kousaie and Phillips, 2012; Paap and Greenberg, 2013, for adults). The mixed findings from these correlational analyses are thus inconclusive as to whether these two measures of executive function tap conflict resolution, inhibitory control, and task monitoring similarly, which could be expected according to Miyake and Friedman (2012) if the same underlying cognitive processes were engaged during task performance. The present study's correlational results indicate the engagement of similar subcomponents of task monitoring across tasks (i.e., correlation of the two conditions) but separable subcomponents of inhibitory control (i.e., non-correlation of difference score). Furthermore, while the home language environment as an index of degree of multilingualism correlated significantly with the Flanker task difference score, this was not the case for that of the Simon task. The partially

diverging cognitive demands posed by the two tasks may thus be critical in whether or not differential performance emerges in multilinguals and monolinguals and whether their performance correlates (Macnamara and Conway, 2014; Ambrosi et al., 2016; Qu et al., 2016; for a more detailed discussion, see Poarch and Van Hell, 2017, and Poarch and Van Hell, in press). This may, all the more so, be the case for individuals such as children in whom executive function development is still ongoing (Anderson, 2002).

CONCLUSION

The present study aimed at replicating earlier research in a population from the same language environment using the same experimental design. The results offer partially corroborating evidence of systematic differences in executive function between multilingual and monolingual children aged 5–13. Given the debate on the findings in the executive function and multilingualism literature, culminating in titles such as “There is no coherent evidence for a bilingual advantage in executive processing” (Paap and Greenberg, 2013), the present findings partially replicate earlier findings and tentatively support the view that multilingualism indeed has an effect on executive function task performance, albeit depending on which tasks is used. The differing performance of the groups across tasks was hypothesized to be driven by factors such as differences in induced cognitive load and task complexity. Furthermore, differences in individuals' language backgrounds, language usage patterns, and other lifestyle variables may have a crucial impact on the course of executive function development in children (Baum and Titone, 2014; Van Hell and Poarch, 2014). Future research may want to draw on more sensitive measures of executive function and aim at testing children longitudinally to better trace the development of executive function over time.

AUTHOR CONTRIBUTIONS

GP conception, design, data collection, statistics, and writing.

ACKNOWLEDGMENTS

The author thanks Alexandra Kemmerer and Sharmin Reza for their assistance with data collection. The publication of this manuscript was supported by the Open Access Publication Fund of the University of Muenster.

REFERENCES

- Abrahamse, E. L., and Van der Lubbe, R. H. J. (2008). Endogenous orienting modulates the Simon effect: critical factors in experimental design. *Psychol. Res.* 72, 261–272. doi: 10.1007/s00426-007-0110-x
- Ambrosi, S., Lemaire, P., and Blaye, A. (2016). Do young children modulate their cognitive control? sequential congruency effects across three conflict tasks in 5-to-6 year-olds. *Exp. Psychol.* 63, 117–126. doi: 10.1027/1618-3169/a000320
- Anderson, J. A. E., Chung-Fat-Yim, A., Bellana, B., Luk, G., and Bialystok, E. (2018a). Language and cognitive control networks in bilinguals and monolinguals. *Neuropsychologia* 117, 352–363. doi: 10.1016/j.neuropsychologia.2018.06.023
- Anderson, J. A. E., Mak, L., Keyvani Chahi, A., and Bialystok, E. (2018b). The language and social background questionnaire: assessing degree of bilingualism in a diverse population. *Behav. Res. Methods* 50, 250–263. doi: 10.3758/s13428-017-0867-9
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychol.* 8, 71–82. doi: 10.1076/chin.8.2.71.8724
- Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L., and Carreiras, M. (2014). Is there a bilingual advantage in the ANT

- task? Evidence from children. *Front. Psychol.* 5:398. doi: 10.3389/fpsyg.2014.00398
- Antoniou, M. (2019). The advantages of bilingualism debate. *Ann. Rev. Linguist.* 5, 1–21. doi: 10.1146/annurev-linguistics-011718-011820
- Bak, T. (2016). Cooking pasta in La Paz. *Linguist. Approach. Biling.* 6, 699–717. doi: 10.1075/lab.16002.bak
- Bak, T. H., and Robertson, I. (2017). Biology enters the scene. a new perspective on bilingualism, cognition, and dementia. *Neurobiol. Aging* 50, 3–4. doi: 10.1016/j.neurobiolaging.2016.10.020
- Baum, S., and Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Appl. Psycholinguist.* 35, 857–894. doi: 10.1017/S0142716414000174
- Best, J. R. (2010). Effects of physical activity on children's executive function: contributions of experimental research on aerobic exercise. *Dev. Rev.* 30, 331–551. doi: 10.1016/j.dr.2010.08.001
- Bialystok, E. (2017). The bilingual adaptation: how minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099
- Bialystok, E., Craik, F. I. M., and Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends Cogn. Sci.* 16, 240–250. doi: 10.1016/j.tics.2012.03.001
- Bishop, D. V. M. (2003). *The Test for Reception of Grammar, Version 2 (TROG-2)*. London: Harcourt Assessment.
- Blom, E., Boerma, T., Bosma, E., Cornips, L., and Everaert, E. (2017). Cognitive advantages of bilingual children in different sociolinguistic contexts. *Front. Psychol.* 8:552. doi: 10.3389/fpsyg.2017.00552
- Blom, E., Küntay, A. C., Messer, M., Verhagen, J., and Leseman, P. (2014). The benefits of being bilingual: working memory in bilingual Turkish- Dutch children. *J. Exp. Child Psychol.* 128, 105–119. doi: 10.1016/j.jecp.2014.06.007
- Botvinick, M. M., Braver, T. S., Barch, D., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 625–652. doi: 10.1037/0033-295X.108.3.624
- Branzi, F. M., Calabria, M., Boscarino, M. L., and Costa, A. (2016). On the overlap between bilingual language control and domain-general executive control. *Acta Psychol.* 166, 21–30. doi: 10.1016/j.actpsy.2016.03.001
- Calabria, M., Branzi, F. M., Marne, P., Hernandez, M., and Costa, A. (2015). Age-related effects over bilingual language control and executive control. *Biling. Lang. Cogn.* 18, 65–78. doi: 10.1017/S1366728913000138
- Calabria, M., Costa, A., Green, D. W., and Abutalebi, J. (2018). Neural basis of bilingual language control. *Ann. N. Y. Acad. Sci.* 1426, 221–235. doi: 10.1111/nyas.13879
- Carlson, S. M., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298. doi: 10.1111/j.1467-7687.2008.00675.x
- Crivello, C., Kuzyk, O., Rodrigues, M., Friend, M., Zesiger, P., and Poulin-Dubois, D. (2016). The effects of bilingual growth on toddlers' executive function. *J. Exp. Child Psychol.* 141, 121–132. doi: 10.1016/j.jecp.2015.08.004
- De Cat, C., Gusnanto, A., and Serratrice, L. (2018). Identifying a threshold for the executive function advantage in bilingual children. *Stud. Second Lang. Acquis.* 40, 119–51. doi: 10.1017/S0272263116000486
- Declerck, M., Grainger, J., Koch, I., and Philipp, A. M. (2017). Is language control just a form of executive control? Evidence for overlapping processes in language switching and task switching. *J. Mem. Lang.* 95, 138–145. doi: 10.1016/j.jml.2017.03.005
- Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750
- Duñabeitia, A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L., et al. (2013). The inhibitory advantage in bilingual children revisited: myth or reality? *Exp. Psychol.* 61, 234–251. doi: 10.1027/1618-3169/a000243
- Engel de Abreu, P. M., Cruz-Santos, A., Tourinho, C. J., Martin, R., and Bialystok, E. (2012). Bilingualism enriches the poor: enhanced cognitive control in low-income minority children. *Psychol. Sci.* 23, 1364–1371. doi: 10.1177/0956797612443836
- Engle, R. W. (2002). Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11, 19–23. doi: 10.1111/1467-8721.00160
- Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267
- European Federation of Academies of Sciences and Humanities (2017). *The European Code of Conduct for Research Integrity*. Berlin: ALLEA.
- Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., and Posner, M. I. (2003). Cognitive and brain consequences of conflict. *Neuroimage* 18, 42–57. doi: 10.1006/nimg.2002.1319
- Fan, J., McCandliss, B. D., Sommer, T., Raz, A., and Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *J. Cogn. Neurosci.* 14, 340–347. doi: 10.1162/089892902317361886
- Filippi, R., Morris, J., Richardson, F. M., Bright, P., Thomas, M. S. C., Karmiloff-Smith, A., et al. (2015). Bilingual children show an advantage in controlling verbal interference during spoken language comprehension. *Biling. Lang. Cogn.* 18, 490–501. doi: 10.1017/S1366728914000686
- Fox, A. V. (2006). *TROG-D: Test zur Überprüfung des Grammatikverständnisses*. [Test for reception of grammar - German version]. Idstein: Schulz-Kirchner.
- Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *J. Exp. Psychol. General* 133, 101–135. doi: 10.1037/0096-3445.133.1.101
- Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Viñas Guasch, N., Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschooler through older adults on card sorting, Simon, and metalinguistic tasks. *Front. Psychol.* 5:11. doi: 10.3389/fpsyg.2014.00011
- Green, D. W., and Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377
- Hahn, C., Cowell, J. M., Wiprzycka, U. J., Goldstein, D., Ralph, M., Hasher, L., et al. (2012). Circadian rhythms in executive function during the transition to adolescence: the effect of synchrony between chronotype and time of day. *Dev. Sci.* 15, 408–416. doi: 10.1111/j.1467-7687.2012.01137.x
- Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychol.* 136, 189–202. doi: 10.1016/j.actpsy.2010.04.011
- JASP Team (2018). *JASP (Version 0.9.0.1) [Computer Software]*. Amsterdam.
- Kapa, L. L., and Colombo, J. (2013). Attentional control in early and later bilingual children. *Cogn. Dev.* 28, 233–246. doi: 10.1016/j.cogdev.2013.01.011
- Keye, D., Wilhelm, O., Oberauer, K., and van Ravenzwaaij, D. (2009). Individual differences in conflict-monitoring: testing means and covariance hypothesis about the Simon and the Eriksen Flanker task. *Psychol. Res.* 73, 762–776. doi: 10.1007/s00426-008-0188-9
- Kim, J. Y., and Wang, S. W. (2017). Relationships between dietary intake and cognitive function in healthy Korean children and adolescents. *J. Lifestyle Med.* 7, 10–17. doi: 10.15280/jlm.2017.7.1.10
- Kousaie, S., and Phillips, N. A. (2012). Conflict monitoring and resolution: are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain Res.* 1446, 71–90. doi: 10.1016/j.brainres.2012.01.052
- Kroll, J. F., and Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *J. Cogn. Psychol.* 25, 497–514. doi: 10.1080/20445911.2013.799170
- Kuula, L., Pesonen, A. K., Martikainen, S., Kajantie, E., Lahti, J., Strandberg, T., and Rääkkönen, K. (2015). Poor sleep and neurocognitive function in early adolescence. *Sleep Med.* 16, 1207–1212. doi: 10.1016/j.sleep.2015.06.017
- Ladas, A. I., Carroll, D. J., and Vivas, A. B. (2015). Attentional processes in low-socioeconomic status bilingual children: are they modulated by the amount of bilingual experience? *Child Dev.* 86, 557–578. doi: 10.1111/cdev.12332
- Macnamara, B. N., and Conway, A. R. A. (2014). Novel evidence in support of the bilingual advantage: influences of task demands and experience on cognitive control and working memory. *Psychon. Bull. Rev.* 21, 520–525. doi: 10.3758/s13423-013-0524-y
- Martin-Rhee, M. M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Biling. Lang. Cogn.* 11, 81–93. doi: 10.1017/S1366728907003227
- Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458
- Morales, J., Calvo, A., and Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *J. Exp. Child Psychol.* 114, 187–202. doi: 10.1016/j.jecp.2012.09.002
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science* 349:aac4716. doi: 10.1126/science.aac4716

- Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Paap, K. R., Johnson, H. A., and Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 73, 265–278. doi: 10.1016/j.cortex.2015.04.014
- Peretz, I., and Zatorre, R. J. (2005). Brain organization for music processing. *Annu. Rev. Psychol.* 56, 89–114. doi: 10.1146/annurev.psych.56.091103.070225
- Poarch, G. J. (2018). “How the use and control of multiple languages can affect attentional processes,” in *The Interactive Mind: Language, Vision and Attention*, eds N. Mani, R. M. Mishra, and F. Huettig (Basingstoke: Macmillan), 53–60.
- Poarch, G. J., and Bialystok, E. (2015). Bilingualism as a model for multitasking. *Dev. Rev.* 35, 113–124. doi: 10.1016/j.dr.2014.12.003
- Poarch, G. J., and Van Hell, J. G. (2012a). Executive functions and inhibitory control in multilingual children: evidence from second language learners, bilinguals, and trilinguals. *J. Exp. Child Psychol.* 113, 535–551. doi: 10.1016/j.jecp.2012.06.013
- Poarch, G. J., and Van Hell, J. G. (2012b). Cross-language activation in children’s speech production: evidence from second language learners, bilinguals, and trilinguals. *J. Exp. Child Psychol.* 111, 419–438. doi: 10.1016/j.jecp.2011.09.008
- Poarch, G. J., and Van Hell, J. G. (2017). “Bilingualism and the development of executive function in children: the interplay of languages and cognition,” in *Executive Function: Development Across the Life Span (Frontiers of Developmental Science)*, eds S. Wiebe and J. Karch (New York, NY: Taylor and Francis), 172–187.
- Poarch, G. J., and Van Hell, J. G. (in press). “Does performance on executive function tasks correlate? Evidence from child trilinguals, bilinguals, and second language learners,” in *Bilingualism, Executive Function, and Beyond: Questions and Insights. Studies in Bilingualism* (John Benjamins), 57.
- Poarch, G. J., Vanhove, J., and Berthele, R. (2018). The effect of bidialectalism on executive function. *Int. J. Biling.* 1–17. doi: 10.1177/1367006918763132
- Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 23, 3–25. doi: 10.1080/00335558008248231
- Qu, L., Low, J. J. W., Zhang, T., Li, H., and Zelazo, P. D. (2016). Bilingual advantage in executive control when task demands are considered. *Biling. Lang. Cogn.* 19, 277–293. doi: 10.1017/S1366728914000376
- Raven, J. C., Raven, J., and Court, J. H. (1998). *Raven’s Colored Progressive Matrices*. San Antonio, TX: The Psychological Corporation.
- Ross, J., and Melinger, A. (2017). Bilingual advantage, bidialectal advantage or neither? Comparing performance across three tests of executive function in middle childhood. *Dev. Sci.* 20:e12405. doi: 10.1111/desc.12405
- Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in children. *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012
- Schroeder, S. R., Marian, V., Shook, A., and Bartolotti, J. (2016). Bilingualism and musicianship enhance cognitive control. *Neural Plast.* 4058620, 1–11. doi: 10.1155/2016/4058620
- Simon, J. R., and Rudell, A. P. (1967). Auditory S–R compatibility: the effect of an irrelevant cue on information processing. *J. Appl. Psychol.* 51, 300–304. doi: 10.1037/h0020586
- Snyder, H. R., Miyake, A., and Hankin, B. L. (2015). Advancing understanding of executive function impairments and psychopathology: bridging the gap between clinical and cognitive approaches. *Front. Psychol.* 6:328. doi: 10.3389/fpsyg.2015.00328
- Surraín, S., and Luk, G. (2017). Describing bilinguals: a systematic review of labels and descriptions used in the literature between 2005–2015. *Biling. Lang. Cogn.* 1–15. doi: 10.1017/S1366728917000682
- Thierry, G., and Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12530–12535. doi: 10.1073/pnas.0609927104
- Thomas-Sunesson, D., Hakuta, K., and Bialystok, E. (2018). Degree of bilingualism modifies executive control in Hispanic children in the US. *Int. J. Biling. Educ. Biling.* 21, 197–206. doi: 10.1080/13670050.2016.1148114
- University of Münster (2002). *Code of Ethics “Rules of Good Scientific Practice.”* Available online at: https://www.uni-muenster.de/imperia/md/content/wwu/senat/pdf/kodex_en_22_11_2017.pdf
- Valian, V. (2015). Bilingualism and cognition. *Biling. Lang. Cogn.* 18, 3–24. doi: 10.1017/S1366728914000522
- Van Hell, J. G., and Poarch, G. J. (2014). How much bilingual experience is needed to affect executive control? *Appl. Psycholinguist.* 35, 925–928. doi: 10.1017/S0142716414000265
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., and Morey, R. D. (2018). Bayesian inference for psychology. Part I: theoretical advantages and practical ramifications. *Psychon. Bull. Rev.* 25, 35–57. doi: 10.3758/s13423-017-1343-3
- Wagenmakers, E.-J., Morey, R. D., and Lee, M. D. (2016). Bayesian benefits for the pragmatic researcher. *Curr. Dir. Psychol. Sci.* 11, 19–23. doi: 10.1177/0963721416643289
- Wöstmann, N. M., Aichert, D. S., Costa, A., Rubia, K., Moller, H. J., and Ettinger, U. (2013). Reliability and plasticity of response inhibition and interference control. *Brain Cogn.* 81, 82–94. doi: 10.1016/j.bandc.2012.09.010
- Yang, H., Hartanto, A., and Yang, S. (2016). The complex nature of bilinguals’ language usage modulates task-switching outcomes. *Front. Psychol.* 7:560. doi: 10.3389/fpsyg.2016.00560
- Yang, S., and Yang, H. (2016). Bilingual effects on deployment of the attention system in linguistically and culturally homogeneous children and adults. *J. Exp. Child Psychol.* 146, 121–136. doi: 10.1016/j.jecp.2016.01.011
- Yang, S., Yang, H., and Lust, B. (2011). Early childhood bilingualism leads to advances in executive attention: dissociating culture and language. *Biling. Lang. Cogn.* 14, 412–422. doi: 10.1017/S1366728910000611
- Zhou, B., and Krott, A. (2016). Data trimming procedure can eliminate bilingual cognitive advantage. *Psychon. Bull. Rev.* 23, 1221–1230. doi: 10.3758/s13423-015-0981-6
- Zuk, J., Benjamin, C., Kenyon, A., and Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS ONE* 9:e99868. doi: 10.1371/journal.pone.0099868

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Poarch. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Does Extreme Language Control Training Improve Cognitive Control? A Comparison of Professional Interpreters, L2 Teachers and Monolinguals

Lize Van der Linden^{1*}, Eowyn Van de Putte², Evy Woumans², Wouter Duyck² and Arnaud Szmalec^{1,2,3}

OPEN ACCESS

Edited by:

Peter Bright,
Anglia Ruskin University,
United Kingdom

Reviewed by:

Claudia C. von Bastian,
University of Sheffield,
United Kingdom
Sara Incera,
Eastern Kentucky University,
United States

*Correspondence:

Lize Van der Linden
lize.vanderlinden@uclouvain.be

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 30 June 2018

Accepted: 28 September 2018

Published: 23 October 2018

Citation:

Van der Linden L, Van de Putte E,
Woumans E, Duyck W and Szmalec A
(2018) Does Extreme Language
Control Training Improve Cognitive
Control? A Comparison
of Professional Interpreters, L2
Teachers and Monolinguals.
Front. Psychol. 9:1998.
doi: 10.3389/fpsyg.2018.01998

¹ Psychological Sciences Research Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium, ² Department of Experimental Psychology, Ghent University, Ghent, Belgium, ³ Institute of Neuroscience, Université catholique de Louvain, Louvain-la-Neuve, Belgium

There is currently a lively debate in the literature whether bilingualism leads to enhanced cognitive control or not. Recent evidence suggests that knowledge of more than one language does not always suffice for the manifestation of a bilingual cognitive control advantage. As a result, ongoing research has focused on modalities of bilingual language use that may interact with the bilingual advantage. In this study, we explored the cognitive control performance of simultaneous interpreters. These highly proficient bilinguals comprehend information in one language while producing in the other language, which is a complex skill requiring high levels of language control. In a first experiment, we compared professional interpreters to monolinguals. Data were collected on interference suppression (flanker task), prepotent response inhibition (Simon task), and short-term memory (digit span task). The results showed that the professional interpreters performed similarly to the monolinguals on all measures. In Experiment 2, we compared professional interpreters to monolinguals and second language teachers. Data were collected on interference suppression (advanced flanker task), prepotent response inhibition (advanced flanker task), attention (advanced flanker task), short-term memory (Hebb repetition paradigm), and updating (*n*-back task). We found converging evidence for our finding that experience in interpreting may not lead to superior interference suppression, prepotent response inhibition, and short-term memory. In fact, our results showed that the professional interpreters performed similarly to both the monolinguals and the second language teachers on all tested cognitive control measures. We did, however, find anecdotal evidence for a (small) advantage in short-term memory for interpreters relative to monolinguals when analyzing composite scores of both experiments together. Taken together, the results of the current study

suggest that interpreter experience does not necessarily lead to general cognitive control advantages. However, there may be small interpreter advantages in short-term memory, suggesting that this might be an important cognitive control aspect of simultaneous interpreting. The results are discussed in the light of ongoing debates about bilingual cognitive control advantages.

Keywords: bilingualism, interpreting, cognitive control, language control, bilingual experience

INTRODUCTION

Recent research has shown that certain cognitively demanding activities, such as playing video games, playing music, and mastering chess, may be beneficial to human cognition, beyond the domain of practice (e.g., Reingold et al., 2001; Bialystok, 2006; Schroeder et al., 2016). Gaining expertise in a certain skill may lead to a transfer of the acquired abilities to other behaviors that involve the same processes, often related to cognitive control. Cognitive control is an umbrella term for the cognitive processes that guide goal-directed behavior. Knowing and using a second language (L2) in daily life, or bilingualism (Grosjean, 2010), may also have beneficial effects on cognition. Bilinguals outperform monolinguals in learning novel words (e.g., Kaushanskaya and Marian, 2009; Nair et al., 2017). Similarly, bilinguals outperform monolinguals on non-verbal tasks that require different cognitive control processes, like conflict resolution, attention, shifting, updating, and working memory, for example (e.g., Bialystok et al., 2006, 2008; Costa et al., 2008; Prior and Macwhinney, 2010; Luo et al., 2013). One explanation for these bilingual advantages is that using multiple languages requires a mechanism to select (words in) the target language while avoiding interference from the other known language. There is in fact compelling evidence that both languages of bilinguals are always simultaneously active in their mind (e.g., Colomé, 2001; Duyck and Warlop, 2009; Van Assche et al., 2009). Bilinguals therefore need to control (inhibit) activation of the non-target language to use the intended language (Green, 1998). The mechanisms that allow this language control are believed to be domain-general and hence, not specific to the linguistic domain. Using multiple languages in daily life might therefore train domain-general cognitive control, in a way similar to mastering chess (Bialystok et al., 2012).

Although there is abundant evidence supporting this bilingual advantage, quite a few recent studies have also questioned its existence (e.g., Morton and Harper, 2007; Hilchey and Klein, 2011; Paap and Sawi, 2014; Paap et al., 2015). Paap and Sawi (2014), for example, compared highly proficient bilinguals and monolinguals on tasks that require conflict resolution, attention, and shifting. Across the three tested cognitive control processes, there was no evidence for a bilingual advantage. These and several similar findings have led some researchers to claim that the bilingual advantage does not exist, and the inconsistent results have caused a lively debate about the correctness of the bilingual advantage hypothesis (see Barac et al., 2014, for a review). To make things even more complex, in a meta-analysis on the issue, de Bruin et al. (2015) showed that the bilingual advantage is a reliable effect across studies, but also, taking into account non-published reports, that a publication bias exists against

null-findings. This publication bias was confirmed by a recent meta-analysis of Lehtonen et al. (2018). Before correcting for the bias, the authors observed a very small bilingual advantage for conflict resolution, shifting, and working memory. However, no evidence for a bilingual advantage remained after controlling for the publication bias.

Woumans and Duyck (2015) suggested that research on the bilingual advantage should move away from the rather unfruitful debate of whether or not the advantage exists. According to these authors, bilingualism may lead to an advantage in cognitive control, but only for some bilingual profiles. Future work should therefore aim to define the precise characteristics of bilingualism that may benefit cognitive control. Bilingual experience can vary in several ways. For example, bilinguals have varying levels of L2 proficiency, they can differ in their language switching frequency, or in the age at which they acquired their L2. One of these many characteristics that can vary across bilinguals might be the key to enhanced cognitive control. Several other researchers made similar claims (e.g., Prior and Gollan, 2011; Green and Abutalebi, 2013; Woumans et al., 2015; Verreyt et al., 2016). According to the adaptive control hypothesis (Green and Abutalebi, 2013), for instance, the interactional context in which bilinguals use their languages is important. Specifically, those bilinguals who use their languages within the same context (i.e., dual-language context) require a high level of cognitive control to keep their languages separated. This is less true for bilinguals who use their languages in different contexts (i.e., single-language context) or for bilinguals who mix their languages within a sentence or conversation (i.e., dense code-switching context). Using multiple languages in a dual-language context might thus require and hence, train cognitive control processes more than using these languages in single-language or dense code-switching contexts. This hypothesis has been corroborated by recent work showing that bilinguals in dual-language contexts outperform bilinguals in single-language contexts in cognitive flexibility (Hartanto and Yang, 2016). Another factor that has been recently suggested as crucial for the development of a bilingual advantage, is the frequency at which bilinguals switch between their languages (e.g., Prior and Gollan, 2011; Woumans et al., 2015; Verreyt et al., 2016). That is, those bilinguals who switch more frequently between their languages may show more cognitive control advantages than those who switch less often. Language switching requires adaptations in language control (reactivating and inhibiting languages), which each time involves the recruitment of cognitive control. The frequent recruitment of cognitive control for language switching might then train this mechanism. In their study, Verreyt et al. (2016) compared two groups of highly proficient bilinguals (non-frequent and frequent

language-switchers) and a group of low proficient bilinguals on conflict resolution. They found a bilingual advantage for the frequent language-switchers over the other groups. These results were further supported by a study of Woumans et al. (2015), who observed a positive correlation between language-switching frequency and conflict resolution. It should be noted, however, that other studies did not obtain evidence for moderating effects of characteristics like language-switching frequency on the bilingual advantage (Yim and Bialystok, 2012; Paap et al., 2017). The bilingual advantage debate therefore continues, and further research and data points are mandatory to understand which specific aspects of bilingualism might lead to enhanced cognitive control. Prior work nevertheless suggests that the bilingual advantage is more likely to emerge in those bilinguals who use their languages in a dual-language context and who switch regularly between their languages. In other words, if a bilingual advantage exists, those bilinguals who require higher levels of language control are more likely to develop it.

What is arguably the most demanding type of bilingualism in terms of language control is simultaneous interpreting. Interpreters have to comprehend incoming speech in a source language and reformulate (translate) this message in the target language, while simultaneously producing a previously translated message. Thus, they have to speak in one language while processing, manipulating and storing considerable amounts of incoming information in the other language. It is estimated that interpreters are speaking in one language while simultaneously comprehending in the other language about 70% of the time (Chernov, 1994). This contrasts with everyday bilingual language practice in which bilinguals typically use only one language at a time. Furthermore and importantly, the languages may not be mixed. The quality of simultaneous interpreting depends in part on the output in the target language. A non-target language intrusion might thus have more negative consequences for interpreters than for other bilinguals, making efficient language control extremely important. This high level of language control requires several cognitive processes (conflict resolution, attention, updating, and short-term memory) to be used in parallel under heavy time pressure (Christoffels et al., 2006; Köpcke and Nespoulous, 2006). As language control is assumed to develop cognitive control, expertise in simultaneous interpreting could thus cause interpreters to become experts in several cognitive control processes (Yudes et al., 2011).

Relatively little is known today about the effects of proficiency in simultaneous interpreting on language control, or, more generally, cognitive control. First, there are some inconsistent results regarding the bilingual advantage for interpreters with respect to the cognitive control processes that are often linked to bilingualism (see **Table 1** for an overview). In a study of Christoffels et al. (2006), professional interpreters and L2 teachers performed similarly on a basic cognitive control task measuring attention. Yudes et al. (2011) found that professional interpreters outperformed both bilinguals and monolinguals on cognitive flexibility, but not on conflict resolution. An advantage in conflict resolution for interpreters was, however, found by Woumans et al. (2015). In their study, monolinguals, unbalanced bilinguals, balanced bilinguals, and student interpreters were

compared. They observed that all bilingual groups outperformed monolinguals on speed of conflict resolution. Furthermore, student interpreters were more accurate than unbalanced, but not than balanced bilinguals. The latter results provide support for the bilingual advantage hypothesis by showing that being highly proficient in multiple languages yields cognitive control advantages, at least in conflict resolution. However, the results of Woumans and colleagues also suggest that experience in simultaneous interpreting may not lead to accumulated advantages in conflict resolution over and above the advantages proper to bilingualism. Morales et al. (2015) found that professional interpreters were better in updating than highly proficient bilinguals, but again they found no difference in terms of conflict resolution. These results provide further support for the finding that professional interpreters have no accumulated advantage in conflict resolution. Experience in simultaneous interpreting might, however, lead to better updating abilities relative to other bilinguals. Therefore, Henrard and Van Daele (2017) compared professional interpreters, translators and monolinguals on a wide range of cognitive control processes (conflict resolution, updating, working memory, speed of information processing, and flexibility). Professional interpreters and translators are both highly proficient bilinguals who have to translate a message from a source language into a target language. However, interpreting is an online process under important time pressure, as interpreters have to comprehend, translate and produce simultaneously. This is not the case for translators, who can process the information in the source language before reformulating the message in the target language. Furthermore, interpreters require a lot of cognitive resources in parallel, as they have to translate while processing a lot of incoming information. Translators, on the other hand, sequentially process the incoming information, translate the message, and produce the output, which requires less cognitive resources. Interpreters therefore might deliberately ignore less relevant information to cope with the time pressure and have to update their memory more than translators. The results of Henrard and Van Daele showed that both bilingual groups outperformed the monolinguals on all cognitive control measures. Moreover, interpreters performed better than translators on all cognitive control aspects, except shifting. Together, these results suggest that experience in interpreting stimulates cognitive control abilities. Research conducted thus far is, however, inconclusive about which cognitive control processes might be specifically enhanced and whether or not there are accumulated advantages for interpreters over other bilingual populations.

There are also some studies that examined the effects of interpreting on other cognitive control aspects, such as short-term memory (STM; Padilla et al., 2005; Christoffels et al., 2006; Signorelli et al., 2011; Timarova et al., 2014; Rosiers et al., 2017). STM refers to the cognitive system to memorize information (e.g., digits) for a brief period of time (Kolb and Wishaw, 2009). The importance of STM for simultaneous interpreting makes intuitive sense. As noted earlier, interpreters have to temporarily memorize information in the source language while translating it in the target language. Christoffels et al. (2006) found that interpreters performed better on STM tasks than

TABLE 1 | Overview of the studies on cognitive control abilities of interpreters.

Paper	Tested cognitive control process	Participants	Main results
Christoffels et al. (2006)	(i) Attention (ii) STM	Professional interpreters L2 teachers Bilingual students	(i) Students outperformed interpreters on attention (ii) Interpreter advantage at the level of STM
Yudes et al. (2011)	(i) Cognitive flexibility (ii) Conflict resolution (iii) STM	Professional interpreters Bilinguals Monolinguals	(i) Interpreter advantage at the level of cognitive flexibility and STM; bilinguals = monolinguals (ii) No group differences at the level of conflict resolution
Woumans et al. (2015)	Conflict resolution	Student interpreters Monolinguals Unbalanced bilinguals Balanced bilinguals	(i) All bilingual groups better conflict resolution in terms of reaction times (ii) Interpreters more accurate than unbalanced, but not than balanced bilinguals
Morales et al. (2015)	(i) Updating	Professional interpreters	(i) Interpreter advantage in updating
	(ii) Conflict resolution	Highly proficient bilinguals	(ii) No group differences in conflict resolution
Henrard and Van Daele (2017)	(i) Conflict resolution	Professional interpreters	(i) Bilingual groups outperformed monolinguals on all tested cognitive control measures (ii) Interpreters outperformed translators on all cognitive control measures, except on cognitive flexibility
	(ii) Speed of information processing	Translators	
	(iii) Cognitive flexibility	Monolinguals	
	(iv) Updating		
Rosiers et al. (2017)	Conflict resolution	Student interpreters	No evidence for a bilingual advantage in conflict resolution
		Student translators	
		Students multilingual communication	
Liu et al. (2004)	STM	Professional interpreters Beginner student interpreters Advanced student interpreters	No group differences in STM
Köpke and Nespoulous (2006)	(i) Conflict resolution	Professional interpreters	(i) No group differences in STM and conflict resolution (ii) Novice interpreters advantage in working memory
	(ii) STM	Student interpreters	
	(iii) Working memory	Students Bilinguals	

both highly proficient L2 teachers and younger unbalanced bilingual students. Other studies, however, failed to find support for better STM in professional interpreters (e.g., Liu et al., 2004; Köpke and Nespoulous, 2006). Liu et al. (2004), for example, found that student interpreters had similar STM as professional interpreters, despite the fact that the professionals excelled the students in interpreting skills. This finding suggests that accumulating expertise in interpreting does not further train STM. The performance of the student and professional interpreters was not compared to a monolingual control group. This leaves open the question whether or not simultaneous

interpreting training develops STM. In another study, Köpke and Nespoulous (2006) assessed the STM of professional interpreters, second-year interpreting students, and two control groups (students and bilinguals). While their data indicated that student interpreters outperformed the control groups, this was not true for professional interpreters who had at least 4 years of professional experience. While these two studies suggest that professional interpreters might not have better developed STM than monolinguals or other bilinguals, they might also be explained by other factors. The authors argued that an effect of expertise in simultaneous interpreting may have

been obscured by a confounding effect of age, for example. Nevertheless, observing better or similar performance in STM for student interpreters than for professional interpreters is rather remarkable if simultaneous interpreting relies heavily on STM that further develops with accumulating experience. In their correlational study, Timarova et al. (2014) also only observed a weak association between STM and expertise in simultaneous interpreting. These findings suggest that STM may not be strongly taxed upon during interpreting.

THE PRESENT STUDY

This study aims to investigate whether special, advanced expertise in L2 benefits cognitive control. We therefore assessed the performance of professional interpreters and L2 teachers on multiple aspects of cognitive control that have been linked to bilingualism. The selection of the different processes was based on the scientific findings about cognitive control in bilinguals and interpreters described above. One major difference between the present study and prior work on interpreters, though, is that we brought all the different aspects of cognitive control together in one study, in multiple groups of advanced L2 users. Indeed, of the relatively few studies examining the cognitive abilities of interpreters, the majority focused on only one or two cognitive control processes (for an exception, see Henrard and Van Daele, 2017).

In Experiment 1, we used three extensively used cognitive control tasks to compare conflict resolution and STM between professional interpreters and monolinguals. Friedman and Miyake (2004) proposed that different conflict resolution tasks may reveal different results because they rely on different conflict resolution types. Two types may be important for bilingualism. Resistance to interference is a type of conflict resolution that allows an individual to focus on the task at hand and to avoid distraction from irrelevant information. Interpreters must resist from being distracted not only by the co-activation of the non-target language, just like typical bilinguals, but also by distractions such as the incoming information in the source language, which competes for attentional resources with the message they are formulating. Furthermore, given that both of their languages have to be active in parallel, interpreters may experience more dual-language competition than typical bilinguals. The second conflict resolution type is prepotent response inhibition. Automatic responses can be caused by developed routines (automatized behavior), or by a triggering response. Bilinguals need prepotent response inhibition to avoid using false cognates, for example. False cognates are word-forms that exist in both languages, but that have a different meaning in each language (e.g., the English-Dutch *room*, which is cream in Dutch). A typical example of this type of conflict resolution in the context of interpreting is the postponement of reformulating (translation) until sufficient information is available to allow for planning (e.g., to avoid interpreting errors caused by syntactic ambiguous sentences). If simultaneous interpreting trains conflict resolution, we anticipate interpreters to outperform monolinguals on both conflict resolution types.

We also compared STM of interpreters and monolinguals. As already noted, the ability to temporarily memorize a considerable amount of information is very important for simultaneous interpreting. Furthermore, bilingualism may also lead to better STM (e.g., Grundy and Timmer, 2016). We therefore predict interpreters to have a better STM than monolinguals.

In Experiment 2, we further tested the bilingual cognitive control advantage by introducing a third group of participants, namely L2 teachers. L2 teachers are, like professional interpreters, highly proficient bilinguals, but, as the monolinguals, they have no experience in simultaneous interpreting. They can therefore be considered as typical, highly proficient bilinguals. Assessing different cognitive control processes within the same groups of interpreters and comparing their performance to that of L2 teachers and monolinguals will allow us to determine which aspects of cognitive control are specifically developed by bilingualism and by simultaneous interpreting, more particularly.

Experiment 1

In Experiment 1, we compared professional interpreters who had at least 4 years of professional experience to monolinguals, using three well-established tasks previously found to be sensitive to the bilingual cognitive control advantage. First, we used the flanker task (Eriksen and Eriksen, 1974) to measure interference suppression. Costa et al. (2008), for instance, found an advantage in interference suppression, reflected in smaller flanker congruency effects for bilinguals than for monolinguals. In their study, the attention network task (ANT) was used, which is a flanker task embedded in a cue reaction time task. It explores three attentional networks, namely cognitive control, alerting, and orienting. With respect to the cognitive control component, which is relevant here, congruent trials were comprised of a target and a flanking arrow pointing in the same direction, whereas a target arrow pointing in one direction and flanking arrows pointing in the other direction were presented on incongruent trials. The difference between congruent and incongruent trials (flanker congruency effect) was taken as a marker of interference suppression.

Second, we assessed prepotent response inhibition with the Simon task (Simon and Wolf, 1963). In this task, participants respond on the color (green or red) of the stimulus, using either their left or right hand, while ignoring its location (left or right). The Simon task includes congruent and incongruent conditions, as this task is based on stimulus-response compatibility. The difference between congruent and incongruent trials (Simon effect) is a marker of prepotent response inhibition. As for the flanker congruency effect, some prior work has reported smaller Simon effects for bilinguals than for monolinguals (e.g., Bialystok et al., 2004) and for interpreters than for monolinguals (Woumans et al., 2015).

Finally, to examine whether simultaneous interpreting improves STM, we used the digit span test. In this task, sequences of digits are presented for immediate serial recall. The length of the sequences gradually increases, making memorization of the sequences more difficult. Some prior work found that bilinguals have a better STM than monolinguals (Bialystok et al., 2008; Morales et al., 2013). Furthermore, there is evidence that

interpreters have better STM than various other populations (e.g., Christoffels et al., 2006).

Materials and Methods

Participants

We recruited 52 participants, divided into two groups (27 monolinguals and 25 interpreters). All participants reported having no language, hearing, uncorrected visual, or neurological problems. Informed consent was obtained under a protocol approved by the ethical committee at Ghent University (Belgium). Objective language proficiency tests could not be used because interpreters had different languages as native language (L1) and L2. Given that self-evaluation correlates strongly with objective measures (Marian et al., 2007; Luk and Bialystok, 2013), participants self-rated their language proficiency. Further, we administered the short untimed 12 item-version of the Advanced Progressive Matrices (Bors and Stokes, 1998) as a measure of intelligence. This version has a strong correlation with the complete 48 item-version (Raven et al., 1998).

Detailed demographic information is reported in **Table 2**. The 27 monolinguals spoke French as L1 and acquired anecdotal knowledge of an L2 (Dutch or English) through formal education. That is, they indicated having low proficiency in L2 and rarely used this language. The 25 interpreters had at least 4 years of professional experience in simultaneous interpreting. They spoke a variety of languages as L1 (23 Dutch, 1 Portuguese, and 1 French) and L2 (1 Dutch, 2 German, 5 French, 1 Danish, 9 English, 1 Portuguese, and 6 Spanish), but they were all highly proficient in Dutch and used this language for their profession.

T-tests comparing the demographic information between the monolinguals and interpreters are reported in **Table 2**. We relied on Bayes factors (BF_{10}) for interpreting our results. Null-hypothesis (H_0) significance tests and their accompanied p -values have several shortcomings and more reliable alternative approaches, such as BF_{10} , have been suggested (Gallistel, 2009; Dienes, 2011; Nuzzo, 2014). BF_{10} compares the fit of the data under H_0 (there is no effect) compared to the alternative hypothesis (there is an effect; H_1). BF_{10} thus provides a quantification of the degree to which the data support either hypothesis. Values greater than 1 indicate increasing evidence for H_1 over H_0 , values smaller than 1 the reverse. We relied on the

guidelines proposed by Jeffreys (1961) for interpreting BF_{10} (see **Table 3**). The monolinguals were matched to the interpreters in terms of age (substantial evidence), male/female ratio (anecdotal evidence), intelligence (anecdotal evidence), and L1 proficiency (anecdotal evidence). As expected, there was decisive evidence that interpreters had a higher proficiency in their L2 than monolinguals and that interpreters used their L2 more frequently than monolinguals.

Stimuli and procedure

Participants were tested individually in a quiet room. They were asked to carry out the intelligence test, two computerized cognitive control tasks (flanker task and Simon task), and the digit span task in a counterbalanced order. Task instructions were given in French for monolinguals and in Dutch for interpreters, because monolinguals were recruited in French-speaking Belgium, and interpreters in Dutch-speaking Belgium.

Flanker task. The stimuli were white arrows on a black screen that were flanked by four white distractor arrows. The distractor arrows could either point in the same (congruent) or the opposite (incongruent) direction as the target arrow (e.g., congruent trial <<<<< and incongruent trial <<<><<).

The task was programmed using Tscope (Stevens et al., 2006). Participants were asked to indicate the direction of the central arrow by pressing the left (a) or right (p) button on an azerty keyboard. Each trial began with a centered 500 ms fixation cross, followed by the stimulus for 1500 ms or until a response was made. There was 500 ms inter trial interval. The task began with 10 practice trials with feedback, followed by two blocks of 100 trials. Each block contained an equal amount of randomly presented congruent and incongruent trials.

Simon task. Participants saw colored dots on the left or right side of the screen. They were asked to indicate as quickly and accurately as possible whether the dot was green or red by pressing the left (right) or right (left) key on the keyboard, respectively. Response mapping was counterbalanced across participants. Position and color elicited either the same (congruent trials) or different responses (incongruent trials).

The task was presented via Tscope (Stevens et al., 2006). Each trial began with a 500 ms fixation cross, followed by a 500 ms

TABLE 2 | Demographic data of the different participant groups of Experiment 1.

	Monolinguals	Interpreters	Test	BF_{10}
N	27	25		
Male/female ratio	5/22	9/16	$\chi^2(1) = 2.20$	0.78
Age (years)	48.37 (8.54)	49.76 (7.99)	$t < 1$	0.32
Raven (raw score)	7.37 (2.76)	8.36 (2.12)	$t(50) = -1.44$	0.65
L1 proficiency (20-point scale)	19.52 (1.40)	19.96 (0.20)	$t(27.15) = -1.62$	0.76
L1 Age of acquisition (years)	0.00 (0.00)	0.08 (0.40)	$t(24.00) = 1.00$	0.43
L1 frequency of use (%)	95.56 (6.94)	66.07 (13.68)	$t(34.97) = 9.68^{***}$	> 100
L2 proficiency (20-point scale)	5.07 (6.01)	17.44 (1.98)	$t(31.99) = -10.12^{***}$	> 100
L2 Age of acquisition (years)	14.11 (7.42)	11.16 (6.16)	$t(50) = 1.55$	0.75
L2 frequency of use (%)	4.44 (6.94)	33.93 (13.68)	$t(34.97) = -9.68^{***}$	> 100

SD are shown between parentheses. BF_{10} , Bayes factor in favor of the alternative hypothesis. $*p < 0.05$; $**p < 0.01$; $***p < 0.001$.

TABLE 3 | Interpretation of Bayes Factors (BF_{10}) as evidence for null hypothesis (H_0) and alternative hypotheses (H_1).

BF_{10}	Support for hypothesis
<0.01	Decisive evidence for H_0
0.03–0.01	Very strong evidence for H_0
0.10–0.03	Strong evidence for H_0
0.33–0.10	Substantial evidence for H_0
0.33–1	Anecdotal evidence for H_0
1	No evidence
1–3	Anecdotal evidence for H_1
3–10	Substantial evidence for H_1
10–30	Strong evidence for H_1
30–100	Very strong evidence for H_1
>100	Decisive evidence for H_1

blank screen. Next, a red or green dot appeared on the left or right side of the screen for 1500 ms or until a response was made, followed by a 500 ms inter trial interval. The task started with 10 practice trials with feedback, followed by two blocks of 100 trials. Each block contained an equal amount of randomly presented congruent and incongruent trials.

Digit span task. Series of two to nine numbers (one to nine) were presented in ascending order, with two trials per sequence length. Each number in a sequence was orally presented at a rate of 1000 ms. At the end of a sequence, participants were asked to immediately recall the sequence. A sequence was scored as correct if the sequence was repeated in its correct serial order. Sequences were presented in French for monolinguals and in Dutch for interpreters. The task ended when two trials at a particular sequence length were incorrectly reproduced. The number of correctly recalled sequences was calculated (maximum score: 16).

Results

Incorrect responses and outliers were excluded for all analyses on reaction times (RTs). Outlier RTs were trimmed individually by calculating a mean RT for each condition and excluding responses that had an RT of 2.5 SD from this mean. Unless stated otherwise, data were analyzed by fitting generalized mixed-effects models with maximum likelihood estimation on individual trials, using the `glmer` function from the `lme4` package in R (Bates et al., 2015). Models on RT data assumed an Inverse Gaussian distribution, and a linear relationship between the predictors and RT (Lo and Andrews, 2015). We initially applied the simplest model, which included the fixed effects, their interactions and the random effect of participants. We included by-participant random slopes if conducted maximum likelihood model comparisons showed that the data justified their inclusion. Planned comparisons were performed using the `multcomp` package (Hothorn et al., 2008). To calculate BF_{10} for main and interaction effects, we used the Bayesian Information Criteria technique (Wagenmakers, 2007). For planned comparisons, we used Bayesian t -tests with a default Cauchy prior width of $r = 0.707$ for effect size on H_1 (Rouder et al., 2009).

Flanker task

The data of one monolingual were excluded because he had an ACC of less than 50% (chance-level) on congruent trials. The ACC data are shown in **Figure 1A**. For ACC, the model included Group (monolingual, interpreter), Congruency (congruent, incongruent) and their interaction as fixed effects, Participant as random effect and by-Participant random slopes of Congruency. We observed decisive evidence for a main effect of Congruency, $\chi^2(1) = 26.58$, $p < 0.001$, $BF_{10} > 100$ (flanker congruency effect). There was anecdotal evidence against an effect of Group, $\chi^2 < 1$, $BF_{10} = 0.39$, and against an interaction of Congruency and Group, $\chi^2(1) = 1.52$, $p = 0.22$, $BF_{10} = 0.78$.

Of the RT data, 2.45% (248 trials) were outliers. The number of outlier RT trials was similar for the interpreters ($n = 125$) and the monolinguals ($n = 118$), $t < 1$. The trimmed RT data are summarized in **Figure 1B**. The same model as for ACC data was used for analyzing RTs. We observed decisive evidence for an effect of Congruency, $\chi^2(1) = 42.54$, $p < 0.001$, $BF_{10} > 100$ (flanker congruency effect). There was very strong evidence against an effect of Group, $\chi^2 < 1$, $BF_{10} = 0.02$, and against an interaction of Congruency and Group, $\chi^2 < 1$, $BF_{10} = 0.01$.

Simon task

Figure 2A summarized the ACC data. For ACC, the model included Group (monolingual and interpreter), Congruency (congruent and incongruent) and their interaction as fixed effects, and Participant as random effect. We observed decisive evidence for a main effect of Congruency, $\chi^2(1) = 83.86$, $p < 0.001$, $BF_{10} > 100$ (Simon effect). There was strong evidence against an effect of Group, $\chi^2(1) = 3.61$, $p = 0.06$, $BF_{10} = 0.06$, and against an interaction of Congruency and Group, $\chi^2 < 1$, $BF_{10} = 0.01$.

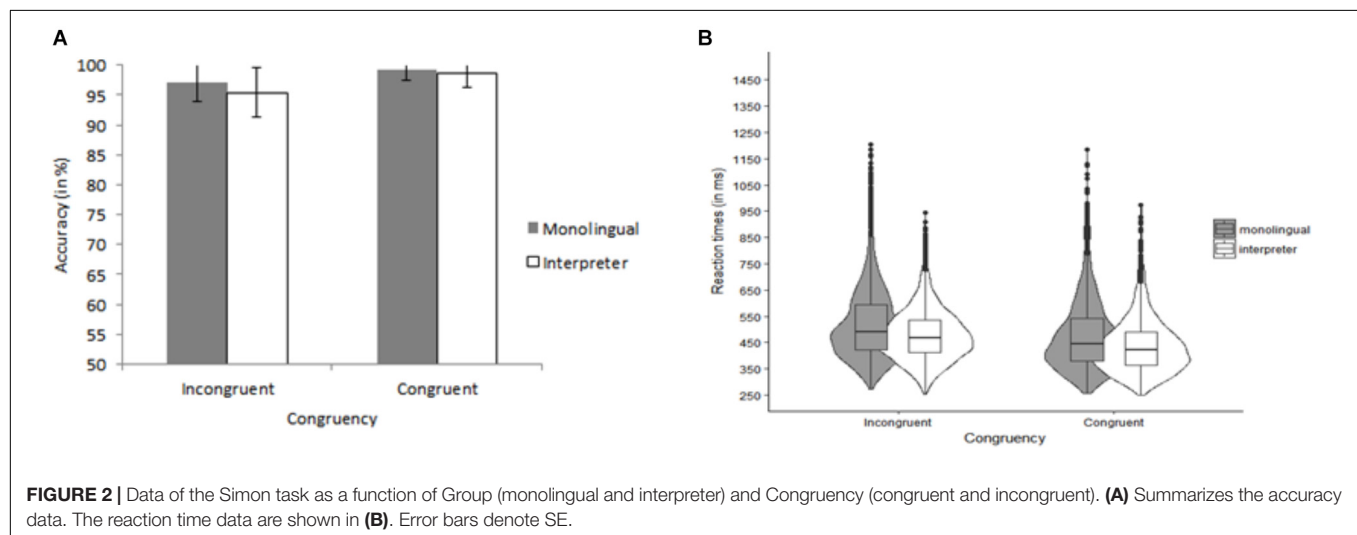
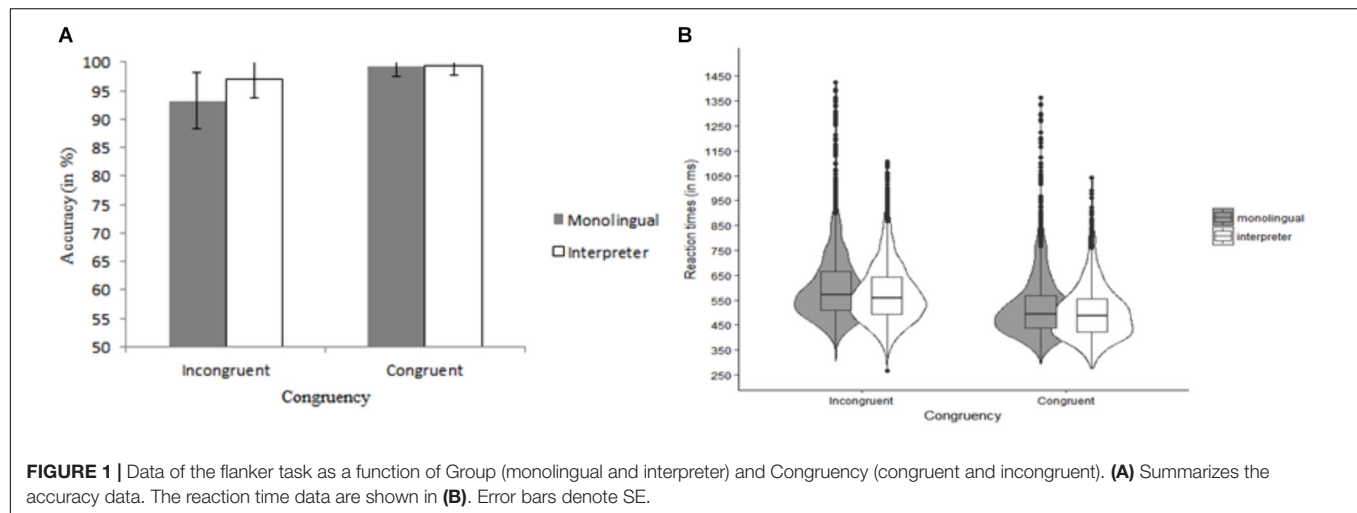
Of the RT data, 2.65% (269 trials) were outliers. The number of excluded trials was similar for interpreters ($n = 136$) and monolinguals ($n = 133$), $t < 1$. **Figure 2B** shows the trimmed RT data. The model on RT contained Group (monolingual and interpreter), Congruency (congruent and incongruent) and their interaction as fixed effects, Participant as random effect and by-Participant random slopes of Congruency. There was decisive evidence for an effect of Congruency, $\chi^2(1) = 37.10$, $p < 0.001$, $BF_{10} > 100$ (Simon effect), and strong evidence against an effect of Group, $\chi^2(1) = 3.49$, $p = 0.06$, $BF_{10} = 0.05$. There was very strong evidence against an interaction of Congruency and Group, $\chi^2 < 1$, $BF_{10} = 0.01$.

Digit span task

Digit span performance is summarized in **Figure 3**. An independent samples t -test revealed anecdotal evidence against a group difference in digit span performance, $t(50) = -1.49$, $p = 0.14$, $BF_{10} = 0.69$.

Summary of Results

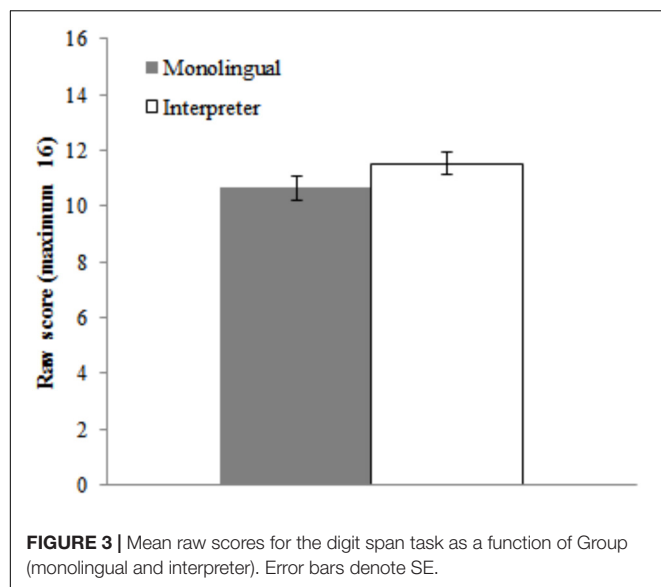
If simultaneous interpreting modulates the bilingual advantage, we would predict better cognitive control for the interpreters. Our data did, however, not reveal evidence for a difference between interpreters and monolinguals on any of the tested cognitive control measures. That is, there were no differences on the flanker congruency effect, indicating similar interference



suppression. We also could not observe a group difference on the Simon effect, indicating similar prepotent response inhibition. Finally, the two groups had comparable performance on the digit span task, indicating similar STM. Note, however, that the lack of evidence in favor of group differences was accompanied by decisive evidence for a flanker congruency effect and for a Simon effect. This indicates that the tasks were valid, and sufficiently sensitive, to measure the underlying cognitive control processes.

One might argue that we could not obtain evidence for an interpreter advantage over monolinguals on conflict resolution and STM because interpreters use a different language control mechanism than other, more typical bilinguals (Yudes et al., 2011). It is beyond doubt that language control is more important for interpreters than for other bilinguals, but the specific cognitive control processes involved to achieve language control may differ. There is evidence that both languages are active in parallel in the mind of interpreters and that interpreters therefore experience interference of the non-target language while speaking, just like other bilinguals (Rodríguez-Fornells et al., 2005; Kaushanskaya and Marian, 2007). Interpreters may,

however, differ from more typical bilinguals in how they manage cross-language activation. Bilinguals are assumed to select the appropriate language and avoid non-target language interference by inhibiting the latter language (Green, 1998; Dijkstra and Van Heuven, 2002). Interpreters, however, have to maintain both languages active in parallel, one for comprehension and one for speaking. There are indeed some indications that interpreters do not use inhibition to control their languages (Ibáñez et al., 2010), but it is still unknown how interpreters then manage their languages. Nevertheless, if bilinguals and interpreters control their two languages differently, this can lead to differences in (some) cognitive control abilities. That is, simultaneous interpreting may train different aspects of cognitive control than more typical bilingualism. The scope of Experiment 2 was therefore to investigate how the potential cognitive control advantages for interpreters differ from cognitive control advantages associated with more typical bilingual language use. We therefore again examined whether professional interpreters have cognitive control advantages over monolinguals, but we assessed more cognitive control processes that are important for



simultaneous interpreting (conflict resolution, attention, STM, and updating). Furthermore, we additionally compared the performance of both groups on each cognitive control measure to that of L2 teachers.

Experiment 2

In experiment 2, we further investigated the cognitive implications of simultaneous interpreting. We compared professional interpreters with a well-matched group of L2 teachers, based on the following methodological considerations. First, both professional interpreters and L2 teachers are rather rare populations that have very high levels of L2 proficiency. Both populations use their languages for their profession, which makes them frequent language-switchers in a dual-language context. Finally, they also share a similar educational background, as they both have a degree in L2 and share an interest in language. One important difference between interpreters and L2 teachers, though, is the amount of interpreting experience they have, and therefore the amount of language control training that can be expected. It is reasonable to assume that interpreters require higher levels of language control than L2 teachers. There may also be qualitative differences between the cognitive control processes involved in language control, which can lead to differences between interpreters and L2 teachers on these cognitive control abilities. Interpreters have to resolve conflict, store considerable amounts of information in STM, and update their memory, without confusing their languages. Assessing these different cognitive control processes within the same groups of interpreters and comparing their performance with that of L2 teachers and monolinguals will allow us to determine which cognitive control aspects are specifically developed by bilingualism and by experience in simultaneous interpreting, more particularly.

The advanced flanker task (Emmorey et al., 2008) was used to measure two types of conflict resolution, as it is a combination of the flanker task and the go/no-go task, measuring

resistance to interference and prepotent response inhibition, respectively. We anticipate a bilingual advantage in both conflict resolution types. If interpreting involves resistance to interference or prepotent response inhibition, we also predict interpreters to outperform L2 teachers because of accumulated practice. Conversely, if interpreters do not use inhibition for language control, in contrast to L2 teachers, we predict L2 teachers to outperform both monolinguals and interpreters. Additionally, the advanced flanker task allowed us examining another cognitive control aspect, namely attention. It almost goes without saying that high levels of attention are important during interpreting, as it enables an individual to speak, listen, and manipulate information simultaneously. We therefore anticipate interpreters to outperform both L2 teachers and monolinguals in attentional abilities.

The third cognitive control process assessed was STM, using Hebb learning. Hebb learning is an immediate serial recall task in which sequences of items (e.g., phonemes) are presented. We chose this task because phoneme recall is not dependent upon prior language knowledge. This is important because functional STM may not be the same in bilinguals' L1 and L2 (Service et al., 2002). In Experiment 1, the monolinguals and interpreters carried out the digit span test in different languages, which may have obscured the detection of possible group differences. Given the importance of STM for simultaneous interpreting, we anticipate interpreters to outperform L2 teachers and monolinguals. We also predict L2 teachers to outperform monolinguals, in line with prior work suggesting bilingual advantages on STM (Bialystok et al., 2008; Morales et al., 2013; Grundy and Timmer, 2016). Furthermore, it has been shown that Hebb learning can be considered as an analog of novel word-form learning (e.g., Szmalec et al., 2009; Smalle et al., 2017). When a particular sequence of phonemes is repeated, performance for the repeating Hebb sequence improves relative to non-repeating filler sequences (Hebb, 1961). This finding (Hebb repetition effect) reflects the gradual transfer of newly acquired serial-order information from STM to long-term memory, which underlies novel word learning. Given the indications that bilinguals outperform monolinguals in learning novel words (Kaushanskaya and Marian, 2009; Nair et al., 2017) and that better STM has been associated with superior word learning abilities in bilinguals (Papagno and Vallar, 1995; Kaushanskaya, 2012), we also investigated whether interpreters outperform other groups on the Hebb repetition effect.

The fourth and final aspect of cognitive control tested here was updating, using the *n*-back task (Collette et al., 2001; Oberauer, 2005; Szmalec et al., 2011). A typical feature of STM is that its capacity is limited (Cowan, 2005). Thus, when confronted with a large stream of incoming information, individuals must temporarily store subsets of information and successively update STM as more information becomes available. This is exactly what needs to be done during simultaneous interpreting: a continuous stream of incoming information in the source language needs to be temporarily held in STM while it is being reformulated in the target language, and then “forgotten” in order to store and reformulate new information in the source language. We therefore predict interpreters to have better updating abilities

than both L2 teachers and monolinguals. Given that prior research has shown that bilinguals outperform monolinguals in updating (e.g., Bialystok et al., 2006), we also anticipate L2 teachers to have better updating abilities than monolinguals.

To summarize, using three tasks we examined the possibilities of bilingual advantages in interpreters at the level of interference suppression, prepotent response inhibition, attention, STM, and updating. We also tested whether interpreters have advantages at the level of the Hebb repetition effect, an analog of novel word learning. We not only investigated whether interpreting leads to improved cognitive control over monolinguals, but also how the cognitive implications of simultaneous interpreting may differ from more typical bilingual language use. Based on the research explained above and assuming the existence of a bilingual advantage, we predict L2 teachers and interpreters to outperform monolinguals on all cognitive control measures. If the bilingual advantage is specifically related to extensive language control, interpreters are anticipated also to outperform L2 teachers.

Materials and Methods

Participants

A total of 59 participants were recruited and divided into three groups: 19 professional interpreters, 20 L2 teachers, and 20 monolinguals. All participants reported having no language, hearing, uncorrected visual, or neurological problems. Informed consent was obtained under a protocol approved by the ethical committee at the Université catholique de Louvain, Belgium. As for Experiment 1, objective language proficiency tests could not be used given that interpreters had different languages as L1 and L2. Participants filled in the Language Experience and Proficiency Questionnaire (LEAP-Q) to obtain self-rated language proficiency (Marian et al., 2007). Further, as in Experiment 1, we administered the short untimed 12 item-version of the Advanced Progressive Matrices (Bors and Stokes, 1998) as a measure of intelligence.

Detailed demographic information is reported in Table 4. All groups were highly proficient in French. The 20 monolinguals had French as L1 and acquired anecdotal knowledge of an L2 (Dutch or English) through formal education. That is, they

indicated that they had low L2 proficiency and rarely used this language (see Table 3). The 20 L2 teachers were highly proficient bilinguals with no experience in simultaneous interpreting. They spoke French ($n = 18$) or Dutch ($n = 3$) as L1 and had at least 4 years of experience in teaching L2 courses (English or French). The 19 interpreters had at least 4 year of professional experience in simultaneous interpreting. They spoke a variety of languages as L1 (8 French, 4 Dutch, 4 English, 1 German, and 2 Spanish) and L2 (6 French, 2 Dutch, 6 English, 1 Italian, 1 Spanish, 1 German, 1 Polish, and 1 Russian).

The three groups were matched on age (substantial evidence), male/female ratio (substantial evidence), years of education (anecdotal evidence), and intelligence (substantial evidence). Planned comparisons showed anecdotal evidence that interpreters and L2 teachers were matched on L1 and L2 proficiency, $t(37) = 1.15$, $p = 0.15$, $BF_{10} = 0.73$ for L1 and $t(26.25) = -1.00$, $p = 0.34$, $BF_{10} = 0.46$ for L2. Furthermore, both L2 teachers and interpreters reported higher L1 proficiency than the monolingual group [monolinguals vs. teachers: $t(38) = -2.86$, $p < 0.01$, $BF_{10} = 6.68$ (substantial evidence); monolinguals vs. interpreters: $t(37) = -4.36$, $p < 0.001$, $BF_{10} > 100$ (decisive evidence)]. The same was true for L2 proficiency [monolinguals vs. teachers: $t(38) = -10.94$, $p < 0.001$, $BF_{10} > 100$ (decisive evidence); monolinguals vs. interpreters: $t(22.29) = -12.76$, $p < 0.001$, $BF_{10} > 100$ (decisive evidence)].

Stimuli and procedure

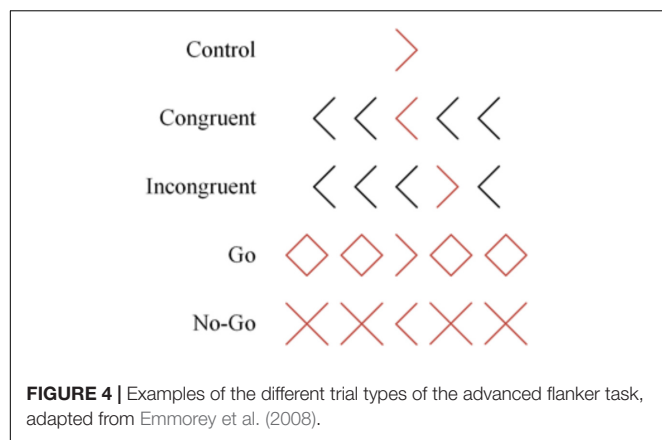
Participants were tested individually in a quiet room. They were asked to carry out the intelligence test and three computerized tasks (advanced flanker task, n -back task, Hebb repetition paradigm) in a counterbalanced order.

Advanced flanker task. The stimuli were red arrows that could be flanked by four distractors (Figure 4). There were three block types. In control blocks, participants saw single red arrows pointing to the left or right. These blocks provide a measure of attention. In flanker blocks, there was an equal number of congruent (flanking black arrows pointing in the same direction as the red target arrow) and incongruent trials (flanking black arrows pointing in the opposite direction as the red target arrow).

TABLE 4 | Demographic data of the different participant groups in Experiment 2.

	Monolinguals	L2 teachers	Interpreters	Test	BF_{10}
N	20	20	19		
Male/female ratio	4/16	4/16	6/13	$\chi^2 < 1$	0.14
Age (years)	44.40 (8.30)	44.15 (8.43)	48.58 (9.76)	$F(2, 56) = 1.53$	0.28
Education (years)	16.45 (2.70)	17.45 (3.78)	18.42 (1.84)	$F(2, 56) = 2.25$	0.51
Raven (raw score)	7.80 (2.02)	7.20 (2.61)	8.05 (1.96)	$F < 1$	0.15
L1 proficiency (10-point scale)	8.80 (0.58)	9.33 (0.60)	9.61 (0.59)	$F(2, 56) = 9.64^{***}$	>100
L1 age of acquisition (years)	0.75 (0.64)	0.90 (2.27)	1.74 (3.63)	$F < 1$	0.17
L1 frequency of use (%)	95.20 (6.43)	59.15 (15.70)	44.89 (14.13)	$F(2, 56) = 81.62^{***}$	>100
L2 proficiency (10-point scale)	2.25 (2.16)	8.33 (1.43)	8.68 (0.62)	$F(2, 56) = 107.71^{***}$	>100
L2 age of acquisition (years)	12.35 (2.50)	8.10 (5.17)	11.05 (6.21)	$F(2, 56) = 4.02^*$	2.02
L2 frequency of use (%)	3.80 (4.92)	29.15 (12.38)	32.95 (13.29)	$F(2, 56) = 42.40^{***}$	>100

SD are shown between parentheses. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.



On incongruent trials, participants had to inhibit interference of the flanking arrows. The difference in performance between congruent and incongruent trials (i.e., flanker congruency effect) reflects a measure of interference suppression. The red arrow could be either presented in the middle or one place to the left or right of the middle position. This was done to prevent participants from focusing solely on the middle stimulus. Finally, in go/no-go blocks, there were an equal proportion of go and no-go trials. On go trials, a central red arrow was flanked by four red diamonds, two on each side. Participants had to indicate the direction of the red arrow as fast as possible. On no-go trials, the arrow was flanked by four red Xs and participants were required to withhold their responses. In this go/no-go block, participants were required to inhibit their responses on no-go trials while responding as rapidly as possible on go trials. The difference in performance between go and no-go trials (i.e., go/no-go congruency effect) provides a measure of prepotent response inhibition.

The task was programmed using Tscope (Stevens et al., 2006). Participants were asked to indicate the direction of the red arrow by pressing the left (d) or right (k) button on a keyboard. Each trial began with a centered 250 ms fixation cross, followed by the stimulus for 2000 ms or until a response was made. There was an inter-trial interval of 500 ms. Each block type was presented twice. Control blocks were always presented as the first and last blocks, with flanker and go/no-go blocks alternating between them in a counterbalanced order. Each block began with 12 practice trials with feedback, followed by 48 trials. Trial types were randomized within each block.

Hebb repetition paradigm. The materials and procedure were based on the study of Szmalec et al. (2009). Sequences of

nine syllables with a consonant-vowel structure were presented visually to the participants for immediate serial recall. Two sets (A and B) of nine syllables that were matched on bigram frequency (in French) were generated using WordGen (see Table 5; Duyck et al., 2004). For half of the participants, set A was used for filler sequences and set B for the Hebb sequence. For the other half it was the reverse. Overall task performance was taken as a measure of STM. The Hebb repetition effect (i.e., the different performance for Hebb and filler trials) provides a measure of long-term memory sequence learning that has been shown to underlie novel word-form acquisition.

The task was developed in E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA, United States). Syllables were presented sequentially for 1000 ms. There was an inter-syllable interval of 500 ms. After the presentation of the sequence, a recall screen was presented on which all syllables were randomly positioned in a circle around a central question mark. Participants were instructed to click with the computer mouse on the syllables in the same order in which they were presented. Participants could click the question mark to indicate an omission, at the position in the sequence of the forgotten syllable. This way, correct responses after an omission are still in the right serial position. When participants clicked nine times (on syllables or the question mark), they were asked to press the space bar to start the following trial. The task started with two practice trials. Participants always saw two consecutive filler sequences, followed by the Hebb sequence. The experiment ended when the participant correctly reproduced two successive Hebb trials, with a maximum of 20 repetitions.

N-back task. A 2-back version was used. Participants saw a long sequence of items and were asked to indicate for each individually presented item whether it was the same as the one that was presented 2 positions before (an example of a match is t-d-m-d; a mismatch is t-h-m-d). Participants were thus required to remember the 2 most recently presented items in their correct serial order. This implied that they had to update the memorized sequence of the 2 most recent items after each trial. On lure trials, a word did not match the word that was presented 2 items before, but one of its neighboring items (an example of an $n + 1$ lure d-h-m-d; an $n - 1$ lure is t-h-d-d). Lure trials typically lead to slower responses and reduced accuracy (McElree, 2001; Gray et al., 2003; Oberauer, 2005; Jonides and Nee, 2006; Kane et al., 2007; Szmalec et al., 2011). This is because continuously updating items in STM hinders distinguishing between relevant and irrelevant items. Although the entire task is an updating task, lure interference effects (i.e., the difference between mismatch and lure trials) were taken as a measure of updating abilities,

TABLE 5 | Syllables and their French bigram frequency used in the Hebb repetition paradigm.

	1	2	3	4	5	6	7	8	9
Set A	VE	DA	FI	GU	JO	ZI	WA	XA	RO
	1416	892	1153	889	253	99	36	104	3642
Set B	CO	CU	MI	BI	PE	JL	MU	PO	XU
	3821	957	1885	1202	1537	8	477	1833	44

because recollection demands are most strongly involved in lure trials. On these trials, participants must make a clear distinction between the current trial (requiring a negative response) and the previous trial (which would lead to a positive response). If updating is not efficient, this should lead to larger lure interference effects.

The procedure and materials were held as close as possible to Szmalec et al. (2011). Participants were asked to indicate as quickly as possible whether or not the presented consonant on the screen matched the item that was presented 2 consonants earlier, by pressing the right (k) or left (d) button on a keyboard, respectively.

The task was developed in E-Prime 2.0. Each trial started with the presentation of a 500 ms consonant, followed by a 2500 ms fixation cross. The task consisted of 20 practice trials that did not contain lure trials, followed by four randomly presented blocks of $45 + 2$ (two stimuli that did not require a response at the beginning of each list of consonants) trials. Each block contained 15 match trials, 24 mismatch trials, 3 $n-1$ lure trials, and 3 $n+1$ lure trials that were presented in a pseudo-random order.

Results

The same data-analyses procedures as in Experiment 1 were used.

Advanced flanker task

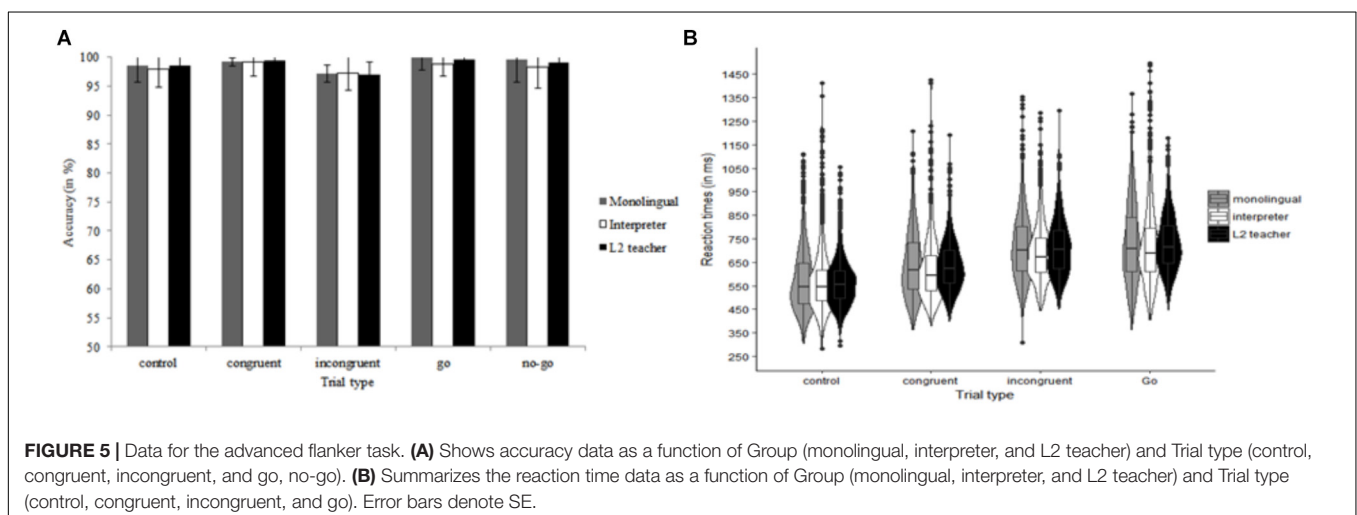
The data of one L2 teacher was excluded because he had an average ACC of only 50% (chance-level) for congruent trials. The ACC data are shown in **Figure 5A**. The final model on ACC for the control block contained Group (interpreter, L2 teacher, monolingual) as fixed effect and Participant as random effect. The model on ACC for the Go/no-go block contained Group (interpreter, L2 teacher, and monolingual) and Trial type (go and no-go) and their interaction as fixed effects, Participant as random effect and by-participant random slopes of Trial type. The final model on ACC for the analyses for the flanker block contained Group (interpreter, L2 teacher, and monolingual) and Congruency (congruent and incongruent) as fixed effects, Participant as random effects and by-Participant random slopes of Congruency. For the control block, we observed decisive

evidence against a main effect of Group, $\chi^2 < 1$, $BF_{10} < 0.01$. For the Go/no-go block, there was very strong evidence against a main effect of Group, $\chi^2(2) = 8.08$, $p = 0.02$, $BF_{10} = 0.01$, and substantial evidence against an effect of Trial type, $\chi^2(1) = 5.59$, $p = 0.02$, $BF_{10} = 0.22$. There was decisive evidence against an interaction of Group and Trial type, $\chi^2 < 1$, $BF_{10} < 0.01$. For the flanker block, we observed decisive evidence against an effect of Group, $\chi^2 < 1$, $BF_{10} < 0.01$. There was decisive evidence in favor of an effect of Congruency, $\chi^2(1) = 37.55$, $p < 0.001$, $BF_{10} < 0.01$ (i.e., flanker congruency effect). There was decisive evidence against an interaction of Group and Congruency, $\chi^2 < 1$, $BF_{10} < 0.01$.

2.26% of the RT data (310 trials) were outliers. A univariate ANOVA indicated that there were no differences between the number of trials excluded for the interpreters ($n = 105$), the monolinguals ($n = 108$), and the L2 teachers ($n = 97$), $F < 1$. The trimmed RT data are summarized in **Figure 5B**. The final model on RT for the control and Go/no-go block contained Group (interpreter, L2 teacher, and monolingual) as fixed effect and Participant as random effect. The final model on RT for the analyses for the flanker block contained Group (interpreter, L2 teacher, and monolingual) and Congruency (congruent and incongruent) as fixed effects, Participant as random effects and by-Participant random slopes of Congruency. For both the control and go block, we observed decisive evidence against a main effect of Group, $\chi^2 < 1$, $BF_{10} < 0.01$. For the flanker block, there was decisive evidence for a main effect of Trial type, $\chi^2(1) = 61.47$, $p < 0.001$, $BF_{10} > 100$ (i.e., flanker congruency effect). There was decisive evidence against a main effect of Group, as well as against an interaction of Congruency and Group, $\chi^2 < 1$, $BF_{10} < 0.01$ for both effects.

Hebb repetition paradigm

Hebb recall performance was calculated with the McKelvie scoring method (McKelvie, 1987). This method takes into account both the position and serial order of recalled items. First, we counted the number of items that were in the correct position from left to right up to the first error. Second, the same step



was repeated from right to left up to the first error. Third, the number of items in any correct sequence of two or more items between the first error from the left and the first error from the right was counted. Finally, any other items that occurred in the correct serial position were counted. The maximal possible score for each sequence was 9. The mean McKelvie score for each Group and each Trial type are presented in **Figure 6**. Analyses were performed at the mean level, because not all participants had the same number of trials due to the stopping criterion. The model on the McKelvie scores included Group (monolingual, L2 teacher, and professional interpreter), Trial type (filler and Hebb) and their interaction as fixed effects, and Participant as random effect. We observed decisive evidence for an effect of Trial type (i.e., Hebb repetition effect), $\chi^2(1) = 82.82$, $p < 0.001$, $BF_{10} > 100$, but strong evidence against a main effect of Group, $\chi^2(1) = 2.24$, $p = 0.33$, $BF_{10} = 0.03$. There was also strong evidence against an interaction of Trial type and Group, $\chi^2(2) = 3.28$, $p = 0.19$, $BF_{10} = 0.04$. On average, the monolingual group needed 14.50 repetitions ($SD = 5.72$) to reach the stopping criterion, the L2 teachers 12.60 repetitions ($SD = 6.89$) and the interpreters 14.63 repetitions ($SD = 5.89$). A univariate ANOVA on the number of repetitions showed that there was substantial evidence against a group difference, $F < 1$, $BF_{10} = 0.13$.

N-back task

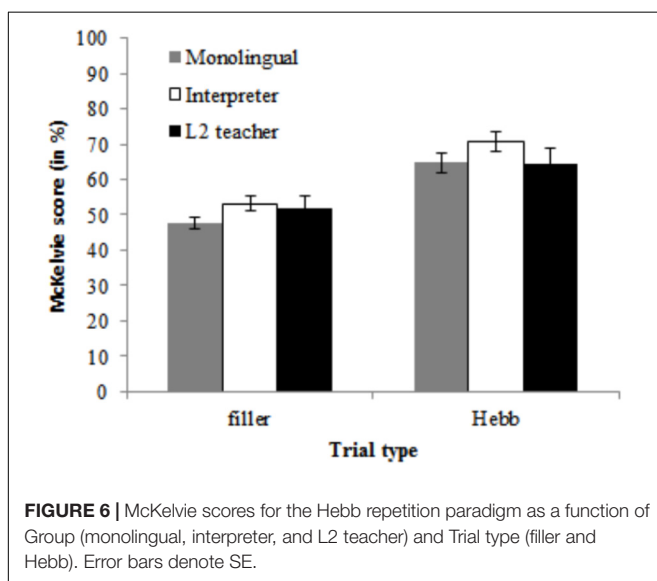
The data of two monolinguals, one L2 teacher, and two professional interpreters were excluded because they had an ACC below 50% (i.e., chance-level) on match trials. The final sample contained 18 monolinguals, 19 L2 teachers, and 17 interpreters. The ACC data are presented in **Figure 7A**. For ACC, the model included Group (interpreter, L2 teacher, and monolingual), Trial type (mismatch, match, $n + 1$ lure, and $n - 1$ lure) and their interaction as fixed effects, Participant and Trial order as random effects and by-Participant random slopes of Trial type. Trial order was included to control for learning effects, as we presented trials in a counterbalanced order (Baayen et al., 2008). We observed

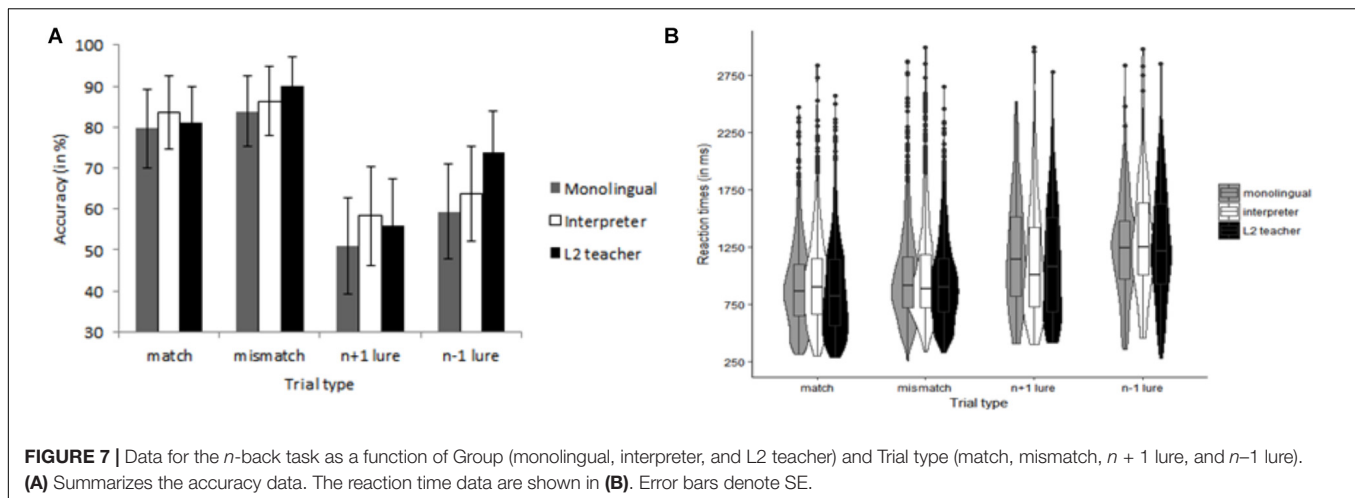
decisive evidence for an effect of Trial type, $\chi^2(3) = 76.36$, $p < 0.001$, $BF_{10} > 100$. There was decisive evidence against an effect of Group, $\chi^2(2) = 1.69$, $p = 0.43$, $BF_{10} < 0.01$, and against an interaction of Trial type and Group, $\chi^2(6) = 4.08$, $p = 0.67$, $BF_{10} < 0.01$. Planned comparisons on Trial type revealed decisive evidence for an $n + 1$ lure effect (mismatch vs. $n + 1$ lures), $t = -19.46$, $p < 0.001$, $BF_{10} > 100$, and for an $n - 1$ lure effect (mismatch vs. $n - 1$ lures), $t = -13.43$, $p < 0.001$, $BF_{10} > 100$.

The RT data are summarized in **Figure 7B**. Here, 2.42% of the RTs (192 trials) were outliers. There were no differences between the number of trials excluded for the interpreters ($n = 52$), the monolinguals ($n = 66$), and the L2 teachers ($n = 74$), $F(51) = 1.77$, $p = 0.18$. The same model as for ACC was used for analyses on RTs. We observed decisive evidence for an effect of Trial type, $\chi^2(3) = 122.93$, $p < 0.001$, $BF_{10} > 100$. There was decisive evidence against an effect of Group, $\chi^2 < 1$, $BF_{10} < 0.01$, and against an interaction of Trial type and Group, $\chi^2(6) = 1.65$, $p = 0.95$, $BF_{10} < 0.01$. Planned comparisons on Trial type revealed decisive evidence for a $n + 1$ lure effect, $t = 7.65$, $p < 0.001$, $BF_{10} > 100$, and a $n - 1$ lure effect, $t = 14.68$, $p < 0.001$, $BF_{10} > 100$.

Summary of Results

The aim of Experiment 2 was to investigate whether the high levels of language control of professional interpreters amplify possible cognitive control advantages often associated with bilingualism. Therefore, we compared three participant groups (professional interpreters, L2 teachers, and monolinguals) on a wide range of cognitive control measures, including interference suppression, response inhibition, attention, STM, and updating. Overall, we did not find support for a bilingual or interpreter advantage. First, our results on the advanced flanker task revealed evidence for similar flanker congruency effects for the three groups. The results on this task also showed that there were no differences between the three groups in the terms of the go/no-go congruency effect or on the control block. Together, these results suggest that all groups had similar performance in terms of interference suppression, prepotent response inhibition, and attention, respectively. Second, the results on the Hebb repetition paradigm also provided strong evidence against group differences. There was no evidence for an overall better performance for L2 teachers or interpreters relative to monolinguals. The interpreters also performed similarly to the L2 teachers. This indicates that all groups had similar STM. Furthermore, the comparable Hebb repetition effect for the three groups suggests that there were no differences in terms of long-term memory sequence learning that underlies novel word-form learning. Finally, there were no differences between the three groups on lure interference in the n -back task, indicating similar updating abilities. Thus, we found no evidence for an interpreter advantage on any tested cognitive control aspect. Indeed, our data showed that the interpreters and the L2 teachers performed very similarly on conflict resolution (interference suppression and prepotent response inhibition), attention, STM, and updating. Furthermore, both bilingual groups did not differ from the monolinguals in terms of their cognitive control performance, indicating that there was no measurable bilingual advantage.





It is worth mentioning that the lack of evidence for a bilingual advantage in Experiment 2 was accompanied in each task by decisive evidence in favor of the expected markers of cognitive control. As such, our participants had clear flanker congruency effects in the advanced flanker task, lure effects in the *n*-back paradigm, and clear Hebb repetition effects. This shows that the tasks used in the current study were valid and sensitive to the underlying cognitive control processes that they were meant to measure.

Cross-Experiment Comparison

We performed additional analyses to further explore the reliability of our null-findings. Although we obtained similar results in two independent experiments, which strengthens the reliability of our results, there may still be smaller bilingual or interpreter advantages that we were not able to detect. If such small group differences exist, we might detect them by combining the data of Experiments 1 and 2. To this end, we calculated standardized *z* scores for the accuracy data of interpreters and monolinguals for the measures of interference suppression (flanker congruency effect in both experiments), prepotent response inhibition (Simon effect for Experiment 1, go/no go congruency effect for Experiment 2), and STM (digit span task performance for Experiment 1, overall performance on the Hebb task for Experiment 2) for each Experiment. The *z* scores for the ACC data are shown in **Figure 8A**. We also calculated the *z* scores for the reaction time data of interpreters and monolinguals for the measures of interference suppression (flanker congruency effect in both experiments), and prepotent response inhibition (Simon effect for Experiment 2, go RTs for Experiment 2) for each Experiment. The *z* scores for the RT data are summarized in **Figure 8B**.

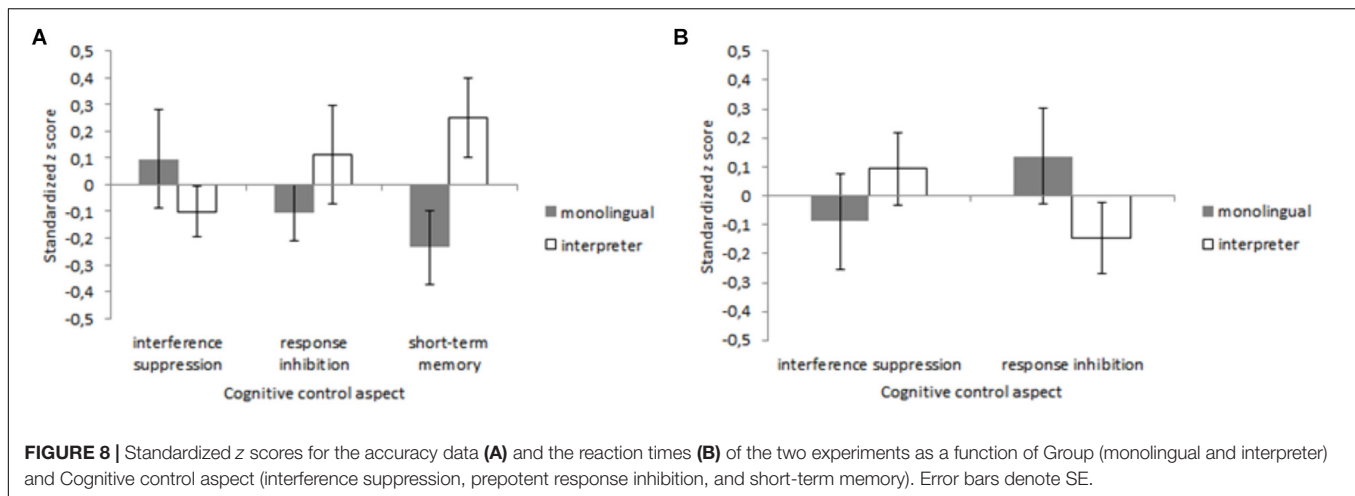
For both the ACC and RT data on interference suppression, independent samples *t*-tests comparing 45 interpreters and 46 monolinguals revealed no evidence in favor of any group differences, $t < 1$, $BF_{10} = 0.33$, and $t < 1$, $BF_{10} = 0.31$, respectively. The same was true for prepotent response inhibition. That is, independent samples *t*-tests comparing 44 interpreters and 47 monolinguals on both the composite score for ACC and

RT revealed no evidence in favor of any group differences, $t(89) = -1.04$, $p = 0.30$, $BF_{10} = 0.35$, and $t < 1$, $BF_{10} = 0.22$, respectively. In contrast, comparing the STM data of 44 interpreters and 47 monolinguals, we observed anecdotal evidence for better STM for interpreters than for monolinguals, $t(89) = 2.40$, $p = 0.02$, $BF_{10} = 2.65$. These results suggest that, although only to a small degree, experience in simultaneous interpreting may to some extent be associated with better STM performance.

DISCUSSION

The main purpose of this study was to investigate whether the high levels of language control of interpreters amplify possible cognitive control advantages often associated with bilingualism. We therefore conducted two experiments in which we compared interpreters to other populations (monolinguals and L2 teachers) on a wide range of cognitive control measures, including conflict resolution, attention, STM, and updating. Based on the adaptive control hypothesis (Green and Abutalebi, 2013), we predicted that the two bilingual groups would outperform the monolingual group on all cognitive control measures. Furthermore, we anticipated that the interpreters would outperform the L2 teachers because the higher language control demands associated with simultaneous interpreting could amplify the bilingual advantage.

In Experiment 1, we used the flanker, Simon, and digit span tasks to compare professional interpreters and monolinguals on interference suppression, prepotent response inhibition, and STM, respectively. We did not find evidence for any cognitive control advantage for interpreters over monolinguals. In Experiment 2, we compared the performance of professional interpreters, L2 teachers and monolinguals on interference suppression, prepotent response inhibition, attention, STM, and updating. We therefore used an advanced flanker task, an *n*-back task and a Hebb repetition paradigm. Again, we did not observe support for a bilingual or interpreter advantage on any of the measures. The combined results of Experiment 1 and 2 indicate



that the interpreters performed like the monolinguals and the L2 teachers on all the tested cognitive control processes. This suggests that there is no bilingual advantage in cognitive control, at least not for L2 teachers and interpreters. To further examine this result, we conducted an additional set of analyses. By merging the data of both experiments by analyzing the standardized composite scores in a cross-experiment comparison, we found further support that experience in simultaneous interpreting does not lead to an advantage in conflict resolution, neither at the level of interference suppression, nor at the level of prepotent response inhibition. The cross-experiment analyses did, however, reveal a small but measurable advantage for interpreters over monolinguals in terms of STM. Given that we had no group of L2 teachers in Experiment 1, we were not able to test whether the STM advantage was related to bilingualism or specifically to simultaneous interpreting. In sum, the combined results of Experiment 1 and 2 that the interpreters performed like the monolinguals and the L2 teachers suggests that there is no bilingual or interpreter advantage at the level of conflict resolution, attention, and updating. The results provide, on the other hand, anecdotal evidence for a (small) bilingual advantage in STM.

The fact that we have not found empirical support for the existence of an advantage for our bilingual groups on most of the tested cognitive control processes (conflict resolution, attention, and updating) is noteworthy. We examined highly proficient interpreters who all had at least 4 years of professional experience and may therefore be assumed to have extensive training in language control. Furthermore, we also recruited highly proficient L2 teachers who were using their languages daily for their professional activities. Our results therefore suggest that neither using languages in a dual-language context, nor having extensive training in language control is a guarantee to develop overall superior cognitive control abilities. The current findings are in line with previous studies that failed to obtain evidence for better performance for bilinguals than for monolinguals on multiple aspects of cognitive control (Paap and Greenberg, 2013; Paap et al., 2015) or for professional interpreters in particular (e.g., Liu et al., 2004;

Christoffels et al., 2006; Yudes et al., 2011; Babcock and Vallesi, 2017).

There are several possible explanations for the lack of enhanced conflict resolution, attention, and updating for interpreters relative to other groups. First, as already noted, interpreters may not use their language control mechanisms as other, more typical bilinguals. Interpreters arguably experience more cross-language interference between their languages and a greater requirement to produce the correct output in the target language. Given these extreme language control demands, interpreters might develop qualitative different methods to manage their languages and to be able to comprehend and produce information in different languages simultaneously. Consequently, interpreters might not develop better cognitive control, because they are not using the same language control mechanisms as other bilinguals. Nevertheless, we also could not observe an advantage for L2 teachers over monolinguals, suggesting that even for more typical bilinguals there might not be a bilingual advantage. This brings us to the second possibility, which is that there is no bilingual advantage at the level of conflict resolution, attention, updating, and novel word learning. In a recent study, Van de Putte et al. (2018) also did not find support for the hypothesis that interpreting experience enhances cognitive control. In their study, interpreters and translators performed similarly on tasks measuring conflict resolution and shifting, both before and after a 9-month training in their profession. However, only after the language training, the authors observed increased activation for the interpreters relative to the translators in the right angular gyrus during the shifting task and in the left superior temporal gyrus during the conflict resolution task. As neural measures were outside the scope of the current study, future work should shed light on the relationship between simultaneous interpreting training, the associated neural changes, and their relation to behavioral cognitive control measures. Third, it should be mentioned that the monolinguals tested in the current study also acquired (passive and anecdotal) knowledge of a L2. It cannot be excluded that the interpreters and L2 teachers tested here improved their cognitive control abilities, but that the improvement is not

linearly related to L2 proficiency or that they reached a ceiling. That is, the dual-language use and higher demands of language control might not further increase the cognitive control benefits that all participants already had due to the fact that they all knew a L2. A fourth and final possible explanation for the absence of a bilingual or interpreter advantage on the aforementioned aspects of cognitive control is that our bilingual groups were too proficient. Paap (2018) proposed the Controlled Dose hypothesis, which states that the bilingual advantage might only be present during the process of L2 acquisition. This hypothesis is based on a general framework of behavioral learning proposed by Chein and Schneider (2012). The acquisition of novel behavior typically proceeds with shifting from relying on the metacognitive system during the formation stage, to recruiting the cognitive control network during the controlled-execution stage and, finally, to relying on the representation system during the automatic-execution stage (see **Figure 9**). According to the Controlled Dose hypothesis (Paap, 2018), there might be a similar shift in engagement of cognitive control for bilinguals. The bilingual advantage may therefore only be present during a particular period of L2 acquisition, when bilinguals are still learning how to juggle their languages. Once bilinguals have sufficient training in language control, language management might become an automatic skill that does not require cognitive control processes. Similar to losing better developed muscles when stopping physical fitness training, the benefits in cognitive control of bilinguals might not persist indelibly when this mechanism is no longer recruited for language control. This hypothesis is new and still needs to be investigated. According to the Controlled Dose hypothesis, a benefit in cognitive control might thus be predicted for interpreters and L2 teachers, but these advantages are likely to be transitory. Less experienced interpreters and L2 teachers who are still becoming more proficient in their job may still train their cognitive control with every linguistic choice they make, so that there can be a cognitive control advantage for these populations.

Prior research has indeed found bilingual advantages for interpreters that were still at the start of their professional career. Woumans et al. (2015), for instance, compared the conflict resolution performance of student interpreters, student balanced bilinguals, student unbalanced bilinguals, and student

monolinguals. They used the Simon task to measure prepotent response inhibition and the ANT to measure interference suppression. All bilingual groups had a smaller Simon effect than the monolinguals, suggesting better prepotent response inhibition. Furthermore, both the interpreters and balanced bilinguals had a smaller congruency effect on the ANT, indicating superior interference suppression. It is possible that an advantage was found in the study of Woumans et al. (2015), but not in the current study, because of the fact that the student interpreters and student balanced bilinguals in the study of Woumans et al. (2015), were still gaining L2 proficiency, whereas the bilingual groups tested here were not. This idea also fits with the Bilingual Expertise hypothesis (Incera and McLennan, 2016; Damian et al., 2018). It has been found that proficient bilinguals take longer to start moving the mouse in a mouse tracking paradigm but then move more efficiently than monolinguals to the correct response. However, no group differences emerge in terms of RTs. It is thus possible that bilinguals change the way in which they approach cognitive control tasks once they have sufficient training in language control, although this does not imply better performance. Our study design does not permit to draw any firm conclusions about the Controlled Dose hypothesis, but together with the findings of prior research it shows that it deserves further investigation. The bilingual profile of the participants in studies of this type should be controlled carefully in the future, as it would enable us to understand when bilingualism provides an advantage in cognitive control and why.

Regardless of the explanation, the results of the current study indicate that neither high levels of L2 proficiency and use in a dual-language context, nor experience with simultaneous interpreting leads to measurable enhancements in conflict resolution, attention, and updating. We did, on the other hand, find some evidence for improved STM for interpreters when compared to monolinguals. Although our findings should be interpreted with caution given the anecdotal evidence in favor of its existence, this interpreter advantage at the level of STM is in line with previous studies which have shown that simultaneous interpreters have better STM

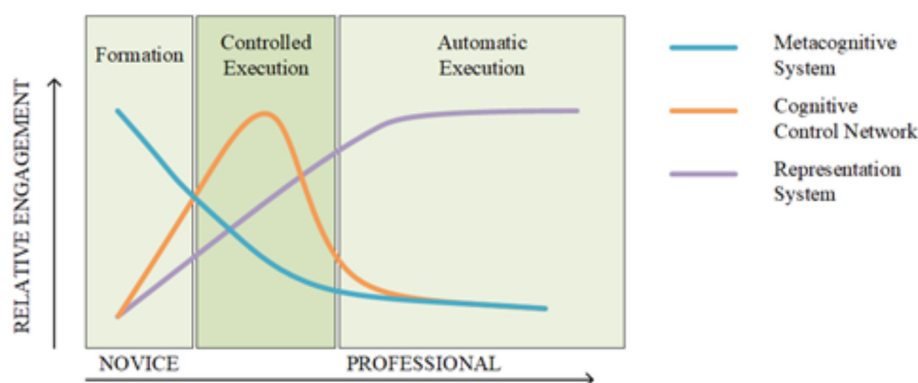


FIGURE 9 | The framework of behavioral learning of Chein and Schneider (2012).

when compared to other populations (Bajo et al., 2000; Padilla et al., 2005; Christoffels et al., 2006; Signorelli et al., 2011; Yudes et al., 2011; Stavrakaki et al., 2012; Babcock and Vallesi, 2017). Christoffels et al. (2006), for instance, examined whether professional interpreters had better STM than bilingual university students and highly proficient L2 teachers. They recruited 13 interpreters, 39 bilingual students, and 15 L2 teachers. Using a word span task that was highly comparable to the digit span task used in Experiment 2, they observed that interpreters outperformed both the students and the L2 teachers, while the students and the L2 teachers performed similarly. The authors also found that interpreters, bilingual students and L2 teachers performed similarly on a basic reaction time task, measuring attention. The authors therefore concluded that working memory is a crucial cognitive control aspect for simultaneous interpreting, whereas attention is not. In the current study, we were not able to find evidence for an interpreter advantage on the digit span task, despite the fact that we used a highly similar design as in the study of Christoffels and colleagues and that we tested groups of comparable size. Nevertheless, the results of the cross-experiment comparison did reveal (small) evidence for an interpreter advantage, in line with the findings of Christoffels et al. (2006). The fact that the advantage in working memory was rather small is further in line with a recent meta-analysis of Grundy and Timmer (2016). In their study, they analyzed the advantage in STM for bilinguals over monolinguals combining data from 88 effect sizes, 27 independent studies, and 2901 participants. Their results revealed a small to medium effect in favor of a bilingual advantage in working memory. So, it appears that bilingualism can give an advantage in working memory, but this advantage is rather small and may therefore be difficult to detect. Across all other cognitive control measures, on the other hand, we found no evidence in favor of an interpreter or bilingual advantage. Together, the current findings therefore further corroborate to the idea that simultaneous interpreting may lead to enhanced STM relative to monolinguals, albeit that this advantage is rather small. An advantage in STM for interpreters over other populations is reasonable given the nature and demands of simultaneous interpreting. Working memory is a crucial component because interpreters have to store content in a source language and reformulate this content in the target language while articulating previous reformulated messages. This high working memory demand appears to alter STM capacity.

The present results should be regarded with a degree of caution, as there are certain limitations worth noting. First, the monolingual group in Experiment 2 had a lower L1 proficiency than the two bilingual groups. This contrasts with prior work, which found that bilinguals have reduced vocabulary knowledge in their L1 relative to monolinguals (Bialystok et al., 2009). The higher L1 proficiency of the bilinguals tested here is likely due to the fact that both professional interpreters and L2 teachers are linguists that received formal education in their L1, which was not true for the monolinguals. This could have influenced the results for our measure of novel word learning. Nevertheless, even with this advantage in L1

proficiency, there were still no differences between the bilinguals and the monolinguals in the Hebb repetition paradigm. Second, in Experiment 2, we compared the performance of interpreters, L2 teachers, and monolinguals on tasks that were selected because they appeal on particular cognitive control processes (conflict resolution, attention, updating, working memory, long-term memory consolidation that underlies word learning). The choice of our tasks raises some questions. First, overall accuracy on the conflict resolution tasks was high. Bialystok (2015) argued that the bilingual advantage is more likely to emerge in more effortful tasks. Although accuracy rates are comparable to past research that did obtain evidence for a bilingual advantage (e.g., Emmorey et al., 2008), it cannot be excluded that the tasks were not sufficiently effortful to detect differences in conflict resolution between our groups. Furthermore, with respect to the *n*-back task, it would also be interesting to examine whether memory updating is better for interpreters or bilinguals in general if words are used instead of consonants. Remembering and updating words is more naturalistic and is more in line with the professional activities of interpreters than consonants. Finally, we decided to use a visual version of the Hebb repetition paradigm. Given the nature of simultaneous interpreting, it would be interesting to examine in future work whether an oral version of the Hebb repetition paradigm, in which the sequences are not presented visually but auditory, elicits better performance for interpreters. During their profession, interpreters hear incoming information which they have to transform and store in their memory. Previous research that reported an interpreter advantage mainly used oral working memory tasks (e.g., Padilla et al., 2005; Christoffels et al., 2006; Signorelli et al., 2011). Nevertheless, the digit span task in Experiment 1 was an oral STM task, where we also failed to observe strong evidence for an interpreter advantage. It should be noted, though, that the testing language of this task was different for monolinguals and interpreters. Although Dutch and French digit names are very similar in terms of word length, it cannot be ruled out that cross-language variability in digit span performance masked possible group differences. Nevertheless, although simultaneous interpreting is likely to train specifically oral STM, the current study suggests that this advantage is rather small.

To conclude, the results of the current study once more point toward the complexity of the phenomenon of bilingualism and the difficulty to determine its cognitive implications. Prior work has suggested that particular characteristics of bilingualism might be important for the advantage to emerge. The amount of language control needed in daily life has been proposed as being the modulating factor. The results of this study provide further insights in this matter by showing that extensive training in language control does not necessarily always lead to general beneficial effects on cognitive control. Although we found ambiguous evidence that interpreters have better STM than monolinguals, there was no evidence for an advantage at the level of conflict resolution, attention, updating, and novel word learning. Further research is needed to determine whether

there might be a certain period during language control training in which cognitive control is (overall) enhanced. Comparing the bilingual advantage between novice and professional interpreters in a longitudinal design could shed more light on the (temporary) importance of cognitive control in the bilingual brain.

DATA AVAILABILITY STATEMENT

The original datasets and R code of the different tasks for this study can be found at Mendeley data (doi: 10.17632/jcr9yswps8.3).

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the ethical committee of Ghent University (Belgium) and the ethical committee of the Psychological Sciences Research Institute at the Université catholique de Louvain (Belgium), with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

REFERENCES

- Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. doi: 10.1016/j.jml.2007.12.005
- Babcock, L., and Vallesi, A. (2017). Are simultaneous interpreters expert bilinguals, unique bilinguals, or both? *Biling. Lang. Cogn.* 20, 403–417. doi: 10.1017/S1366728915000735
- Bajo, M. T., Padilla, F., and Padilla, P. (2000). “Comprehension processes in simultaneous interpreting,” in *Translation in Context*, eds A. Chesterman, N. Gallardo San Salvador, and Y. Gambier (Amsterdam: John Benjamins), 127–142.
- Barac, R., Bialystok, E., Castro, D. C., and Sanchez, M. (2014). The cognitive development of young dual language learners: a critical review. *Early Child. Res. Q.* 29, 699–714. doi: 10.1016/j.ecresq.2014.02.003
- Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01
- Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Can. J. Exp. Psychol.* 60, 68–79. doi: 10.1037/cjep.2006008
- Bialystok, E. (2015). Bilingualism and the development of executive function: the role of attention. *Child Dev. Perspect.* 9, 117–121. doi: 10.1111/cdep.12116
- Bialystok, E., Craik, F. I., and Ruocco, A. C. (2006). Dual-modality monitoring in a classification task: the effects of bilingualism and ageing. *Q. J. Exp. Psychol.* 59, 1968–1983. doi: 10.1080/17470210500482955
- Bialystok, E., Craik, F. I. M., Green, D. W., and Gollan, T. H. (2009). Bilingual minds. *Psychol. Sci. Public Interest Suppl.* 10, 89–129. doi: 10.1177/1529100610387084
- Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290
- Bialystok, E., Craik, F. I. M., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 859–873. doi: 10.1037/0278-7393.34.4.859
- Bialystok, E., Craik, F. I. M., and Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends Cogn. Sci.* 16, 240–250. doi: 10.1016/j.tics.2012.03.001
- Bors, D. A., and Stokes, T. L. (1998). Raven’s advanced progressive matrices: norms for first-year university students and the development of a short form. *Educ. Psychol. Meas.* 58, 382–398. doi: 10.1177/0013164498058003002

AUTHOR CONTRIBUTIONS

LVdL and EVdP contributed to the acquisition of the data. LVdL carried out the analysis and interpretation of the data, the preparation of the figures, and the writing of the manuscript. All co-authors contributed to the conception of the study, the interpretation of the data, and the content and editing of the manuscript.

FUNDING

This work was supported by the Fonds National de la Recherche Scientifique F.R.S.-FNRS (1.A.935.15F to LVdL), by the Fund for Scientific Research (G058914N to WD), and by the special research fund of Ghent University (GOA – Concerted Research Action BOF13/GOA/032 to WD).

ACKNOWLEDGMENTS

The authors would like to thank Marie Lourtie and Joëlle Coen for their help with data collection.

- Chein, J. M., and Schneider, W. (2012). The brain’s learning and control architecture. *Curr. Dir. Psychol. Sci.* 21, 78–84. doi: 10.1177/0963721411434977
- Chernov, G. V. (1994). “Message redundancy and message anticipation in simultaneous interpretation,” in *Bridging the Gap: Empirical Research in Simultaneous Interpretation*, eds S. Lambert and B. Moser-Mercer (Amsterdam: John Benjamins), 139–153.
- Christoffels, I. K., de Groot, A. M. B., and Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: the role of expertise and language proficiency. *J. Mem. Lang.* 54, 324–345. doi: 10.1016/j.jml.2005.12.004
- Collette, F., Van Der Linden, M., Delfiore, G., Degueldre, C., Luxen, A., and Salmon, E. (2001). The functional anatomy of inhibition processes investigated with the Hayling task. *Neuroimage* 14, 258–267. doi: 10.1006/nimg.2001.0846
- Colomé, A. (2001). Lexical activation in bilinguals’ speech production: language-specific or language-independent? *J. Mem. Lang.* 45, 721–736. doi: 10.1006/jmla.2001.2793
- Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86. doi: 10.1016/j.cognition.2006.12.013
- Cowan, N. (2005). *Working Memory Capacity*. New York, NY: Psychology Press.
- Damian, M. F., Ye, W., Oh, M., and Yang, S. (2018). Bilinguals as “experts”? Comparing performance of mono- to bilingual individuals via a mousetracking paradigm. *Biling. Lang. Cogn.* 1–18. doi: 10.1017/S1366728918000901 Available at: <https://www.cambridge.org/core/journals/bilingualism-language-and-cognition/article/bilinguals-as-experts-comparing-performance-of-mono-to-bilingual-individuals-via-a-mousetracking-paradigm/38D3B449EF2DDFBC1FE91CF7BEE0CB00>
- de Bruin, A., Treccani, B., and Della Sala, S. (2015). Cognitive advantage in bilingualism: an example of publication bias? *Psychol. Sci.* 26, 99–107. doi: 10.1177/0956797614557866
- Dienes, Z. (2011). Bayesian versus orthodox statistics: which side are you on? *Perspect. Psychol. Sci.* 6, 274–290. doi: 10.1177/1745691611406920
- Dijkstra, T., and Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Biling. Lang. Cogn.* 5, 175–197. doi: 10.1017/S1366728902003012
- Duyck, W., Desmet, T., Verbeke, L. P. C., and Brysbaert, M. (2004). WordGen: a tool for word selection and nonword generation in Dutch, English, German, and French. *Behav. Res. Methods Instrum. Comput.* 36, 488–499. doi: 10.3758/BF03195595

- Duyck, W., and Warlop, N. (2009). Translation priming between the native language and a second language new evidence from dutch-french bilinguals. *Exp. Psychol.* 56, 173–179. doi: 10.1027/1618-3169.56.3.173
- Emmorey, K., Luk, G., Pyers, J. E., and Bialystok, E. (2008). The source of enhanced cognitive control in bilinguals: evidence from bimodal bilinguals. *Psychol. Sci.* 19, 1201–1206. doi: 10.1111/j.1467-9280.2008.02224.x
- Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267
- Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *J. Exp. Psychol. Gen.* 133, 101–135. doi: 10.1037/0096-3445.133.1.101
- Gallistel, C. R. (2009). The importance of proving the null. *Psychol. Rev.* 116, 439–453. doi: 10.1037/a0015251
- Gray, J. R., Chabris, C. F., and Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nat. Neurosci.* 6, 316–322. doi: 10.1038/nn1014
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81. doi: 10.1017/S1366728998000133
- Green, D. W., and Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377
- Grosjean, F. (2010). Bilingualism, biculturalism, and deafness. *Int. J. Biling. Educ. Biling.* 13, 133–145. doi: 10.1080/13670050903474051
- Grundy, J. G., and Timmer, K. (2016). Bilingualism and working memory capacity: a comprehensive meta-analysis. *Second Lang. Res.* 33, 325–340. doi: 10.1177/0267658316678286
- Hartanto, A., and Yang, H. (2016). Disparate bilingual experiences modulate task-switching advantages: a diffusion-model analysis of the effects of interactional context on switch costs. *Cognition* 150, 10–19. doi: 10.1016/j.cognition.2016.01.016
- Hebb, D. (1961). “Distinctive features of learning in the higher animal,” in *Brain Mechanisms and Learning*, ed. J. F. Delafresnaye (Oxford: Blackwell), 37–46.
- Henrard, S., and Van Daele, A. (2017). Different bilingual experiences might modulate executive tasks advantages: comparative analysis between monolinguals, translators, and interpreters. *Front. Psychol.* 8:1870. doi: 10.3389/fpsyg.2017.01870
- Hilchey, M. D., and Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychon. Bull. Rev.* 18, 625–658. doi: 10.3758/s13423-011-0116-7
- Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biom. J.* 50, 346–363. doi: 10.1002/bimj.200810425
- Ibáñez, A. J., Macizo, P., and Bajo, M. T. (2010). Language access and language selection in professional translators. *Acta Psychol.* 135, 257–266. doi: 10.1016/j.actpsy.2010.07.009
- Incera, S., and McLennan, C. T. (2016). Mouse tracking reveals that bilinguals behave like experts. *Biling. Lang. Cogn.* 19, 610–620. doi: 10.1017/S1366728915000218
- Jeffreys, H. (1961). *Theory of Probability*. Oxford: Oxford University Press.
- Jonides, J., and Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience* 139, 181–193. doi: 10.1016/j.neuroscience.2005.06.042
- Kane, M. J., Conway, A. R. A., Miura, T. K., and Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: a question of construct validity. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 615–622. doi: 10.1037/0278-7393.33.3.615
- Kaushanskaya, M. (2012). Cognitive mechanisms of word learning in bilingual and monolingual adults: the role of phonological memory. *Bilingualism* 15, 470–489. doi: 10.1017/S1366728911000472
- Kaushanskaya, M., and Marian, V. (2007). Bilingual language processing and interference in bilinguals: evidence from eye tracking and picture naming. *Lang. Learn.* 57, 119–163. doi: 10.1111/j.1467-9922.2007.00401.x
- Kaushanskaya, M., and Marian, V. (2009). The bilingual advantage in novel word learning. *Psychon. Bull. Rev.* 16, 705–710. doi: 10.3758/PBR.16.4.705
- Kolb, B., and Wishaw, I. Q. (2009). *Fundamentals of Human Neuropsychology*. New York, NY: Worth Publishers.
- Köpke, B., and Nespoulous, J.-L. (2006). Working memory performance in expert and novice interpreters. *Interpreting* 8, 1–23. doi: 10.1075/intp.8.1.02kop
- Lehtonen, M., Soveri, A., Laine, A., Jarvenpää, J., de Bruin, A., and Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychol. Bull.* 144, 394–425. doi: 10.1037/bul0000142
- Liu, M., Schallert, D. L., and Carroll, P. J. (2004). Working memory and expertise in simultaneous interpreting. *Interpreting* 6, 19–42. doi: 10.1075/intp.6.1.04liu
- Lo, S., and Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Front. Psychol.* 6:1171. doi: 10.3389/fpsyg.2015.01171
- Luk, G., and Bialystok, E. (2013). Bilingualism is not a categorical variable: interaction between language proficiency and usage. *J. Cogn. Psychol.* 25, 605–621. doi: 10.1080/20445911.2013.795574
- Luo, L., Craik, F. I. M., Moreno, S., and Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: evidence from aging. *Psychol. Aging* 28, 28–34. doi: 10.1037/a0030875
- Marian, V., Blumenfeld, H. K., and Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Res.* 50, 940–967. doi: 10.1044/1092-4388(2007)067
- McElree, B. (2001). Working memory and focal attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 817–835. doi: 10.1037/0278-7393.27.3.817
- McKelvie, S. (1987). Learning and awareness in the hebb digits task. *J. Gen. Psychol.* 114, 75–88. doi: 10.1080/00221309.1987.9711057
- Morales, J., Calvo, A., and Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *J. Exp. Child Psychol.* 114, 187–202. doi: 10.1016/j.jecp.2012.09.002
- Morales, J., Padilla, F., Gómez-Ariza, C. J., and Bajo, M. T. (2015). Simultaneous interpretation selectively influences working memory and attentional networks. *Acta Psychol.* 155, 82–91. doi: 10.1016/j.actpsy.2014.12.004
- Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x
- Nair, V. K., Biedermann, B., and Nickels, L. (2017). Understanding bilingual word learning: the role of phonotactic probability and phonological neighborhood density. *J. Speech Lang. Hear. Res.* 60, 3551–3560. doi: 10.1044/2017_JSLHR-L-15-0376
- Nuzzo, R. (2014). Scientific method: statistical errors. *Nature* 506, 150–152. doi: 10.1038/506150a
- Oberauer, K. (2005). Binding and inhibition in working memory: individual and age differences in short-term recognition. *J. Exp. Psychol. Gen.* 134, 368–387. doi: 10.1037/0096-3445.134.3.368
- Paap, K. R. (2018). “Bilingualism in cognitive science,” in *Handbook of Bilingualism*, eds A. De Houwer and L. Oretga (Cambridge: Cambridge University Press).
- Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Paap, K. R., Johnson, H. A., and Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69, 265–278. doi: 10.1016/j.cortex.2015.04.014
- Paap, K. R., Myuz, H. A., Anders, R. T., Bockelman, M. F., Mikulinsky, R., and Sawi, O. M. (2017). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *J. Cogn. Psychol.* 29, 89–112. doi: 10.1080/20445911.2016.1248436
- Paap, K. R., and Sawi, O. (2014). Bilingual advantages in executive functioning: problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. *Front. Psychol.* 5:962. doi: 10.3389/fpsyg.2014.00962
- Padilla, F., Bajo, M. T., and Macizo, P. (2005). Articulatory suppression in language interpretation: working memory capacity, dual tasking and word knowledge. *Biling. Lang. Cogn.* 8, 207–219. doi: 10.1017/S1366728905002269
- Papagno, C., and Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Q. J. Exp. Psychol. A* 48, 98–107. doi: 10.1080/14640749508401378
- Prior, A., and Gollan, T. H. (2011). Good language-switchers are good task-switchers: evidence from Spanish-English and Mandarin-English bilinguals. *J. Int. Neuropsychol. Soc.* 17, 682–691. doi: 10.1017/S1355617711000580

- Prior, A., and Macwhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism* 13, 253–262. doi: 10.1017/S1366728909990526
- Raven, J., Raven, J. C., and Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 2: The Coloured Progressive Matrices*. Oxford: Oxford Psychologists Press.
- Reingold, E. M., Charness, N., Pomplun, M., and Stampe, D. M. (2001). Visual span in expert chess players: evidence from eye movements. *Psychol. Sci.* 12, 48–55. doi: 10.1111/1467-9280.00309
- Rodriguez-Fornells, A., van der Lugt, A., Rotte, M., Britti, B., Heinze, H. J., and Munte, T. F. (2005). Second language interferes with word production in fluent bilinguals: brain potential and functional imaging evidence. *J. Cogn. Neurosci.* 17, 422–433. doi: 10.1162/0898929053279559
- Rosiers, A., Woumans, E., Duyck, W., and Eyckmans, J. (2017). Is it all in the Mind? *Investigating the Presumed Cognitive Advantage of Aspiring Interpreters*, eds P. Franz and L. Minhua. INTERPRETING.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* 16, 225–237. doi: 10.3758/PBR.16.2.225
- Schroeder, S. R., Marian, V., Shook, A., and Bartolotti, J. (2016). Bilingualism and musicianship enhance cognitive control. *Neural Plast.* 2016:4058620. doi: 10.1155/2016/4058620
- Service, E., Simola, M., Metsänheimo, O., and Maury, S. (2002). Bilingual working memory span is affected by language skill. *Eur. J. Cogn. Psychol.* 14, 383–408. doi: 10.1080/09541440143000140
- Signorelli, T. M., Haarmann, H. J., and Obler, L. K. (2011). Working memory in simultaneous interpreters: effects of task and age. *Int. J. Biling.* 16, 198–212. doi: 10.1177/1367006911403200
- Simon, J., and Wolf, J. D. (1963). Choice reaction time as a function of angular stimulus-response correspondence and age. *Ergonomics* 6, 99–105. doi: 10.1080/00140136308930679
- Smalle, E. H. M., Page, M. P. A., Duyck, W., Edwards, M., and Szmalec, A. (2017). Children retain implicitly learned phonological sequences better than adults: a longitudinal study. *Dev. Sci.* 21:e12634. doi: 10.1111/desc.12634
- Stavrakaki, S., Megari, K., Kosmidis, M. H., Apostolidou, M., and Takou, E. (2012). Working memory and verbal fluency in simultaneous interpreters. *J. Clin. Exp. Neuropsychol.* 34, 624–633. doi: 10.1080/13803395.2012.667068
- Stevens, M., Lammertyn, J., Verbruggen, F., and Vandierendonck, A. (2006). Tscope: a C library for programming cognitive experiments on the MS windows platform. *Behav. Res. Methods* 38, 280–286. doi: 10.3758/BF03192779
- Szmalec, A., Duyck, W., Vandierendonck, A., Mata, A. B., and Page, M. P. A. (2009). The Hebb repetition effect as a laboratory analogue of novel word learning. *Q. J. Exp. Psychol.* 62, 435–443. doi: 10.1080/17470210802386375
- Szmalec, A., Verbruggen, F., Vandierendonck, A., and Kemps, E. (2011). Control of interference during working memory updating. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 137–151. doi: 10.1037/a0020365
- Timarova, S., Cenková, I., Meylaerts, R., Hertog, E. S., Szmalec, A., and Duyck, W. (2014). Simultaneous interpreting and working memory capacity. *Interpreting* 16, 139–168. doi: 10.1075/intp.16.2.01tim
- Van Assche, E., Duyck, W., Hartsuiker, R. J., and Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychol. Sci.* 20, 923–927. doi: 10.1111/j.1467-9280.2009.02389.x
- Van de Putte, E., De Baene, W., García-Pentón, L., Woumans, E., Dijkgraaf, A., and Duyck, W. (2018). Anatomical and functional changes in the brain after simultaneous interpreting training: a longitudinal study. *Cortex* 99, 243–257. doi: 10.1016/j.cortex.2017.11.024
- Verreyt, N., Woumans, E., Vandelanotte, D., Szmalec, A., and Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism* 19, 181–190. doi: 10.1017/S1366728914000352
- Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychon. Bull. Rev.* 14, 779–804. doi: 10.3758/BF03194105
- Woumans, E., Ceuleers, E., Van der Linden, L., Szmalec, A., and Duyck, W. (2015). Verbal and nonverbal cognitive control in bilinguals and interpreters. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 1579–1586. doi: 10.1037/xlm0000107
- Woumans, E., and Duyck, W. (2015). The bilingual advantage debate: moving toward different methods for verifying its existence. *Cortex* 73, 356–357. doi: 10.1016/j.cortex.2015.07.012
- Yim, O., and Bialystok, E. (2012). Degree of conversational code-switching enhances verbal task switching in Cantonese–English bilinguals. *Biling. Lang. Cogn.* 15, 873–883. doi: 10.1017/S1366728912000478
- Yudes, C., Macizo, P., and Bajo, T. (2011). The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Front. Psychol.* 2:309. doi: 10.3389/fpsyg.2011.00309

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Van der Linden, Van de Putte, Woumans, Duyck and Szmalec. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Bidialectalism and Bilingualism: Exploring the Role of Language Similarity as a Link Between Linguistic Ability and Executive Control

Jessica Oswald^{1*}, Alisa Schättin², Claudia C. von Bastian³ and Alessandra S. Souza^{1,2}

¹ University Research Priority Program "Dynamics of Healthy Aging", University of Zurich, Zurich, Switzerland, ² Cognitive Psychology Unit, Department of Psychology, University of Zurich, Zurich, Switzerland, ³ Department of Psychology, University of Sheffield, Sheffield, United Kingdom

OPEN ACCESS

Edited by:

Peter Bright,
Anglia Ruskin University,
United Kingdom

Reviewed by:

Evelyne Mercure,
University College London,
United Kingdom
Steve Majerus,
Université de Liège, Belgium

*Correspondence:

Jessica Oswald
jessica.oschwald@gmail.com;
jessica.oschwald@uzh.ch

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 25 June 2018

Accepted: 28 September 2018

Published: 23 October 2018

Citation:

Oswald J, Schättin A,
von Bastian CC and Souza AS (2018)
Bidialectalism and Bilingualism:
Exploring the Role of Language
Similarity as a Link Between Linguistic
Ability and Executive Control.
Front. Psychol. 9:1997.
doi: 10.3389/fpsyg.2018.01997

The notion of bilingual advantages in executive functions (EF) is based on the assumption that the demands posed by cross-language interference serve as EF training. These training effects should be more pronounced the more cross-language interference bilinguals have to overcome when managing their two languages. In the present study, we investigated the proposed link between linguistic and EF performance using the similarity between the two languages spoken since childhood as a proxy for different levels of cross-language interference. We assessed the effect of linearly increasing language dissimilarity on linguistic and EF performance in multiple tasks in four groups of young adults (aged 18–33): German monolinguals ($n = 24$), bidialectals ($n = 25$; German and Swiss German dialect), bilinguals speaking two languages of the same Indo-European ancestry ($n = 24$; e.g., German-English), or bilinguals speaking two languages of different ancestry ($n = 24$; e.g., German-Turkish). Bayesian linear-mixed effects modeling revealed substantial evidence for a linear effect of language similarity on linguistic accuracy, with better performance for participants with more similar languages and monolinguals. However, we did not obtain evidence for the presence of a similarity effect on EF performance. Furthermore, language experience did not modulate EF performance, even when testing the effect of continuous indicators of bilingualism (e.g., age of acquisition, proficiency, daily foreign language usage). These findings question the theoretical assumption that life-long experience in managing cross-language interference serves as EF training.

Keywords: bidialectalism, bilingualism, language similarity, executive functions, linguistic processing

INTRODUCTION

Bilingualism, or the active use of two languages from an early age on, has been suggested to have both positive effects on non-linguistic and negative effects on linguistic processing (Bialystok et al., 2012). On the one hand, the increased attentional demand bilinguals face when they have to select words in the appropriate language while inhibiting their second language (L2) is assumed

to serve as lifelong training of executive control (Kroll and Bialystok, 2013) leading to better executive functions (EF) in bilinguals as compared to monolinguals. Bilingual advantages have been reported for several aspects of EF (Miyake et al., 2000): *inhibition* of prepotent responses (Bialystok et al., 2008; Salvatierra and Rosselli, 2010), *shifting* between mental sets and tasks (Prior and MacWhinney, 2010; Wiseheart et al., 2016), and *updating* and monitoring of working memory (WM) contents (Luo et al., 2013; Blom et al., 2014). On the other hand, the need to maintain more than one lexicon is assumed to come with disadvantages in lexical access (e.g., Bialystok, 2009), leading to worse linguistic performance in bilinguals as compared to monolinguals. Accordingly, it has been found that bilingual children and adults have a smaller receptive vocabulary (Bialystok et al., 2010; Bialystok and Luk, 2012), have lower scores in picture naming tasks (Gollan et al., 2005; Ivanova and Costa, 2008), and perform worse in word-fluency tasks (Gollan et al., 2002; Portocarrero et al., 2007) than monolinguals.

Whereas bilingual linguistic disadvantages are well-supported in the literature, bilingual EF advantages have been challenged by several recent replication failures (Morton and Harper, 2007; Paap and Greenberg, 2013; Antón et al., 2014; Duñabeitia et al., 2014; Gathercole et al., 2014; Kirk et al., 2014; Kousaie et al., 2014; Paap and Liu, 2014; de Bruin et al., 2015; Paap et al., 2016, 2014; von Bastian et al., 2016). This has led to a discussion of the variables potentially modulating the observation of bilingual EF advantages (Kroll and Bialystok, 2013; Baum and Titone, 2014; Valian, 2015). Recently, researchers have paid increasingly more attention to the multifaceted aspects of the bilingual experience, such as the age of L2 acquisition, language proficiency, and frequency of language use. Although these variables have been shown to modulate the performance of bilinguals in linguistic tasks (Portocarrero et al., 2007; Gollan et al., 2008; Luo et al., 2010; Blumenfeld et al., 2016), the importance of these factors in explaining bilingual EF advantages is still under debate. Two large-scale studies failed to observe a relation between age of L2 acquisition, language proficiency, language usage, and number of learned languages in multiple indicators of EF (Paap, 2014; von Bastian et al., 2016), and other studies did not observe a relation between age of acquisition and inhibitory control (Linck et al., 2008; Pelham and Abrams, 2014). A couple of studies have, however, reported effects of balance of language usage on inhibitory control and on shifting (Woumans et al., 2015; Yow and Li, 2015; Verreyt et al., 2016), and of L2 proficiency on conflict monitoring (Singh and Mishra, 2013). A further aspect of the bilingual experience that has received less attention in the literature is the similarity of the two languages spoken.

The Role of Language Similarity

Evidence from event-related potential studies suggests that, in bilinguals, both languages are constantly activated even if only one of them is currently in use (Kroll et al., 2012). An explanation for this parallel activation is proposed by the BIA+ model, which suggests that bilinguals have a shared mental lexicon for both languages. Consequently, when recognizing a word, lexical representations that share orthographic, phonologic and/or semantic similarity with the target word are automatically

activated regardless of the language they correspond to (Dijkstra and van Heuven, 2002). This non-selective activation is assumed to demand general executive control mechanisms to manage cross-linguistic activation (Coderre and van Heuven, 2014). Furthermore, this parallel activation leads to bidirectional cross-language interactions, such that the first language (L1) adapts to the grammar and words of L2, and vice versa (Kroll et al., 2014). Importantly, empirical evidence has shown that cross-language interactions vary as a function of overlap during word production (Schwartz et al., 2007) and reading (Van Assche et al., 2011). If cross-language interactions vary with the similarity of the two languages spoken, language similarity may have a profound impact on how much executive control is required to effectively use L1 and L2. Basically, language similarity could affect executive control demands in language selection in two ways. First, similar L1 and L2 could lead to stronger cross-language interference. If so, selecting the appropriate language should become more difficult the more similar the two languages are, thereby requiring more executive control to inhibit the interfering language, to reduce the costs of switching between languages, and to monitor the contents that get access to WM (Linck et al., 2008; Barac and Bialystok, 2012; Coderre and van Heuven, 2014). In this case, bilinguals with similar languages would train to exert executive control more intensively, leading to enhanced performance in EF tasks. Alternatively, it may be that similar languages yield more adaptation between languages, thereby facilitating lexical access and language comprehension due to their shared grammar, syntax, and phonology. If so, speaking two highly similar languages should reduce the need to exert executive control compared to speaking two more dissimilar languages. In this latter scenario, dissimilar languages would require stronger attentional control, increase the cost of switching between languages, and demand more monitoring of WM contents, thereby yielding more training of EF. In this case, bilinguals speaking similar languages would be less advanced in EF than those speaking dissimilar languages. These opposing views can be disentangled by assessing the impact of language similarity on both linguistic and EF tasks. By taking linguistic performance as a measure of the degree to which the two languages interfere with each other and, hence, of how much executive control is required for managing L1 and L2, it is possible to predict how language similarity modulates bilingual EF advantages.

There is evidence of facilitating effects of language similarity in bilingual children (Bialystok et al., 2003, 2005; Barac and Bialystok, 2012). However, a study with young adults has not found any differences in language switching costs as a function of language similarity (Costa et al., 2006). Thus, more evidence is needed to test for the effect of language similarity on linguistic performance in adulthood. Regarding EF performance, only a small number of studies assessed the impact of language similarity, with mixed results. Three studies found no effect of language similarity on EF performance: one study tested bilingual children on a shifting task (Barac and Bialystok, 2012); one study tested young adults (Linck et al., 2008) and another study tested older adults (Kirk et al., 2014) on an inhibition task. Yet another study with a sample of young adults yielded inconclusive results

on an inhibition task (Coderre and van Heuven, 2014): bilinguals with dissimilar languages showed the smallest interference score in a Stroop task, but they also responded more slowly on the task. In an attempt to replicate this result, Paap et al. (2015) assessed Stroop performance in young adult monolinguals and three groups of bilinguals with varying script similarity. However, script similarity affected neither Stroop interference nor overall reaction times (RT). Instead, orthographic overlap between the two languages spoken was associated with slower RTs in the Stroop task (but not with Stroop interference). Taken together, the evidence for an effect of language similarity on EF performance is mixed.

An extreme form of language similarity is bidialectalism (i.e., speaking a dialect in addition to a standard language). Dialects are naturally tightly related to their originating languages, while still having a distinct grammar and phonology (Chambers and Trudgill, 1998). Yet, only few studies have related bidialectalism to bilingualism. Antoniou et al. (2016) assessed performance in several EF tasks in children that were monolinguals, bidialectals, or bilinguals. Bilinguals and, to some extent, bidialectals outperformed monolinguals in a composite measure of WM and inhibitory control. Noteworthy, the EF advantage of bidialectals was weaker than that of the bilinguals, and only reached significance after covarying children's verbal capacity. In contrast, Ross and Melinger (2017) tested monolingual, bidialectal, and bilingual children in tasks measuring inhibitory control and shifting. Bidialectalism did not yield a benefit in either measure, but bilinguals responded more accurately in one inhibition task. In a sample of older adults, Kirk et al. (2014) found that bidialectals performed similarly as monolinguals in a Simon task. To the best of our knowledge, there is no study focusing on young adulthood.

Taken together, the little research to date on the effects of language similarity on linguistic and EF tasks led to mixed results. The variability across studies may be due to several factors. First, most studies assessed either EF or linguistic performance, but not both (with exception of Barac and Bialystok, 2012). This makes it difficult to establish the link between language processing demands and executive control. Second, most studies assessed performance in only one task, and studies vary in terms of the ability assessed (e.g., inhibition, shifting). Recently, it has been proposed that bilingualism may have a subtle impact on diverse EF measures (Kroll and Bialystok, 2013). Hence, a broader assessment of EF might be required to uncover the effects of language similarity. Third, single-task assessments may also confound task-specific variance with ability-level effects (Shipstead et al., 2012). Thus, studies including multiple tasks measuring the same ability may provide more reliable and generalizable results.

The Present Study

As reviewed above, there is little research on the effects of language similarity on EF and linguistic processing, particularly among young adults. In the present study, we investigated performance in these two domains simultaneously to examine whether and how language similarity mediates the relationship between language control and executive control. We hypothesize

that there are two possible scenarios. Similar languages may lead to more linguistic interference and, thus, require increased executive control relative to dissimilar languages (Linck et al., 2008; Barac and Bialystok, 2012; Coderre and van Heuven, 2014). Alternatively, similar languages may interfere less with each other due to cross-linguistic adaptation. Adaptation should facilitate linguistic processing in more similar languages, thus requiring less executive control compared to dissimilar languages (Barac and Bialystok, 2012). Either way, both views imply that language similarity has opposite effects on EF and linguistic tasks: the conditions that lead to better linguistic performance should yield least training of EFs, thereby limiting EF benefits. The main goal of the present study was to provide a first assessment of this link by measuring how language similarity affects both linguistic and EF performance.

We assessed the effect of language similarity in EF and linguistic tasks by comparing performance of monolinguals, bidialectals, and bilinguals with language combinations that varied in the similarity to Standard German. As this was the first attempt to test the impact of language similarity on the link between EF and linguistic performance, no evidence was available as to what linguistic properties on what level (i.e., orthographical, phonological, semantic) would be most critical for language similarity to yield the hypothesized effects. Hence, we chose a more general measure of language overlap by categorizing languages as similar based on their language family. We considered languages within the Indo-European family as more closely related to each other than they are to languages of any other family based on some overlapping vocabulary and similarities in their general macro-structural syntax and grammar. In contrast, Indo-European and Non-Indo-European languages tend to differ in those aspects to a larger degree (Campbell, 2008; Comrie, 2008; Longobardi et al., 2013). Accordingly, our assumption was that individuals who speak languages from the Indo-European family would broadly deal with languages that share more linguistic properties than individuals whose languages stem from different families¹. Several studies have shown bidirectional cross-linguistic interactions indicating that the multiple languages an individual speaks affect each other (Hohenstein et al., 2006; Brown and Gullberg, 2008, 2010, 2011; Ameel et al., 2009; Van Assche et al., 2011). Based on the BIA+ model, shared linguistic properties should lead to the activation of more lexical representations in the bilingual lexicon that show overlap with the target word, thus resulting in increased cross-linguistic interactions for more similar languages (Dijkstra and van Heuven, 2002). Hence, the broad similarity of the languages may facilitate or hinder language processing and in turn demand less or more EF. Taking the present definition of language similarity, we assume that Standard German and the Swiss-German dialect share the highest degree of overlap (e.g., shared vocabulary, phonology, syntax, etc.). In contrast, when one compares Standard German to other languages

¹This classification is at a broad ordinary level, as we had no means of computing the degree of language overlap between all language combinations reported by our participants.

from the Indo-European family (e.g., German and English, or German and French), there are much less shared properties (e.g., vocabulary), but there remain some macro-structural similarities such as phonological and syntactic processes that could impact the learning and daily usage of these language combinations. Languages from different language families (e.g., German and Turkish, or German and Chinese), conversely, are assumed to share even less properties than languages within the same language family (e.g., different vocabulary, different phonology, etc.). We assumed that there should be less cross-linguistic interactions between these more dissimilar languages than between the (relatively) more similar languages within the Indo-European family.

Having these considerations in mind, participants were classified as belonging to one of four groups. The *monolingual* group comprised native speakers of Standard German only. The *bidialectal* group comprised native speakers of Standard German and the Swiss German dialect. The Swiss German dialect is very closely related to Standard German, as both belong to the German languages within the Indo-European language family, and are located on neighboring branches of the family tree (Simons and Fennig, 2018). The bidialectals in our group used both the Swiss German dialect and Standard German in most social contexts. Bilinguals were speakers of Standard German (and, in most cases, also of the Swiss German dialect) and learned another language. We included a group of bilinguals proficient in Standard German and another Indo-European language (hereafter *similar bilinguals*, e.g., English, French, or Italian), and a group of bilinguals proficient in Standard German and a Non-Indo-European language (*dissimilar bilinguals*, e.g., Arabic, Turkish, or Chinese). Performance of these four groups was compared in three tasks aimed at assessing their linguistic ability and in several measures of EF that have been linked to inhibitory control, monitoring, shifting, mixing, and WM (von Bastian et al., 2016).

In sum, our four language groups differ progressively in terms of which additional languages they spoke. Monolinguals spoke only Standard German, bidialectals spoke Standard German and the Swiss German dialect, and bilinguals spoke Standard German, the Swiss German dialect, and another language that was of the same Indo-European family (similar bilinguals) or not (dissimilar bilinguals). Hence, this partition assumes that speaking additional languages has an additive effect with speaking the dialect. This might not be the case, and it is also conceivable that speaking the dialect and speaking an additional language have opposite effects that cancel each other, thereby diluting group differences. This was a risk of our design. However, this would be an actual concern only if we would find no effects of language similarity neither in linguistic nor in EF performance, or if the effect was constrained only to the comparison between monolinguals and bidialectals with no further differences for the bilingual groups. To foreshadow our results, we did obtain a monotonic effect of language similarity on linguistic performance, which is in line with the assumption that speaking the dialect and an additional language have additive effects.

MATERIALS AND METHODS

Participants signed up for the study via an online form determining their eligibility for study participation (physically and psychologically healthy, not color-blind, and speaker of Standard German). Eligible participants were invited via e-mail to complete an online language history questionnaire (completed in Standard German). Next, they were invited for a laboratory session where they completed a battery of tasks taking approximately 2 h, with a 10-min break midway. All tasks were presented in Standard German. During the laboratory session, participants first completed the Ishihara test for color blindness (Ishihara, 2003), followed by two paper-pencil tasks measuring linguistic ability (a word completion test and a verbal fluency test). Then, they were asked to complete a computerized test battery comprising 11 cognitive tasks. The test battery was programmed with Tatool, an open-source software for programming psychological experiments (von Bastian et al., 2013). To avoid order and fatigue effects, half of the participants completed the computer-based tasks in reversed order (von Bastian and Oberauer, 2013). Participants were randomly assigned to the task order, equally balanced across language groups.

Participants

One-hundred and eleven young adults voluntarily took part in the study. Participation was compensated with extra-course credit or 40 CHF (about 40 USD). Participants were students at a Swiss university or held a diploma comparable to a Swiss high-school certificate (Matura). Written informed consent was obtained for all participants. Participants were tested in groups of up to five. The experimental protocol was approved by the Ethics Committee of the Faculty of Arts and Social Sciences of the University of Zurich (in accordance with the Helsinki declaration), and participants were debriefed at the end of the study. Twelve participants were excluded from the analysis for various reasons: (a) they were not proficient in German ($n = 1$), (b) reported language combinations that did not match the pre-defined language groups ($n = 7$)², or (c) did not fulfill the requirements of our definition of bilingualism described below ($n = 4$). Thus, the final sample consisted of 99 participants, aged 18–33 years ($M = 23.5$, $SD = 3.69$). The sample characteristics are listed in **Table 1**.

Participants were classified into one of four language groups. *Monolinguals* ($n = 25$) were native speakers of only Standard German and had limited knowledge of the Swiss German dialect. *Bidialectals* ($n = 26$) were native speakers of the Swiss German dialect and Standard German. In the German speaking part of Switzerland, the Swiss German dialect is used in most daily interactions and is typically the language children will learn first at home. However, Standard German is used when interacting

²One participant was native speaker of a dialect of German other than Swiss German (i.e., Austrian), and hence this person could not be classified into either the monolingual or bidialectal groups. The other six participants qualified as similar bilinguals but used a different (non-Latin) script in one of their languages. We excluded these participants to keep script similarity constant in the similar bilingual group.

TABLE 1 | Demographic characteristics by language group.

Measure	Monolinguals <i>n</i> = 25			Bidialectals <i>n</i> = 26			Similar bilinguals <i>n</i> = 24			Dissimilar bilinguals <i>n</i> = 24			Evidence for/against group differences		
	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>		<i>BF</i> ₁₀	% error	
Gender (f/m)	24/5	3.55	20/5	23/46	4.32	21/5	22.50	3.24	16/8	23.38	3.47	19/5	0.15	±0.00	
Age	24.48												0.20	±0.00	
Education (1–4) ^a	2.08	0.95		1.69	0.97		1.54	0.88		1.63	0.92		0.32	±0.00	
Parents' education (1–8) ^b	5.54	1.42		4.42	1.47		5.02	1.58		4.19	2.10		2.07	±0.00	
Migration background (n) ^c	22			2					8			10	>1,000	±0.00	
RAPM	0.63	0.21		0.70	0.16		0.58	0.29		0.61	0.27		0.19	±0.00	

RAPM, Raven's Advanced Progressive Matrices. Values printed in bold indicate at least substantial evidence for ($BF_{10} \geq 3$) or against ($BF_{10} \leq 0.33$) group differences.

^aHighest completed degree [1, matura degree (higher education entrance qualification); 2, university of applied sciences degree; 3, university degree; 4, doctoral degree].

^bAverage of mother's and father's educational degree (1, no or less than 9 years of school; 2, mandatory school of approx. 9 years; 3, vocational education; 4, further professional education; 5, matura degree; 6, university of applied sciences; 7, university degree; 8, doctoral degree).

^cCountry of residence differs from country of origin.

with non-dialect speakers, in the news, on most of the available TV channels, and in some other social contexts. Thus, Swiss Germans are highly proficient in both the Swiss German dialect and Standard German. Moreover, the Swiss German dialect is a spoken dialect only, with Standard German being the written language that is obligatory in formal contexts. Hence, all children must learn Standard German when entering school at the age of 6 or 7 years. Participants in the monolingual and bidialectal groups had also formal foreign language education during secondary school (most commonly English or French starting on average after the age of 10), but achieved much lower proficiency in these languages (see **Figure 1**). The two remaining groups were bilinguals. Participants qualified as bilinguals if they (1) learned at least one language (henceforth L2) in addition to Standard German and/or Swiss German up to the age of seven (i.e., before entering school and any formal foreign language education), and (2) indicated that they were still actively using their L2. This definition of bilingualism is in line with the inclusion criteria used in several prior studies (e.g., see Bialystok et al., 2006; Costa et al., 2009; Hernández et al., 2010; Luk et al., 2011; Gold et al., 2013). In sum, what separates our bilingual participants from their monolingual and bidialectal counterparts is the early onset of bilingualism, and also the greater proficiency they achieved in their languages. Bilinguals with an L2 from the Indo-European language family were classified as *similar bilinguals* (*n* = 24), and bilinguals with an L2 from a Non-Indo-European language family were classified as *dissimilar bilinguals* (*n* = 24). Similar bilinguals were native speakers of English (7), French (3), Italian (3), Polish (2), Portuguese (2), Spanish (3), Rhaeto-Romanic (3), or Albanian (1). Dissimilar bilinguals were native speakers of Chinese (2), Korean (1), Laotian (1), Tagalog (1), Tamil (1), Tibetan (1), Turkish (7), Hungarian (6), Finnish (1), Arabic (2), or Malayalam (1).

The language groups were matched in terms of gender, age, educational background, and Raven's Advanced Progressive Matrices (RAPM) scores as confirmed by substantial evidence against group differences in univariate Bayesian ANOVAs (see **Table 1**). The evidence regarding group differences in socioeconomic status (SES) was ambiguous. Paired contrasts between groups revealed, however, that all groups were comparable in terms of their SES, except for the parents of monolinguals having, on average, higher educational degrees than parents of bidialectals. Groups differed though regarding their migration background: Most monolinguals and more than a third of the bilinguals, but less than 8% of the bidialectals were currently residing in a different country than their country of origin.

Language and Demographic Assessment

Demographic and language background information were assessed with a questionnaire administered online using SoSci Survey (Leiner, 2018). The questionnaire was based on the language history questionnaire from Li et al. (2006), translated to German, and adapted for the purposes of this study by two of the authors (JO and AS). After assessing demographic variables and SES, participants were asked to report all languages they have learned (up to a maximum of five languages) in

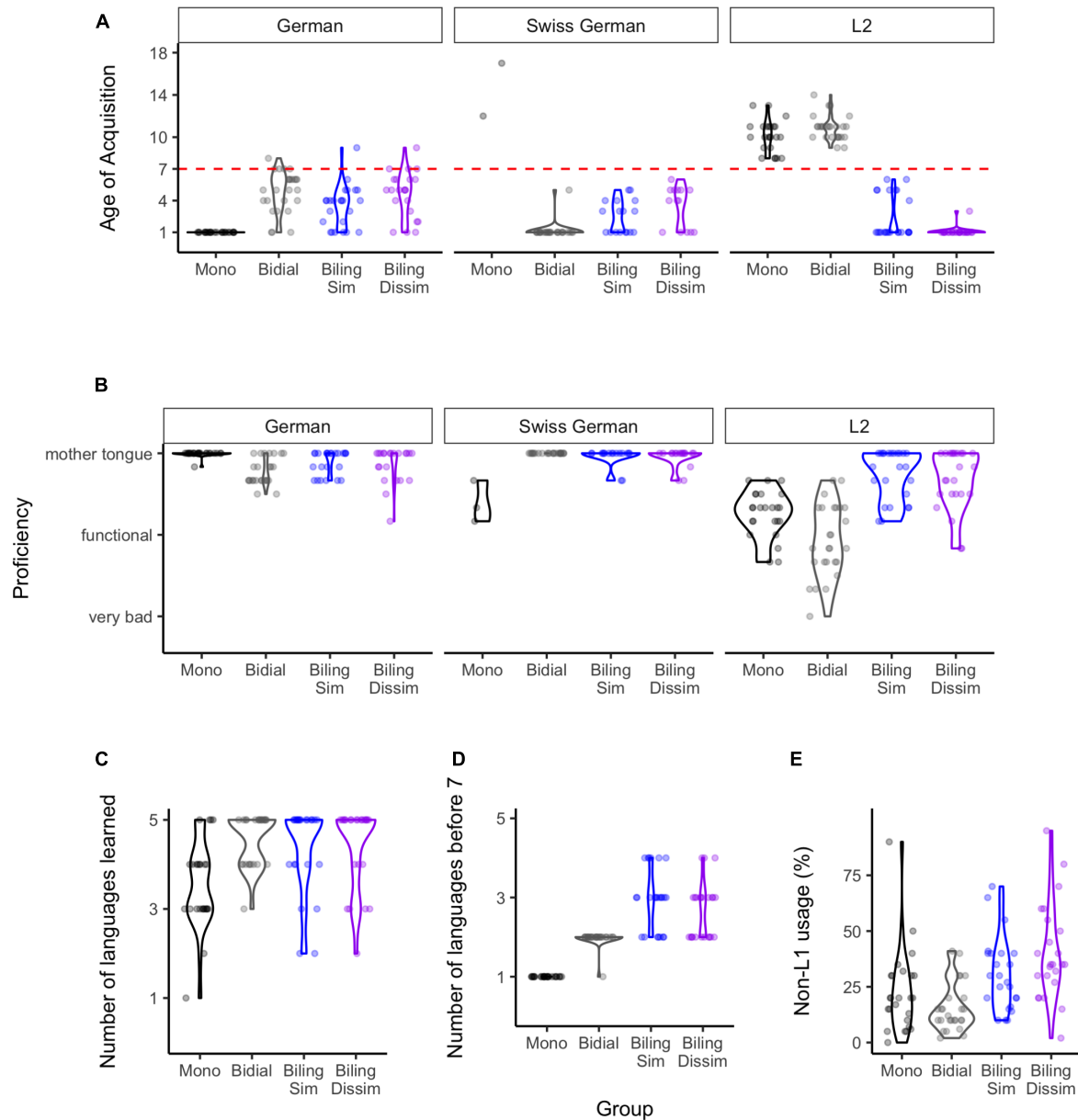


FIGURE 1 | Distribution of the self-reported language variables in each group. **(A,B)** show age of acquisition and proficiency, respectively, in German, Swiss German, and the L2. **(C)** shows the total number of languages learned, whereas **(D)** shows the number of languages learned before the age of seven. Finally, **(E)** presents the daily percentage usage of languages other than German/Swiss German.

the order in which they had learned them, starting with their native language. Participants were explicitly instructed to consider the Swiss German dialect as a separate language. In addition, they were asked to indicate detailed information on the usage of each language. For the present purposes, we extracted the self-reported age of acquisition, proficiency, and percentage of daily language usage for German, Swiss German dialect, and each participant's L2 (language other than German and/or the Swiss German dialect acquired earliest) from the questionnaire. Previous research obtained high correlations of self-rated proficiency measures with objective assessments of

language proficiency (Luk and Bialystok, 2013). Accordingly, we used the above listed self-reported measures to describe the language experience of our groups, and also as continuous predictors of performance in our tasks.

Figure 1 presents the distribution of language background variables in each language group. Panel A indicates the self-reported age of acquisition of each language of interest here, namely German, Swiss German dialect, and the L2. There are clear differences between the language groups, particularly with regards to age of L2 acquisition, which was the inclusion criterion for bilinguals in this study. Note that with regards to the Swiss

German dialect, only two monolinguals reported having learned the dialect. Complementarily, panel B presents self-reported proficiency in each of these languages. Again, language groups differed substantially particularly with regards to L2 proficiency: bilinguals reported higher proficiency than monolinguals or bidialectals. The two monolinguals that reported learning Swiss German also reported that their proficiency on the dialect was lower than that of a native speaker. Panel C shows that most participants in the study learned more than one language at some point in their lives (note that Swiss German is included here as an additional language). Panel D indicates, however, that the age of acquisition of the learned languages differed between groups: monolinguals acquired only one language by the age of 7, whereas bidialectals acquired two languages (i.e., the Swiss German dialect and Standard German), and bilinguals (similar and dissimilar groups) acquired two or more languages (i.e., the Swiss German dialect and/or Standard German and the L2). Lastly, panel E shows that participants in all groups reported using a foreign language (i.e., another language besides the Swiss German dialect or Standard German) for a substantial part of their day. This is probably the case because all participants were university students, and they were confronted with English on a daily basis. Importantly, the item did not differentiate between active (e.g., speaking) and passive (e.g., listening) non-L1 usage. In sum, our bilinguals learned more languages at an earlier age and with higher proficiency than monolinguals or bidialectals.

Linguistic and EF Assessment

Linguistic ability was assessed with three tasks, and the five EF abilities (inhibition, monitoring, mixing, shifting, and WM) each with two tasks using different materials to reduce the influence of task-specific variance. Furthermore, we included a short non-verbal reasoning test to assess group comparability on this ability. All tasks were preceded by practice trials which were excluded from the final analysis. Dependent measures were coded so that larger values indicate better performance.

Linguistic Ability

Bilinguals have been consistently found to be disadvantaged in linguistic tasks compared to monolinguals: they produce less words in semantic fluency tasks (Gollan et al., 2002; Portocarrero et al., 2007) and react slower and less accurately in lexical decision tasks (Ransdell and Fischler, 1987; Lehtonen et al., 2012). Moreover, Ransdell and Fischler (1989) found that bilinguals benefitted less from accessing concrete in comparison to abstract words. Hence, we assessed linguistic ability through performance in the three tasks: verbal fluency, lexical decision-making, and the concreteness effect in a word recognition task. All tasks were conducted in Standard German language. In the *verbal fluency task* (administered in paper-and-pencil format), participants were asked to write down as many German words as they could think of in response to a categorical prompt (i.e., animals, fruits, clothes, musical instruments, objects on wheels, and furniture) within 2 min for each category. Words from the same semantic subcategory (e.g., poodle and labrador from the subcategory dogs), or words with the same meaning (e.g., “Orange” and “Apfelsine” both of which refer to an orange in Standard German)

were coded as one word only. Linguistic accuracy in the verbal fluency task was measured via the sum of unique words (average across two coders) generated across all semantic categories. Participants also completed a word-fragment completion test (Jäger et al., 1997) that was administered as a warm-up for the following verbal fluency test. Data of this task were discarded and not further analyzed. For the remaining two tasks, we derived two performance measures, namely the accuracy with which the task was completed (hereafter referred to as linguistic accuracy) and the speed of processing (linguistic speed). In the *lexical decision task*, participants indicated with a key press whether a visually presented string was a word (right arrow key) or a non-word (left arrow key). The stimulus remained onscreen until a response was made (see **Figure 2A**). Participants completed 128 trials consisting of a pseudo-random sequence of 64 German words and 64 non-words, matched regarding their number of letters and syllables, and their frequency (only words) using a semantic atlas for German words (Schwibbe et al., 1981). We calculated linguistic speed using the mean log-transformed RTs (multiplied by -1, so that higher values represent better performance), and linguistic accuracy via detection performance computed as $d' = z(H) - z(FA)$, with H being the hit rate, FA being the false alarm rate, and z reflecting the z-transformation of these values. In the *word recognition task*, participants were instructed to memorize 30 German nouns presented sequentially (3 s each) on the screen (see **Figure 2B**). Half of the nouns were concrete (e.g., elephant) and half of them abstract (e.g., theory). Subsequently, participants were sequentially shown 60 probe words, including the 30 previously presented words (old) and 30 new words (new), randomly intermixed. Participants decided whether the probe word was old (right arrow key) or new (left arrow key). Each probe word remained onscreen until a response was made. The concreteness benefit (i.e., the performance difference in responding to abstract and concrete words) in log-transformed RTs was used as a measure of linguistic speed, and the concreteness benefit in detection performance (d') was used as a measure of linguistic accuracy. Both measures were coded so that a larger value reflects a larger concreteness benefit.

Inhibition

Bilinguals' extensive practice inhibiting their currently irrelevant language (Green, 1998) is assumed to yield advantages in inhibiting irrelevant information in non-linguistic tasks. We used two tasks to assess inhibition. In the *flanker task*, participants indicated as fast and accurately as possible whether the central letter (target) in a string of seven letters was a vowel (left arrow key) or consonant (right arrow key). The stimulus remained onscreen until a response was given, followed by an inter-trial interval (ITI) of 250 ms (see **Figure 2C**). Participants completed 144 trials. In one third of the trials, the letters flanking the target were congruent (target and flankers require the same response, e.g., “SSSTSS” or “EEEEEEE”), incongruent (target and flankers require the opposite response, e.g., “SSSASS” or “EEETEEE”), or neutral (flankers are irrelevant to the task, e.g., “###S###” or “###A###”). As an inhibition index, we computed the difference between the log-transformed RTs in neutral and incongruent trials. In the *Simon task*, each trial started with a fixation cross

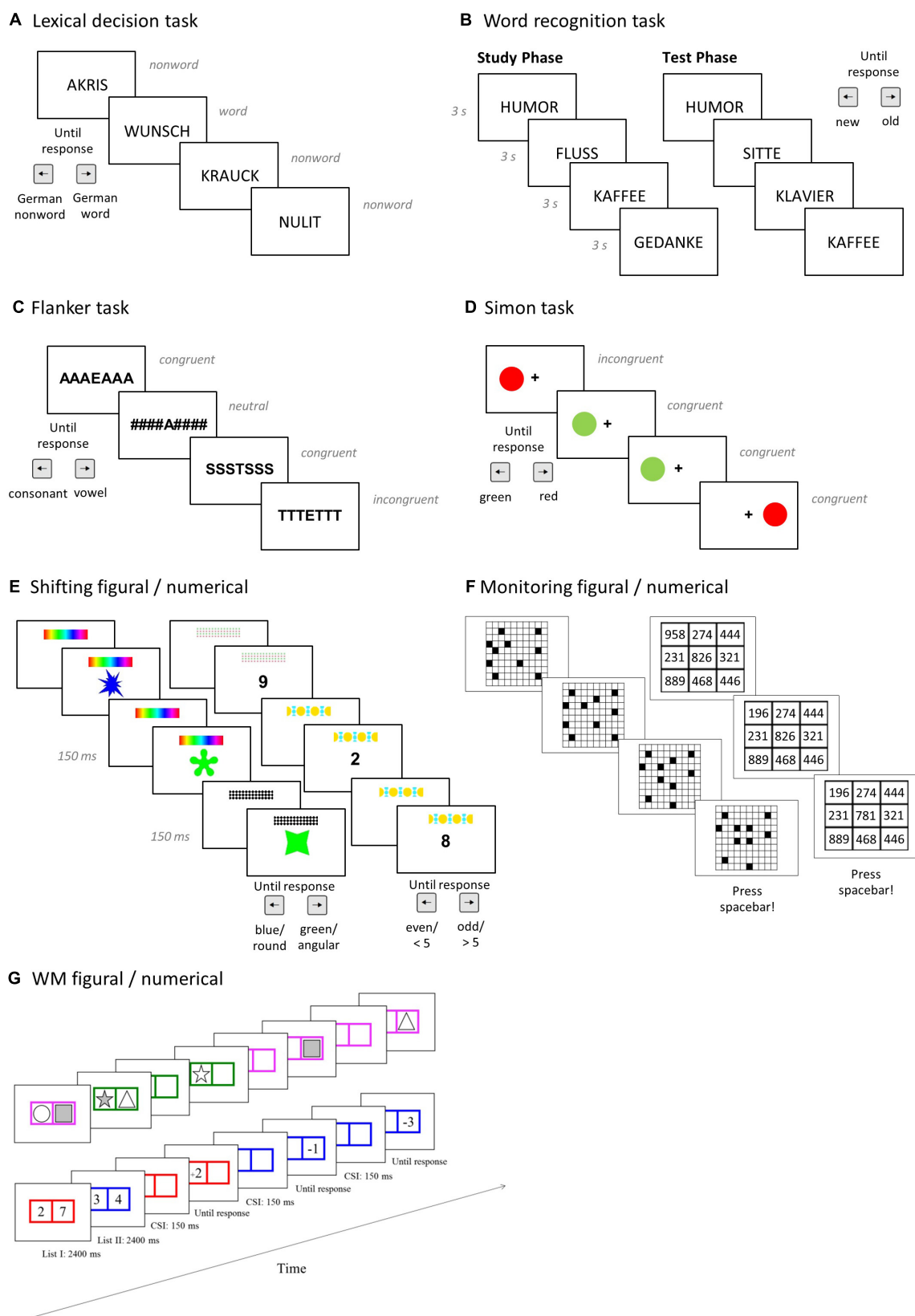


FIGURE 2 | Graphical representation of the tasks administered. See text for details. CSI, cue-stimulus interval. **(A,B)** Linguistic tasks. **(C,D)** Inhibition tasks. **(E)** The figural and numerical versions of the shifting tasks. **(F)** The figural and numerical versions of the monitoring tasks. **(G)** The figural and numerical versions of the working memory (WM) tasks.

presented centrally for 250 ms, followed by a colored circle appearing on the left or right side of the screen (see **Figure 2D**). Participants had to indicate as fast and as accurately as possible whether the circle was red (right arrow key) or green (left arrow key). The circle remained onscreen until a response was made, followed by an ITI of 250 ms. Participants completed 200 trials: 75% were congruent, that is, the location of the response (e.g., left) and the spatial location of the stimulus (e.g., left) matched, and 25% were incongruent trials in which the spatial location of the response and of the stimulus did not match. As an inhibition index, we computed the difference between the log-transformed RTs in congruent and incongruent trials.

Shifting and Mixing

Bilinguals are also assumed to benefit from the extensive practice in switching between languages that generalizes to shifting between tasks, yielding smaller non-linguistic task-switch costs (for a review see Yang et al., 2016). In addition, bilinguals are assumed to excel in monitoring which task to apply in which situation (Soveri et al., 2011; Wiseheart et al., 2016), which is reflected by mixing costs that can also be assessed with shifting tasks. Therefore, we used the figural and a numerical switching tasks from von Bastian et al. (2016) consisting of single-task blocks, where only one task is performed, and a mixed-task block in which two tasks switch unpredictably. In the *color-shape task*, participants classified bivalent figural stimuli according to their color (blue or green) or shape (round or angular) by pressing the left (for blue or round) or right arrow key (for green or angular). The task included 32 angular and 32 round shapes, with half of each colored in blue or green, respectively. In the *parity-magnitude task*, participants classified digits from 1 to 9 (excluding 5) according to their parity (even or odd) or magnitude (smaller or larger than 5) by pressing the left (for even or smaller than 5) or right arrow key (for odd or larger than 5). In both task versions, the upcoming task rule was indicated by an abstract cue (e.g., patterned bar) presented on the top of the screen. After a cue-stimulus interval (CSI) of 150 ms, a shape or digit (depending on the task version) appeared in the center of the screen until participant's response (see **Figure 2E**). Participants completed two single-task blocks (one for each task) of 64 trials each, followed by a mixed-task block of 129 trials, and again the two single-task blocks (in reversed order). The mixed-task block contained 50% repeat trials (i.e., trials in which the task in the current and preceding trial was the same), and 50% switch trials (i.e., the tasks in the current and the preceding trial were different). The first trial in the mixed block was excluded from analysis, as it was neither a repeat nor a switch trial. To assess shifting ability, switching cost scores were calculated by subtracting the average RT in switch trials from the average RT in repeat trials (both from the mixed-task block and log-transformed). To assess mixing, mixing cost scores were computed by subtracting the average repeat trials RT in the mixed-task block from the average RT in the single-task block (both log-transformed).

Monitoring

Monitoring was measured with tasks requiring participants to sustain attention to a stream of inputs to detect certain patterns or relations. Participants completed two tasks from von Bastian et al. (2016) in which they had to monitor independently changing objects, and react whenever a predefined relation between these objects occurred (Oberauer et al., 2003; von Bastian and Oberauer, 2013). In the *squares task*, a display of 20 dots in a 10 × 10 grid was shown and, every 2 s, two dots randomly changed their position. Participants had to press the space key whenever four dots formed a square. In the *digits task*, a 3 × 3 grid with three-digit numbers in each cell was presented and, every 2 s, the numbers in one cell changed. Participants were instructed to press the spacebar whenever the last digits of the numbers in a row, column, or diagonal were identical (see **Figure 2F**). Both task versions comprised 16 trials, each presenting 2 to 8 changes before the predefined relation between objects appeared. The monitoring score was d' .

WM

WM is assumed to be tightly related to executive control (Engle, 2002), which makes it one candidate EF domain for bilingual benefits (Bialystok, 2017). WM was assessed with a figural and numerical version of the list-switching paradigm (Oberauer, 2005; Oberauer et al., 2013; Gade et al., 2017) in which participants had to retain two memory lists in WM for ongoing processing (see **Figure 2G**)³. In the *figural task*, participants memorized two lists distinguished by a pink or green frame. Each list consisted of a row with two colored boxes each containing a filled shape (selected from a pool of 20 shapes with no replacement). The lists were presented sequentially (for 2400 ms each) with a 250 ms inter-list blank interval. Next, 13 memory probes followed. For each probe, the relevant list was cued by the color of the row of boxes and, 150 ms thereafter, the probe appeared in one of the boxes. Participants indicated a match between the probe and the memory item in the same list position (left arrow key, 50% of the trials) or a mismatch (right arrow key, 50% of the trials). Mismatch probes were shapes presented in another list or list position (50%), or not presented in the current trial at all (50%). The *numerical task* followed a similar task structure: participants memorized a red and a blue list, each consisting of two digits (ranging from 1 to 9). Again, a series of 13 memory probes followed in which the relevant list was cued by color, and 150 ms later an arithmetic operation was shown in one of the boxes (e.g., +2). Participants had to retrieve the item shown in this position of the relevant list, apply the arithmetic operation to it, and enter the result (which was always between 1 and 9) using the keyboard. Participants entered the results of the operation, but they were asked to remember the original value of the item. In both task versions, each sequence of memory list encoding followed by 13 probes was considered

³Before and after this task, participants also completed single-list blocks which were similar with the exception that only one list was memorized and consequently there was no switching between lists. In the present study, we considered only performance in the two-list block, because this condition places a higher demand on WM capacity which has been suggested as important for observing a bilingual effect (see Bialystok et al., 2004).

a run. Participants completed 12 runs, each containing 50% list-repeat trials (same list was tested in the current and the previous trial) and 50% list-switch trials (current and previous list were different). The WM scores were proportion of correct responses in both the figural and numerical version.

Reasoning

To evaluate whether the language groups matched regarding their non-linguistic fluid intelligence, we administered a short (Arthur and Day, 1994) computerized version of Raven's Advanced Progressive Matrices (Raven, 1990). Participants had 15 min to complete 12 patterns. For each pattern, they had to choose 1 out of 8 response alternatives. The number of correctly solved items (out of 12) served as dependent measure.

STATISTICAL ANALYSIS

Data Preprocessing

For RT based scores, we removed RTs associated with incorrect responses. Next, RTs were trimmed by removing outliers. Outliers were defined as RTs being three median absolute deviations away from the overall median (Leys et al., 2013). RTs were log-transformed to better approach normality before computing the relevant RT-based scores. To eliminate the unwanted source of variance introduced by having administered two test orders, we arbitrarily chose one order as the reference, and corrected the data of the other order for the mean difference between them (von Bastian and Oberauer, 2013; von Bastian et al., 2016). Lastly, all task scores were z-transformed.

Bayesian Linear Mixed-Effects Modeling

We analyzed our data with Bayesian linear mixed-effects models. The advantage of using Bayesian statistics is that the evidence supporting both the alternative and the null hypothesis can be assessed. We used the BayesFactor package (Morey and Rouder, 2015) implemented in R (R Core Team, 2017), with the default prior settings (i.e., $r = 0.50$). The *lmBF* function implemented in the package computes the strength of the evidence for a specified model (M_1) against a Null model (M_0). For example, M_1 may state that performance of monolinguals differs from bidialectals (alternative hypothesis), whereas M_0 states that there is no group effect (null hypothesis). The ratio of the likelihood of these two models given the data is the Bayes factor (BF). The BF is the factor by which prior beliefs should be updated in light of the data. For example, a BF for M_1 over M_0 (hereafter, BF_{10}) of 5 translates into the data being five times more likely under the alternative than under the null hypothesis. Likewise, $BF_{10} = 0.2$ means that the data are 5 times more likely under null hypothesis than the alternative hypothesis. When $BF_{10} = 1$, the data are equally likely under both hypotheses and, hence, the evidence is ambiguous. It is common to consider $BF_{10} \geq 3$ as providing substantial evidence for the alternative hypothesis over the null, and $BF_{10} \leq 0.33$ as providing substantial evidence for the null over the alternative hypothesis (Wagenmakers et al., 2011).

We tested for an effect of language similarity in each cognitive ability separately, using two approaches. First, we coded language

similarity with a linear contrast over language group (using the *poly* function in R) and entered this variable as a fixed predictor in the models. This contrast implements the assumption that language groups differ in a monotonically decreasing fashion regarding language similarity. Second, to facilitate comparability to previous studies on effects of language similarity, we compared adjacent levels of language similarity (aka. sliding contrast; i.e., monolinguals vs. bidialectals, bidialectals vs. similar bilinguals, and similar bilinguals vs. dissimilar bilinguals). In addition, to test for bilingual effects more commonly investigated in the literature, we also contrasted the monolingual group with the similar and dissimilar groups (simple contrasts). For each model, we included random intercepts for participant and for task. We also included parents' education level as a proxy for SES as a continuous predictor in all analyses (von Bastian et al., 2016). Two participants failed to provide information regarding their parents' education level (one monolingual and one similar bilingual). To keep these participants in the sample, we replaced their missing values with the average SES of their respective groups. Excluding these participants from the analyses altogether did not substantially change the pattern of results. All analyses were computed with a high number of iterations (i.e., 400,000) to ensure that the error in estimating the BF was below 5%.

The data and analysis scripts for performing the analyses reported here are available at the Open Science Framework (OSF) at <https://osf.io/uf2hs>. The computer-based tasks used here (except the Raven's) are freely and publicly available on Tatool Web (www.tatool-web.com). **Supplementary Materials** are available at the journal website (URL) and also at the OSF.

RESULTS

Descriptive statistics for all (non-transformed) measures as a function of language group are listed in **Table 2**. Zero-order correlations between measures and reliabilities are listed in **Table 3**. Split-half reliabilities (for difference scores, d' and RTs; corrected with the Spearman-Brown prophecy formula) and Cronbach's alpha (for accuracies) were within the acceptable range for all scores, except for the accuracy and speed scores derived from the word recognition task and the flanker inhibition score. All measures assessing the same ability were significantly positively intercorrelated, except for the linguistic accuracy scores (although the correlation between the verbal fluency and lexical decision task was marginally significant: $p = 0.051$, $r = 0.20$), linguistic speed scores (for which the correlation was negative), and the flanker and Simon inhibition scores. The evidence for all expected effects (concreteness effects, inhibition, mixing, and shifting costs) was substantial (see **Table 4**). To rule out that problems with reliability or lack of correlations between tasks were masking the effects of interest, we additionally ran all analyses on the level of individual tasks (see **Supplementary Table S1**).

Figure 3 presents the (z-transformed) measures as a function of language group and cognitive ability, and the predictions of the linear contrast over language similarity. The predictions represent the mean and the 95% highest-density interval (HDI)

TABLE 2 | Descriptive statistics for each task score by language group.

Measure	Monolinguals		Bidialectals		Similar bilinguals		Dissimilar bilinguals	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Linguistic accuracy								
Lexical decision	4.18	0.51	4.07	0.45	3.92	0.56	3.62	0.66
Word recognition	0.36	0.79	0.70	0.72	0.41	0.87	0.11	0.63
Verbal fluency	108.28	23.26	110.37	17.90	107.35	20.33	99.40	17.02
Linguistic speed								
Lexical decision	−6.56	0.14	−6.58	0.13	−6.57	0.18	−6.55	0.13
Word recognition	0.04	0.06	0.04	0.06	0.01	0.05	0.01	0.05
Inhibition								
Flanker	0.00	0.06	−0.02	0.04	−0.01	0.04	−0.01	0.04
Simon	−0.20	0.06	−0.18	0.08	−0.16	0.05	−0.17	0.06
Monitoring								
Figural	2.43	0.51	2.53	0.40	2.62	0.53	2.45	0.43
Numerical	2.25	0.55	2.58	0.66	2.61	0.85	2.66	0.82
Mixing								
Figural	−0.60	0.16	−0.51	0.21	−0.48	0.17	−0.55	0.22
Numerical	−0.31	0.19	−0.28	0.21	−0.26	0.16	−0.25	0.20
Shifting								
Figural	−0.33	0.22	−0.28	0.17	−0.37	0.19	−0.27	0.17
Numerical	−0.35	0.20	−0.36	0.19	−0.37	0.16	−0.33	0.17
Working memory								
Figural	0.78	0.08	0.83	0.08	0.82	0.09	0.80	0.11
Numerical	0.89	0.16	0.95	0.11	0.94	0.08	0.96	0.06

See text for dependent measures of the variables listed.

TABLE 3 | Correlations between measures and reliabilities.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Linguistic accuracy															
1. Lexical decision	0.67														
2. Word recognition	0.10	−0.13													
3. Verbal fluency	0.20	0.09	0.80												
Linguistic speed															
4. Lexical decision	−0.11	−0.02	0.17	0.98											
5. Word recognition	0.11	0.11	0.12	−0.18	0.39										
Inhibition															
6. Flanker	−0.10	0.05	−0.07	−0.01	−0.01	0.40									
7. Simon	0.07	0.11	0.00	0.03	−0.29	0.00	0.81								
Monitoring															
8. Figural	0.25	0.02	0.32	0.00	−0.04	0.06	0.16	0.65							
9. Numerical	0.09	0.05	0.02	0.26	−0.17	−0.34	0.10	0.22	0.69						
Mixing															
10. Figural	−0.06	0.10	−0.12	−0.10	−0.11	0.13	0.08	0.07	0.13	0.96					
11. Numerical	−0.07	0.11	−0.07	0.06	0.08	0.00	−0.05	0.16	0.17	0.48	0.97				
Shifting															
12. Figural	0.01	−0.14	0.03	0.11	−0.03	−0.21	0.05	−0.14	0.01	−0.64	−0.50	0.90			
13. Numerical	0.04	0.02	−0.01	−0.09	−0.11	0.05	0.16	−0.09	0.02	−0.07	−0.64	0.39	0.93		
Working memory															
14. Figural	0.37	0.21	0.05	−0.04	−0.11	−0.08	0.05	0.31	0.31	0.25	0.29	−0.24	−0.19	0.81	
15. Numerical	0.08	−0.01	0.18	0.02	−0.05	−0.16	−0.06	0.16	0.32	0.22	0.25	−0.17	−0.15	0.38	0.89

Correlation coefficients printed in bold were significant ($p < 0.05$). Reliabilities are printed along the diagonal.

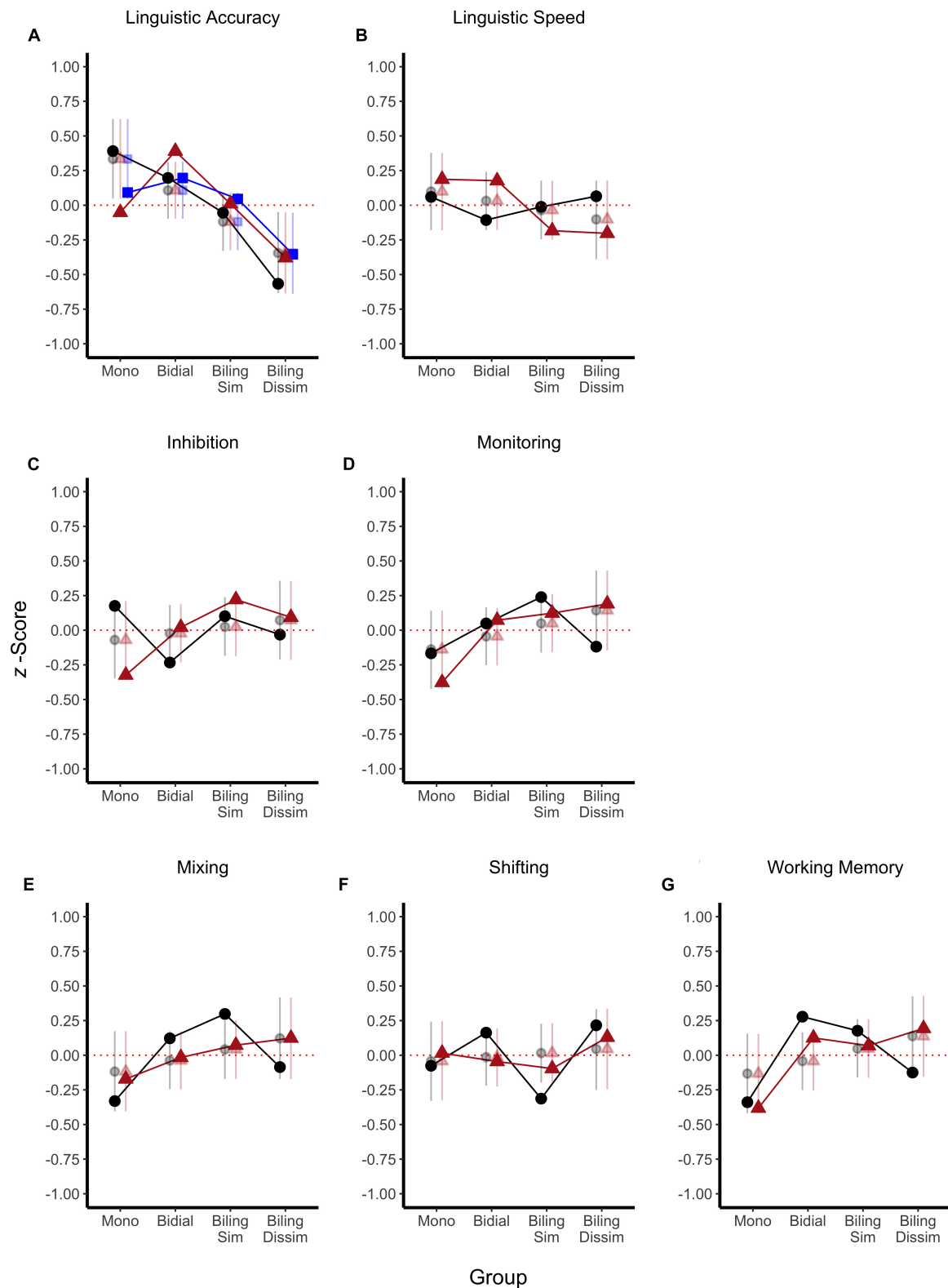


FIGURE 3 | Task scores (z-transformed; solid dots) and model predictions (transparent dots and error bars) as a function of language group. Predictions are means (dot) and 95% HDI (error bars) of the posterior of the linear contrast over language similarity. Each panel (A–G) shows a different ability. Mono, monolinguals; Bidial, bidialectals; Biling, bilinguals; Sim, similar; Dissim, dissimilar; Lexic, lexical decision; Rec, recognition; Fluency, verbal fluency.

TABLE 4 | Evidence for concreteness effects, inhibition, mixing and shifting costs.

Ability/Task	Trial type I		Trial type II		Difference (I-II)		Evidence	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>BF</i> ₁₀	% error
Linguistic	<i>abstract</i>		<i>concrete</i>		<i>concreteness</i>			
Word Rec. Acc	2.52	0.78	2.92	0.79	0.40	0.78	> 1,000	±0.00
Word Rec. Speed	6.73	0.15	6.70	0.15	0.03	0.06	752.90	±0.00
Inhibition	<i>neutral/congruent</i>		<i>incongruent</i>		<i>inhibition costs</i>			
Flanker	6.29	0.12	6.31	0.11	−0.01	0.05	4.29	±0.00
Simon	6.07	0.12	6.24	0.12	−0.18	0.07	> 1,000	±0.00
Mixing	<i>repetition (single)</i>		<i>repetition (mixed)</i>		<i>mixing costs</i>			
Figural	6.20	0.10	6.74	0.23	−0.54	0.20	> 1,000	±0.00
Numerical	6.30	0.11	6.59	0.25	−0.28	0.19	> 1,000	±0.00
Shifting	<i>repetition (mixed)</i>		<i>switch (mixed)</i>		<i>shifting costs</i>			
Figural	6.74	0.23	7.07	0.18	−0.31	0.19	> 1,000	±0.00
Numerical	6.59	0.25	6.94	0.21	−0.35	0.18	> 1,000	±0.00

M, mean; *SD*, standard deviation; *Word Rec.*, Word Recognition; *Acc*, accuracy. Values of trial type I and II are based on log-transformed reaction times (except for accuracies of the word recognition task), and are uncorrected for order-effects. The difference of trial type I and II reflects the concreteness effects, inhibition, mixing and switch costs, and is corrected for order-effects. Values printed in bold indicate at least substantial evidence for the alternative hypothesis ($BF_{10} \geq 3$).

of the parameter posterior distribution. The HDI reflects the range of credible values of the parameter given the data. **Figure 4** presents the posterior distribution of the slope of the linear contrast over language similarity for each cognitive ability. Panels in **Figure 4** show the mean (circle underneath the curve) and the 95% HDI (bar underneath the curve) of the slope, and the proportion of the HDI that is above and below 0. **Table 5** presents the BF_{10} for the effect of language similarity assessed by (a) the linear contrast over language group, (b) the sliding contrast comparing every two adjacent levels of the language similarity factor, as well as (c) the comparison of each group against the monolingual group (simple contrast).

Linguistic Ability

Linguistic accuracy (see **Figures 3A, 4A**) and speed (**Figures 3B, 4B**) decreased as language similarity decreased. This trend was only credibly different from 0 for accuracy though (see **Figure 4A**), which was also reflected by the evidence being substantial for the presence of a linear effect of language similarity on accuracy but not speed (see **Table 5**). At the level of pairwise group comparisons, however, the pattern was more nuanced: Comparison of monolinguals against bidialectals and similar bilinguals showed substantial evidence for the null, whereas the comparison of monolinguals against dissimilar bilinguals yielded substantial evidence for a bilingual cost in linguistic accuracy. Other comparisons yielded ambiguous evidence for or against differences. In linguistic speed, pairwise comparisons of adjacent groups yielded substantial evidence against group differences, but the comparison of the monolingual group against the bilingual groups tended to show more ambiguous evidence against differences in this measure (see **Table 5**). As the word recognition accuracy score was unreliable (-0.13) and uncorrelated to the other linguistic-accuracy scores, we re-ran the analyses without this task. There was still substantial evidence for the linear effect of language similarity in accuracy ($BF_{10} = 34.44 \pm 0.74\%$). Removing SES from the linguistic

processing analyses (see **Table 6**) did not change the pattern of results, except that the evidence against a linear effect on linguistic processing speed became substantial.

EF Measures

Inspection of **Figures 3C–G** indicate a weak linear trend for better EF performance as language similarity decreases. **Figures 4C–G** show that the posterior distributions of the slopes tended to be positive, with 65.9% (**Figure 4F**) to 91.3% (**Figure 4D**) of the posterior values being larger than 0. However, **Figures 4C–G** also show that 0 was within the 95% HDI of all 6 slopes, hence indicating that a null effect is credible given the data. As shown in **Table 5**, although weak if not ambiguous, evidence was in favor of the null hypothesis over the alternative for a linear effect of language similarity for all EFs. In line with the linear trend analyses, the comparisons of adjacent levels of the language similarity factor showed mostly ambiguous evidence for or against group differences (see **Table 5**). There was substantial evidence against monolinguals and bidialectals performing differently in inhibition and shifting measures, but bidialectals outperformed monolinguals in WM performance. Evidence was largely ambiguous regarding the performance differences between bidialectals and similar bilinguals, except for monitoring, for which there was substantial evidence against group differences. The comparisons of similar and dissimilar bilinguals yielded mostly weak evidence for the null hypothesis, with exception of monitoring for which the evidence was substantial for the null over the alternative hypothesis. Lastly, when contrasting monolinguals against similar and dissimilar bilinguals, the evidence was mostly ambiguous. For inhibition and shifting, the evidence was even substantially supporting no group differences when comparing extreme groups (i.e., monolingual vs. dissimilar bilinguals).

Given the low correlation between the flanker and Simon inhibition scores and the low reliability of the flanker inhibition score, we also ran the analyses for each task separately (and for

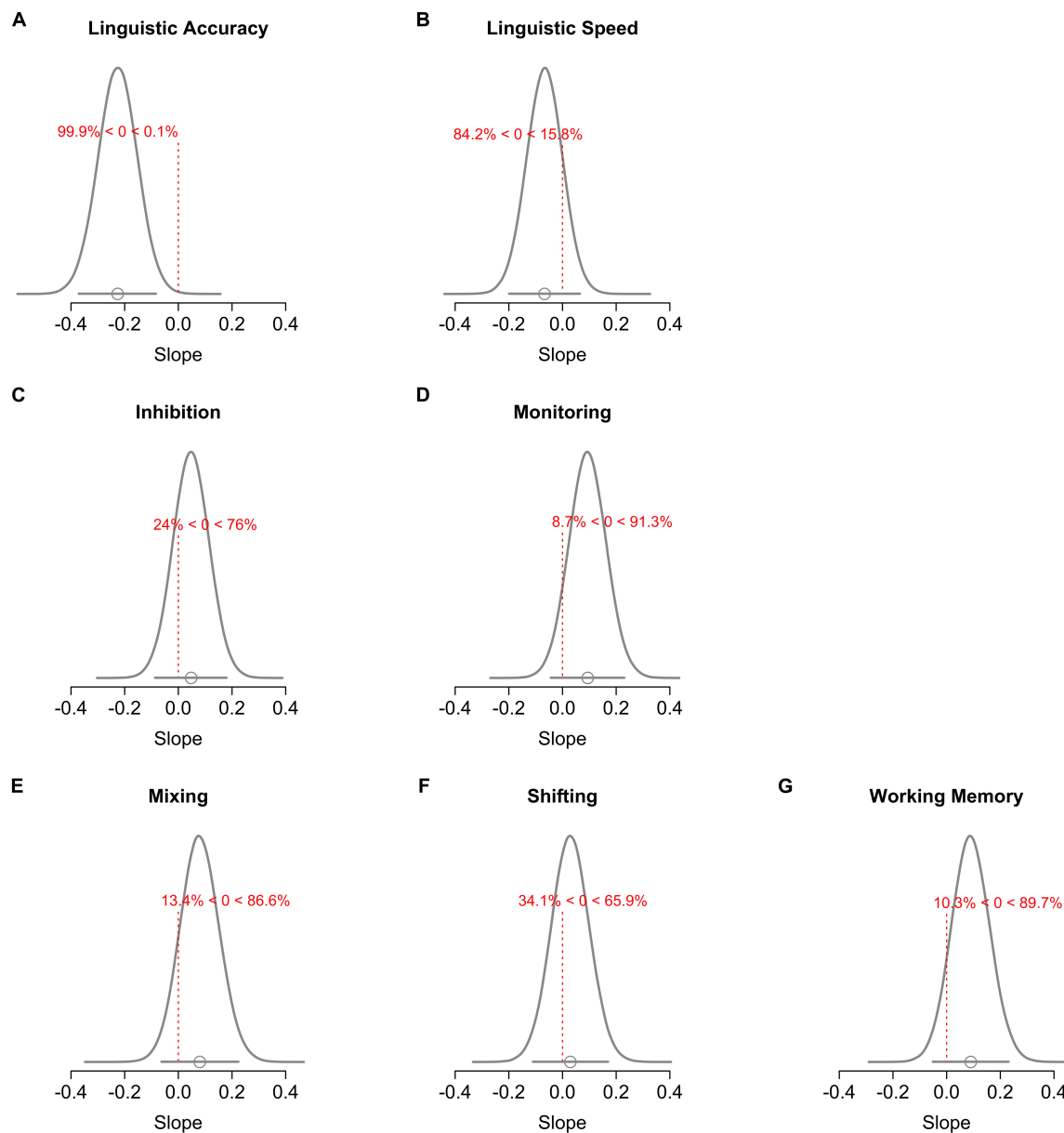


FIGURE 4 | Posterior distribution of the slope of the language similarity effect for each cognitive ability. Each panel (A–G) shows the mean (dot) and the 95% HDI (bar underneath the curve) of the slope, and the proportion of the HDI that is below or above 0 (which represents the null).

all other tasks as well; see **Supplementary Table S1**). For both tasks, the evidence remained overall ambiguous (BF_{10} between 0.53 and 1.88). We also ran all of the analyses reported here without entering SES (see **Table 6**). A similar pattern of results emerged, with the main difference being that inhibition and shifting yielded substantial evidence against an effect of language similarity (linear trend), and that the evidence for an effect of bidialectalism on WM was reduced to an ambiguous range ($BF_{10} = 2.27 \pm 0.88\%$). To rule out that SES drove the effect of bidialectalism on WM, we examined the main effect of SES and tested for an interaction between group (monolinguals vs. bidialectals) and SES: yielding evidence in the direction of the

absence of both a main effect of SES ($BF_{10} = 0.53 \pm 1.46\%$) and of an SES \times group interaction ($BF_{10} = 0.40 \pm 1.39\%$).

Language Experience: Continuous Predictors

As the group design was mainly aimed to assess differences in linguistic ability and EF functioning due to language similarity, it might not have adequately captured effects of other aspects of the bilingual experience (e.g., the effect of age of acquiring a second language, the proficiency, or the frequency of using it in a daily context). Therefore, we additionally ran the analyses reported above using continuous

TABLE 5 | Evidence (BF_{10}) for and against the effect of language similarity on each ability.

Ability	Sliding contrast								Simple contrast			
	Linear trend		Mono vs. Bidial		Bidial vs. Sim		Sim vs. Dissim		Mono vs. Sim		Mono vs. Dissim	
	BF_{10}	% error	BF_{10}	% error	BF_{10}	% error	BF_{10}	% error	BF_{10}	% error	BF_{10}	% error
Linguistic ability												
Linguistic accuracy	30.20	± 1.46	0.27	± 1.67	0.53	± 2.30	1.64	± 0.72	0.29	± 0.91	7.11	± 0.84
Linguistic speed	0.45	± 2.17	0.29	± 2.20	0.28	± 1.11	0.27	± 1.82	0.39	± 1.04	0.36	± 1.23
Executive functions												
Inhibition	0.35	± 1.41	0.27	± 1.73	0.49	± 0.96	0.34	± 0.96	0.35	± 1.44	0.27	± 2.63
Monitoring	0.75	± 0.91	1.07	± 2.78	0.26	± 1.88	0.33	± 0.74	1.83	± 1.10	0.43	± 1.19
Mixing	0.68	± 0.81	0.71	± 3.97	0.45	± 0.86	0.37	± 1.37	1.11	± 0.69	0.66	± 0.91
Shifting	0.37	± 0.98	0.32	± 0.94	0.72	± 1.56	1.02	± 1.20	0.34	± 1.24	0.32	± 0.99
Working memory	0.74	± 0.78	3.85	± 1.83	0.39	± 1.37	0.35	± 1.05	1.24	± 0.89	0.74	± 0.75

Mono, Monolinguals; Bidial, Bidialectals; Sim, Similar Bilinguals; Dissim, Dissimilar Bilinguals. All models included the average parents' education level as a proxy for socio-economic status as an additional predictor. Values printed in bold indicate at least substantial evidence for ($BF_{10} \geq 3$) or against ($BF_{10} \leq 0.33$) the alternative hypothesis.

TABLE 6 | Evidence (BF_{10}) for and against the effect of language similarity on each ability (not controlled for SES).

Ability	Sliding contrast								Simple contrast			
	Linear trend		Mono vs. Bidial		Bidial vs. Sim		Sim vs. Dissim		Mono vs. Sim		Mono vs. Dissim	
	BF_{10}	% error	BF_{10}	% error	BF_{10}	% error	BF_{10}	% error	BF_{10}	% error	BF_{10}	% error
Linguistic ability												
Linguistic accuracy	26.86	± 1.26	0.26	± 0.88	0.53	± 1.08	1.73	± 0.58	0.27	± 0.61	10.41	± 0.45
Linguistic speed	0.28	± 1.73	0.25	± 1.10	0.28	± 1.34	0.26	± 1.38	0.36	± 0.75	0.34	± 0.92
Executive functions												
Inhibition	0.25	± 1.00	0.25	± 0.64	0.50	± 0.53	0.31	± 0.63	0.37	± 3.02	0.26	± 0.57
Monitoring	0.52	± 0.78	0.80	± 0.63	0.30	± 2.01	0.32	± 0.61	1.21	± 0.66	0.59	± 0.75
Mixing	0.47	± 0.57	0.60	± 0.61	0.44	± 0.66	0.37	± 0.88	1.57	± 2.49	0.48	± 1.76
Shifting	0.25	± 0.61	0.33	± 3.66	0.52	± 2.26	0.90	± 0.72	0.38	± 1.75	0.38	± 0.60
Working memory	0.59	± 0.70	2.27	± 0.88	0.38	± 0.62	0.32	± 1.16	1.35	± 0.66	0.80	± 2.93

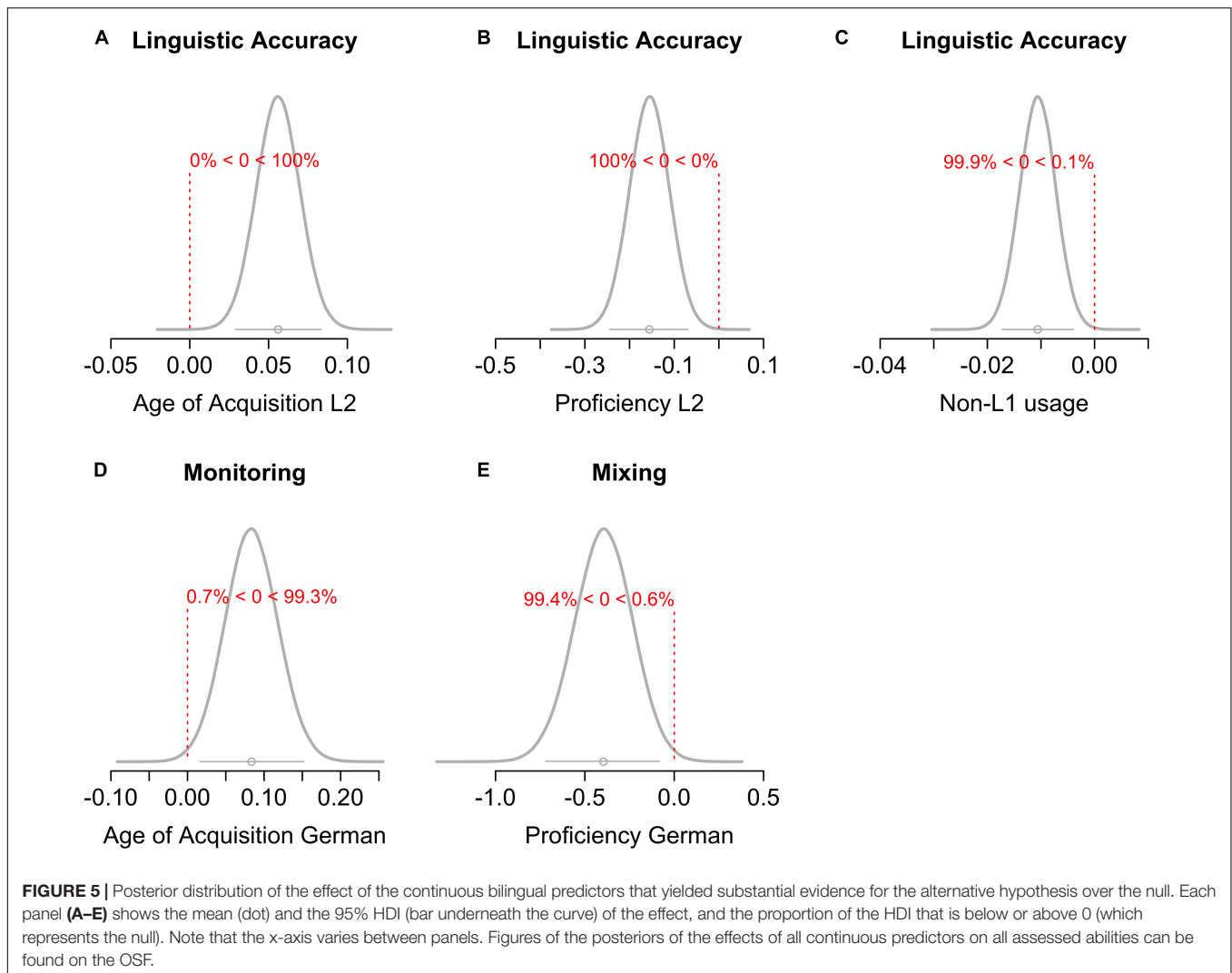
Mono, Monolinguals; Bidial, Bidialectals; Sim, Similar Bilinguals; Dissim, Dissimilar Bilinguals. Values printed in bold indicate at least substantial evidence for ($BF_{10} \geq 3$) or against ($BF_{10} \leq 0.33$) the alternative hypothesis.

measures of bilingualism as predictors instead of language group (see **Figure 1** for all continuous predictors and their distribution in the language groups). We ran a separate model for each predictor to avoid issues with multicollinearity. All models included SES as covariate. **Figure 5** shows the posterior distributions of the continuous predictors that yielded substantial effects. Figures with the posterior distributions of all continuous predictors for all abilities can be found on the OSF (**Supplementary Figures S1–S3**). The results largely reflected the findings reported for the group comparisons (see **Supplementary Table S2**). Substantial evidence for an effect of bilingualism emerged only for linguistic accuracy but not for linguistic speed. Specifically, a younger age of L2 acquisition ($BF_{10} = 784.22 \pm 1.43\%$) and higher L2 proficiency ($BF_{10} = 104.47 \pm 0.86\%$) was associated with lower linguistic accuracy (see **Figures 5A,B**). Moreover, a higher proportion of daily usage ($BF_{10} = 34.83 \pm 0.68\%$) of other languages besides German or the Swiss German dialect was associated with lower linguistic accuracy (**Figure 5C**). Notably,

however, the effect was very small ($M = -0.01$). For EF, substantial evidence was only present for an effect of age of acquisition of German on monitoring ($BF_{10} = 6.11 \pm 1.23\%$; see **Figure 5D**) and German proficiency on mixing ability ($BF_{10} = 8.24 \pm 0.94\%$; see **Figure 5E**). As can be seen in **Figures 5D,E**, these effects were in the direction of a monolingual disadvantage, such that younger age of learning German and higher German proficiency was associated with lower EF performance.

DISCUSSION

Theoretical claims about the effects of bilingualism on EF rest on the assumption that the heightened demands for language control in bilinguals require general executive control processes, thereby providing lifelong EF training. Our main goal was to examine the putative link between the difficulty in managing two languages (reflected by costs in linguistic performance)



and EF performance. For this purpose, we assessed several aspects of linguistic and EF processing in the same group of participants, and related their performance to their self-reported language experience. We classified participants in one of four groups with respects to language similarity (i.e., monolinguals, bidialectals, similar bilinguals, and dissimilar bilinguals) and tested for a linear effect of language similarity on linguistic and EF performance.

We predicted that language similarity should have opposite effects on linguistic and EF performance: language combinations that facilitate language control should yield comparatively better linguistic processing, but reduce the opportunities to train executive control, leading to limited EF advantages. We obtained evidence that language similarity was linearly related to linguistic accuracy, with similar languages yielding better performance than dissimilar languages. In line with the predictions, the estimated slope of the effect of language similarity on EF was indeed opposite to the slope of its effect on linguistic accuracy (i.e., better EF performance with more dissimilar languages). However, despite this trend, the

evidence was overall ambiguous and tended to support the null hypothesis, with one exception: we found substantial evidence for a positive association between bidialectism and WM.

Language Similarity, Linguistic Ability, and EF

The observed effect of language similarity on linguistic processing is consistent with several other reports in the literature (but see Costa et al., 2006). Previous studies have found that bilingual children with more similar languages (e.g., Spanish-English) outperform bilingual peers with less similar languages (e.g., Chinese-English) in linguistic tasks (Bialystok et al., 2003, 2005; Barac and Bialystok, 2012). This pattern also extends to young adults: in a sentence production task, dissimilar bilinguals (Mandarin-English) showed larger bilingual disadvantages, as measured by higher frequency effects, than similar bilinguals (Spanish-English; Runqvist et al., 2013). Taken together with our results, these findings support the notion that linguistic adaptation is more pronounced for

more similar languages. Furthermore, our results indicate that L2 can have an impact on L1, corroborating recent studies showing bidirectional cross-linguistic effects (e.g., Hohenstein et al., 2006; Brown and Gullberg, 2008, 2010, 2011; Ameer et al., 2009). This is in line with the assumption that both languages are activated simultaneously in the bilingual mind, leading to cross-interactions between languages. Our findings indicate that these adaptations facilitate language processing when languages are similar, thereby arguably reducing executive control demands. Hence, bilinguals speaking more dissimilar languages (which face the most challenging linguistic condition) should show larger EF benefits. However, the evidence was ambiguous (BF_{10} between 0.35 and 0.75 for a linear trend) for effects of language similarity on all of the EFs assessed.

Besides language similarity, the bilingual linguistic disadvantage found here might also reflect an effect of language usage (i.e., bilinguals generally use each of their languages less often than monolinguals use their L1). Support for this hypothesis comes from studies showing that lexical access in both L1 and L2 is delayed in bilinguals relative to monolinguals, specifically for less frequently used words (e.g., Gollan et al., 2008; Ivanova and Costa, 2008). However, in the present study, some monolinguals also indicated to use languages other than their L1 on a frequent basis. Furthermore, in a follow-up analysis examining the effect of continuous variables of the bilingual experience, we found that besides higher frequency of using non-L1 languages, higher proficiency and lower age of acquisition in L2 were also associated with lower linguistic accuracy across groups. Thus, taken together, frequency of language usage alone cannot entirely explain the group differences reported here.

Bidialectalism and Its Association With WM

Considering bidialectalism as an extreme case of language similarity is one novel feature of the present study. Recent studies in the field have suggested that bidialectalism may involve similar language control demands as bilingualism (Kirk et al., 2014; Antoniou et al., 2016) and may, thus, yield similar EF benefits. To our best knowledge, our study is the first to test how bidialectism affects both linguistic abilities and EFs in a sample of young adults. Our results showed evidence that bidialectism was associated with better WM performance than monolingualism. However, we did not observe additive effects of speaking an additional language (i.e., bidialectals did not differ from bilinguals), and the WM benefits did not generalize to any other of the assessed EFs.

One may argue that the presence of this effect solely for WM is in line with recent notions that, given the central role of executive control in WM (Baddeley, 1986; Cowan, 1995; Engle, 2002; Oberauer and Hein, 2012), any effects of speaking an L2 can be expected to be stronger for WM tasks than for any other EF tasks (Bialystok, 2017). Bidialectal effects on executive control may, hence, simply be too subtle to be

detected in non-WM tasks with less executive control demands. However, this explanation is contrary to findings from Miyake and Friedman (2012; see also Friedman et al., 2008) showing that the contribution of WM performance to a general executive control factor (i.e., “common EF”) is not particularly greater than the contribution made by shifting or inhibition performance to that factor.

Therefore, taken together with the effect on WM not being modulated by language similarity, the absence of an effect on other EFs could suggest that bidialectism does not practice common executive control as much as it does processes that are more specific to WM – for example the access and retrieval of currently relevant information. However, the present data do not allow for directly testing this proposition. Future studies specifically designed to disentangle effects of bidialectism and bilingualism on executive control from effects on retrieval of information may shed further light on the specific cognitive mechanisms affected by bilingualism.

Limitations

Although we found substantial evidence for a linear trend of language similarity on linguistic performance, evidence was mostly weak for the pairwise comparisons of adjacent levels of this factor. This is likely due to the fact that performance differences between two adjacent levels were too small to be distinguished from within-group variability by data from only about 24 participants per group. Therefore, future studies aiming at examining the effects of contrasting only two levels of language similarity on linguistic performance may need substantially larger group sizes.

Similarly, many of the effects on EF (linear trend and group contrasts) yielded evidence within the ambiguous range (i.e., BF_{10} between 0.34 and 1.83). The ambiguous results obtained here are in line with recent concerns that studies with small sample sizes provide ambiguous and unreliable evidence for effects of bilingualism on EF performance. Even with a total sample of 99 students and using a linear contrast approach, which is more powerful to detect experimental effects, we were unable to firmly reject or support the hypothesis that language similarity influences EF performance. In fact, for any EF considered here, the absolute slope of the language similarity effect was substantially smaller than that for linguistic accuracy, which indicate that the effects, if they were true, would be harder to detect with the small sample sizes common in the bilingual advantage literature.

Furthermore, we observed substantial effects of language similarity on linguistic accuracy but not speed. As effects in accuracy and speed can show trade-offs (Wickelgren, 1977), future studies may consider using sequential sampling models (such as the diffusion model; Ratcliff, 1978), which integrate information across these two measures to derive psychologically meaningful parameters. This may be a fruitful venue to examine how language experience affects different cognitive processes involved in decision-making.

Regarding the inclusion criteria of the present study, we chose age of acquisition (cutoff 7 years) and continuous active language usage as requirements for categorizing participants as bilingual, with the aim to most closely align our definition with previous studies testing for bilingual EF advantages (Bialystok et al., 2006; Costa et al., 2009; Hernández et al., 2010; Prior and MacWhinney, 2010; Luk et al., 2011; Gold et al., 2013). However, some studies have also found advantages in EF for bilinguals with a later age of acquisition (Pelham and Abrams, 2014; Vega-Mendoza et al., 2015). Thus, it is possible that factors of language experience other than the early acquisition of a language might bear more explanatory value for the proposed training effect on EF functioning (see below for a more detailed discussion of this topic).

Even though we paid close attention to match the groups and included random effects to take individual differences between participants and also between tasks of a measured ability into account (Coderre and van Heuven, 2014), we still faced the problem of group differences that were unrelated to bilingual status. These differences primarily affected the monolinguals. First, monolinguals differed from bidialectals and bilinguals regarding their immigrational background. Second, monolinguals in the present study may be considered as less monolingual than those in other studies reporting a bilingual advantage in EF (e.g., Luo et al., 2013) as they had, on average, learned at least three languages at a later point in their lives. Specifically, two monolinguals indicated to use a non-native language around half of the time or more. Additional analyses using continuous predictors of bilingualism revealed that higher non-L1 language usage was associated with lower linguistic accuracy. Therefore, if anything, excluding these individuals from the monolingual group would have resulted in an even stronger linguistic advantage of monolinguals over bidialectals and bilinguals. Although a monolingual group with less language exposure may have been desirable for the present study, this matches the Swiss (and European) demographic: learning two foreign languages is required by the Swiss educational system, and people in Switzerland speak on average two languages besides their native language (Werlen, 2009). Thus, we cannot rule out that a comparison to a strictly monolingual and non-immigrant sample would have led to stronger effects of bilingualism. However, in an increasingly globalized world the number of people speaking more than one language is rising, and monolinguals not exposed to other languages at all are rare (Grosjean, 2010). For example, estimates from survey data suggest that more than half of the European population are able to hold a conversation in at least one additional language besides their L1 (European Commission, 2012), and approximately 21% of the U. S. American population speak a language other than English at home (US Census Bureau, 2015). If dialects were also counted as separate languages, these percentages would rise even higher. Hence, effects for individuals without any L2 exposure at any time in their lives are not informative for the majority of the population. Moreover, as discussed previously (von Bastian et al., 2016), monolinguals without any

L2 exposure will likely differ from bilinguals in other aspects than just language exposure that would then be confounded with any effects of bilingualism. Thus, any differences observed for such extreme groups may disappear when accounting for the full range of individuals (e.g., see Unsworth et al., 2015).

Lastly, as the present study was a first attempt to investigate the effect of language similarity on the link between linguistic and EF performance, we chose to define the similarity of two languages based on their common ancestry (i.e., language family), a perhaps overly simplified classification. This definition resulted in a large heterogeneity in our group of participants with regards to the exact language combinations they had acquired. Future research using a more fine-grained operationalization of language similarity is required to identify the specific language properties (e.g., lexical, phonological, or grammatical overlap) underlying the effects on linguistic accuracy found in the present study. Deriving quantitative predictions based on the precise degree of overlap between languages would allow for testing specific hypotheses regarding which aspects of language overlap are relevant in yielding cross-linguistic interactions and potential knock-on effects on EF advantages.

Bilingualism Advantages: Challenges and Opportunities

As mentioned above, one limitation of this study is that we classified individuals as monolinguals or bilinguals based only on their age of acquisition and active usage of their L2. The choice of these criteria was based on previous literature at the time of designing this study. Arguably, however, it is possible that (an)other language background variable(s) would have been more predictive for linguistic and EF performance. For example, **Figure 1** illustrates that all participants learned some other language(s) at some point later in their life (panel C), and most participants, including some monolinguals, used other languages on a daily basis (panel E). Furthermore, considerable heterogeneity existed within language groups with regards to these factors. Thus, depending on the definition one has of bilingualism, the participants in the present study could be regrouped in different ways. This is one of the limitations of groups created based on observed variables (i.e., quasi-experimental designs).

Although typically not in the focus of bilingual advantage research, it is possible that, in previous studies relative to the present study, the number of languages learned and/or the non-L1 language usage (or any other potential indicator of bilingualism) were more closely aligned with age of acquisition and active usage – and, so, possibly the actual driving forces behind the bilingual advantages found. Indeed, assuming bilingual EF advantages exist in the first place, it is not unlikely that no single language background variable can explain bilingual advantages in their entirety but that (only) a certain combination of language experiences leads to EF advantages. However, it is yet unclear what combination of variables might be important to describe bilingualism, let alone what instruments are best suited to measure bilingualism (Surrain and Luk, 2017).

Moreover, bilingualism may not be a static trait but evolve dynamically over time depending on many other circumstances including necessity, context, and social and societal norms of language usage. One challenge is, therefore, to better capture the multidimensional and dynamic reality of bilingualism. To some extent, this development is already happening, with most studies using a multi-method approach combining self-reported measure (e.g., the Language Experience and Proficiency Questionnaire, LEAP-Q, Marian et al., 2007) with performance-based assessments (e.g., the Multilingual Naming Test, MINT, available in multiple languages, Gollan et al., 2012). Moreover, recent studies have adopted a multidimensional understanding of bilingualism (Luk and Bialystok, 2013) by using a range of different variables, such as proficiency and age of L2 acquisition, as continuous predictors (e.g., Paap et al., 2014; von Bastian et al., 2016). Similarly, in the present study, we ran a series of additional analyses using continuous indicators of bilingualism as predictors of linguistic and EF functioning, resulting in largely the same pattern of results as for the group design. Without clear theoretical predictions as to which language background variables (or combinations thereof) should relate to bilingual EF advantages, however, the selection and reporting of those variables will remain relatively unsystematic and, so, studies will be at risk of turning into analytical fishing expeditions.

Over and above the challenges of selecting, assessing, and analyzing dimensions of bilingualism, relatively vague theorizing poses an additional challenge. To meet this challenge in the present study, we tested the broad notion of bilingual EF advantages by examining one specific, theoretically derived, mechanism – the similarity of the two languages spoken – as a proxy of the demands of cross-language interference and its effect on both linguistic and EF performance. We did not find evidence to support the predicted relationship and, thus, our findings question the theoretical validity of the cross-language interference serving as the link between language experience and EF advantages.

As any single study, our findings are not definitive and require replication, and possibly theoretical and methodological refinement. In bilingual advantage research, replication poses a particular challenge due to the many sources of variation in measurement and sampling. Open Science practices, as followed in the present study, can support both theory development and replication attempts (e.g., Munafò et al., 2017). First, by providing our dataset alongside the analysis scripts, other researchers can directly test alternative analytical procedures and alternative hypotheses using our data. For example, the participants in this sample could be regrouped according to a different definition of bilingualism or of language similarity. Second, by providing open materials that can be used with open-source experimental software such as Tootool Web, other researchers can attempt an exact replication of our study with larger sample sizes or with a refined definition of language similarity. Third, pooling the present data set with data from such replication attempts will lead to more precise estimates of

the effects of language experience on linguistic and non-linguistic tasks.

CONCLUSION

We found that the similarity of the two languages spoken by bidialectals and bilinguals affects linguistic processing in a linear fashion, with performance worsening the more dissimilar the two languages are. However, the increased difficulty of managing two more dissimilar languages did not translate into substantial evidence for increased EF benefits. We contend that any fruitful future investigation in the field needs to test clear theoretical links between language demands and EFs, as advanced here.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the ‘Ethical Principles of Psychologists and Code of Conduct’ of the American Psychological Association (APA) and the ‘Ethische Richtlinien für Psychologinnen und Psychologen der Schweizerischen Gesellschaft für Psychologie (SGP) of the SGP with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved following the self-assessment checklist for ethical treatment of human participants provided by the Ethics Committee of the Faculty of Arts and Social Sciences of the University of Zurich.

AUTHOR CONTRIBUTIONS

The research reported in this paper was derived from two master theses submitted to the University of Zurich by JO under the supervision of ASS and from AS under the supervision of CvB. All authors contributed to the conception and design of the study. JO and AS collected the data. JO and ASS performed the statistical analysis. JO wrote the first draft of the manuscript. ASS and CvB wrote sections of the manuscript. All authors approved the submitted version.

FUNDING

This research was supported by a grant from the Suzanne and Hans Biäsch Foundation to ASS and CvB.

ACKNOWLEDGMENTS

During the work on this paper, JO was a pre-doctoral fellow of the International Max Planck Research School on the Life Course (LIFE, www.imprs-life.mpg.de; participating institutions: Max Planck Institute for Human Development, Freie Universität

Berlin, Humboldt-Universität zu Berlin, University of Michigan, University of Virginia, University of Zurich).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01997/full#supplementary-material>

FIGURE S1 | Posterior distribution of the effect of age of acquisition (AoA) of German, of the Swiss German dialect, and of the second language (L2). Each panel (A–G) shows these effects for each cognitive ability. The dot and the bar underneath the curve shows the mean and the 95% HDI of the posterior, respectively.

FIGURE S2 | Posterior distribution of the effect of proficiency (Prof.) in German, in the Swiss German dialect, and in the second language (L2). Each panel (A–G)

shows the posterior of these effects for each cognitive ability. The dot and the bar underneath the curve shows the mean and the 95% highest density interval (HDI) of the posterior, respectively.

FIGURE S3 | Posterior distribution of the effect of the number of languages learned (N. Lang.), the number of languages learned below age 7 (N. Lang. < 7), and the % of language usage other than the first language (Non-L1 Usage). Each panel (A–G) presents the posterior of these effects for each cognitive ability. The dot and the bar underneath the curve shows the mean (dot) and the 95% HDI of the posterior, respectively.

TABLE S1 | Evidence (BF_{10}) for (and against) the effect of language similarity on each task. Values printed in bold indicate at least substantial evidence for ($BF_{10} = 3$), and values printed in gray against ($BF_{10} = 0.33$) the alternative hypothesis.

TABLE S2 | Evidence (BF_{10}) for (and against) a linear effect of the continuous language demographic variables on each ability. Values printed in bold indicate at least substantial evidence for ($BF_{10} = 3$), and values printed in gray against ($BF_{10} = 0.33$) the alternative hypothesis.

REFERENCES

- Ameel, E., Malt, B. C., Storms, G., and Van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *J. Mem. Lang.* 60, 270–290. doi: 10.1016/j.jml.2008.10.001
- Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., et al. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Front. Psychol.* 5:398. doi: 10.3389/fpsyg.2014.00398
- Antoniou, K., Grohmann, K. K., Kambanaros, M., and Katsos, N. (2016). The effect of childhood bilingualism and multilingualism on executive control. *Cognition* 149, 18–30. doi: 10.1016/j.cognition.2015.12.002
- Arthur, W., and Day, D. V. (1994). Development of a short form for the raven advanced progressive matrices test. *Educ. Psychol.* 54, 394–403. doi: 10.1177/0013164494054002013
- Baddeley, A. D. (1986). *Working Memory*. New York, NY: Oxford University Press.
- Barac, R., and Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: role of language, cultural background, and education. *Child Dev.* 83, 413–422. doi: 10.1111/j.1467-8624.2011.01707.x
- Baum, S., and Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Appl. Psychol.* 35, 857–894. doi: 10.1017/S0142716414000174
- Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Biling. Lang. Cogniti.* 12, 3–11. doi: 10.1017/S1366728908003477
- Bialystok, E. (2017). The bilingual adaptation: how minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099
- Bialystok, E., Craik, F., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn. Mem. Cogniti.* 34, 859–873. doi: 10.1037/0278-7393.34.4.859
- Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290
- Bialystok, E., Craik, F. I. M., and Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends Cogniti. Sci.* 16, 240–250. doi: 10.1016/j.tics.2012.03.001
- Bialystok, E., Craik, F. I. M., and Ryan, J. (2006). Executive control in a modified antisaccade task: effects of aging and bilingualism. *J. Exp. Psychol. Learn. Mem. Cogniti.* 32, 1341–1354. doi: 10.1037/0278-7393.32.6.1341
- Bialystok, E., and Luk, G. (2012). Receptive vocabulary differences in monolingual and bilingual adults. *Biling. Lang. Cogniti.* 15, 397–401. doi: 10.1017/S136672891100040X
- Bialystok, E., Luk, G., and Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: interactions among languages and writing systems. *Sci. Stud. Read.* 9, 43–61. doi: 10.1207/s1532799xssr0901_4
- Bialystok, E., Luk, G., Peets, K. F., and Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Biling. Lang. Cogniti.* 13, 525–531. doi: 10.1017/S1366728909990423
- Bialystok, E., Majumder, S., and Martin, M. M. (2003). Developing phonological awareness: is there a bilingual advantage? *Appl. Psychol.* 24, 27–44. doi: 10.1017/S014271640300002X
- Blom, E., Küntay, A. C., Messer, M., Verhagen, J., and Leseman, P. (2014). The benefits of being bilingual: working memory in bilingual Turkish–Dutch children. *J. Exp. Child Psychol.* 128, 105–119. doi: 10.1016/j.jecp.2014.06.007
- Blumenfeld, H. K., Bobb, S. C., and Marian, V. (2016). The role of language proficiency, cognate status and word frequency in the assessment of Spanish–English bilinguals’ verbal fluency. *Int. J. Speech. Lang. Pathol.* 18, 190–201. doi: 10.3109/17549507.2015.1081288
- Brown, A., and Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1–L2 encoding of manner in speech and gesture: a study of Japanese speakers of English. *Stud. Second Lang. Acquis.* 30, 225–251. doi: 10.1017/S0272263108080327
- Brown, A., and Gullberg, M. (2010). Changes in encoding of path of motion in a first language during acquisition of a second language. *Cogniti. Linguist.* 21, 263–286. doi: 10.1515/COGL.2010.010
- Brown, A., and Gullberg, M. (2011). Bidirectional cross-linguistic influence in event conceptualization? Expressions of path among Japanese learners of English. *Biling. Lang. Cogniti.* 14, 79–94. doi: 10.1017/S1366728910000064
- Campbell, L. (2008). “The history of linguistics,” in *The Handbook of Linguistics*, eds M. Aronoff and J. Rees-Miller (Oxford: Oxford University Press), 81–104. doi: 10.1002/9780470756409.ch4
- Chambers, J. K., and Trudgill, P. (1998). *Dialectology*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511805103
- Coderre, E. L., and van Heuven, W. J. B. (2014). The effect of script similarity on executive control in bilinguals. *Front. Psychol.* 5:1070. doi: 10.3389/fpsyg.2014.01070
- Comrie, B. (2008). “Languages of the world,” in *The Handbook of Linguistics*, eds M. Aronoff and J. Rees-Miller (Oxford: Blackwell Publishing), 20–42. doi: 10.1002/9780470756409.ch2
- Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don’t. *Cognition* 113, 135–149. doi: 10.1016/j.cognition.2009.08.001
- Costa, A., Santesteban, M., and Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *J. Exp. Psychol. Learn. Mem. Cogniti.* 32, 1057–1074. doi: 10.1037/0278-7393.32.5.1057
- Cowan, N. (1995). *Attention and Memory: An Integrated Framework*. New York, NY: Oxford University Press.

- de Bruin, A., Bak, T. H., and Della Sala, S. (2015). Examining the effects of active versus inactive bilingualism on executive control in a carefully matched non-immigrant sample. *J. Mem. Lang.* 85, 15–26. doi: 10.1016/j.jml.2015.07.001
- Dijkstra, T., and van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Biling. Lang. Cognit.* 5, 175–197. doi: 10.1017/S1366728902003012
- Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., et al. (2014). The inhibitory advantage in bilingual children revisited. *Exp. Psychol.* 61, 234–251. doi: 10.1027/1618-3169/a000243
- Engle, R. W. (2002). Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11, 19–23. doi: 10.1111/1467-8721.00160
- European Commission (2012). *Europeans and Their Languages*. 1–145. Available at: http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_386_en.pdf
- Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., and Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *J. Exp. Psychol. Gen.* 137, 201–225. doi: 10.1037/0096-3445.137.2.201
- Gade, M., Souza, A. S., Druey, M. D., and Oberauer, K. (2017). Analogous selection processes in declarative and procedural working memory: N-2 list-repetition and task-repetition costs. *Mem. Cognit.* 45, 26–39. doi: 10.3758/s13421-016-0645-4
- Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Guasch, N. V., et al. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Front. Psychol.* 5:11. doi: 10.3389/fpsyg.2014.00011
- Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., and Smith, C. D. (2013). Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *J. Neurosci.* 33, 387–396. doi: 10.1523/JNEUROSCI.3837-12.2013
- Gollan, T. H., Montoya, R. I., Cera, C., and Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: aging, bilingualism, and the weaker links hypothesis. *J. Mem. Lang.* 58, 787–814. doi: 10.1016/j.jml.2007.07.001
- Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., and Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Mem. Cognit.* 33, 1220–1234. doi: 10.3758/bf03193224
- Gollan, T. H., Montoya, R. I., and Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology* 16, 562–576. doi: 10.1037/0894-4105.16.4.562
- Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., and Cera, C. M. (2012). Self-ratings of spoken language dominance: a multilingual naming test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. *Biling. Lang. Cognit.* 15, 594–615. doi: 10.1017/S1366728911000332
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cognit.* 1, 67–81. doi: 10.1017/S1366728998000133
- Grosjean, F. (2010). *Bilingual: Life and Reality*. Cambridge: Harvard University Press. doi: 10.4159/9780674056459
- Hernández, M., Costa, A., Fuentes, L. J., Vivas, A. B., and Sebastián-Gallés, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism* 13, 315–325. doi: 10.1017/S1366728909990010
- Hohenstein, J., Eisenberg, A., and Naigles, L. (2006). Is he floating across or crossing afloat? Cross-influence of L1 and L2 in Spanish-English bilingual adults. *Biling. Lang. Cognit.* 9, 249–261. doi: 10.1017/S1366728906002616
- Ishihara, S. (2003). *Ishihara's Tests for Colour-Blindness*. Tokyo: Kanehara.
- Ivanova, I., and Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychol. (Amst)*. 127, 277–288. doi: 10.1016/j.actpsy.2007.06.003
- Jäger, A. O., Süß, H. M., and Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test: BIS-Test, Form 4*. Göttingen: Hogrefe.
- Kirk, N. W., Fiala, L., Scott-Brown, K. C., and Kempe, V. (2014). No evidence for reduced Simon cost in elderly bilinguals and bidialectals. *J. Cognit. Psychol.* 26, 640–648. doi: 10.1080/20445911.2014.929580
- Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., and Taler, V. (2014). Executive function and bilingualism in young and older adults. *Front. Behav. Neurosci.* 8:250. doi: 10.3389/fnbeh.2014.00250
- Kroll, J. F., and Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *J. Cognit. Psychol.* 25, 497–514. doi: 10.1080/20445911.2013.799170
- Kroll, J. F., Bobb, S. C., and Hoshino, N. (2014). Two languages in mind: bilingualism as a tool to investigate language, cognition, and the brain. *Curr. Dir. Psychol. Sci.* 23, 159–163. doi: 10.1177/0963721414528511
- Kroll, J. F., Dussias, P. E., Bogulski, C. A., and Valdes Kroff, J. R. (2012). “Juggling two languages in one mind: what bilinguals tell us about language processing and its consequences for cognition,” in *Psychology of Learning and Motivation*, ed. B. Ross (San Diego, CA: Academic Press), 229–262. doi: 10.1016/B978-0-12-394393-4.00007-8
- Lehtonen, M., Hultén, A., Rodríguez-Fornells, A., Cunillera, T., Tuomainen, J., and Laine, M. (2012). Differences in word recognition between early bilinguals and monolinguals: behavioral and ERP evidence. *Neuropsychologia* 50, 1362–1371. doi: 10.1016/j.neuropsychologia.2012.02.021
- Leiner, D. J. (2018). *SoSci Survey*. Available at: <http://www.sosicisurvey.com>
- Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L. (2013). Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* 49, 764–766. doi: 10.1016/j.jesp.2013.03.013
- Li, P., Sepanski, S., and Zhao, X. (2006). Language history questionnaire: a web-based interface for bilingual research. *Behav. Res. Methods* 38, 202–210. doi: 10.3758/BF03192770
- Linck, J. A., Hoshino, N., and Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *Ment. Lex.* 3, 349–374. doi: 10.1075/ml.3.3.06lin
- Longobardi, G., Guardiano, C., Silvestri, G., Boattini, A., and Coelin, A. (2013). Toward a syntactic phylogeny of modern Indo-European languages. *J. Hist. Linguist.* 3, 122–152. doi: 10.1075/jhl.3.1.07lon
- Luk, G., and Bialystok, E. (2013). Bilingualism is not a categorical variable: interaction between language proficiency and usage. *J. Cognit. Psychol.* 25, 605–621. doi: 10.1080/20445911.2013.795574
- Luk, G., De Sa, E., and Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Biling. Lang. Cognit.* 14, 588–595. doi: 10.1017/S1366728911000010
- Luo, L., Craik, F. I. M., Moreno, S., and Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: evidence from aging. *Psychol. Aging* 28, 28–34. doi: 10.1037/a0030875
- Luo, L., Luk, G., and Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition* 114, 29–41. doi: 10.1016/j.cognition.2009.08.014
- Marian, V., Blumenfeld, H. K., and Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Res.* 50, 940–967. doi: 10.1044/1092-4388(2007)067
- Miyake, A., and Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Curr. Dir. Psychol. Sci.* 21, 8–14. doi: 10.1177/0963721411429458
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cognit. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734
- Morey, R. D., and Rouder, J. N. (2015). *BayesFactor: Computation of Bayes Factors for Common Designs (R package version 0.9.12-2)*. Available at: www.bayesfactorppl.r-project.org/
- Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., et al. (2017). A manifesto for reproducible science. *Nat. Hum. Behav.* 1, 1–9. doi: 10.1038/s41562-016-0021
- Oberauer, K. (2005). Control of the contents of working memory—A comparison of two paradigms and two age groups. *J. Exp. Psychol. Learn. Mem. Cognit.* 31, 714–728. doi: 10.1037/0278-7393.31.4.714
- Oberauer, K., and Hein, L. (2012). Attention to information in working memory. *Curr. Dir. Psychol. Sci.* 21, 164–169. doi: 10.1177/0963721412444727
- Oberauer, K., Souza, A. S., Druey, M. D., and Gade, M. (2013). Analogous mechanisms of selection and updating in declarative and procedural working memory: experiments and a computational model. *Cognit. Psychol.* 66, 157–211. doi: 10.1016/j.cogpsych.2012.11.001

- Oberauer, K., Süß, H.-M., Wilhelm, O., and Wittman, W. W. (2003). The multiple faces of working memory storage, processing, supervision, and coordination. *Intelligence* 31, 167–193. doi: 10.1016/S0160-2896(02)00115-0
- Paap, K. R. (2014). The role of componential analysis, categorical hypothesising, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *J. Cognit. Psychol.* 26, 242–255. doi: 10.1080/20445911.2014.891597
- Paap, K. R., Darrow, J., Dalibar, C., and Johnson, H. A. (2015). Effects of script similarity on bilingual advantages in executive control are likely to be negligible or null. *Front. Psychol.* 5:1539. doi: 10.3389/fpsyg.2014.01539
- Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognit. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Paap, K. R., Johnson, H. A., and Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *J. Cognit. Psychol.* 26, 615–639. doi: 10.1080/20445911.2014.944914
- Paap, K. R., and Liu, Y. (2014). Conflict resolution in sentence processing is the same for bilinguals and monolinguals: the role of confirmation bias in testing for bilingual advantages. *J. Neurolinguistics* 27, 50–74. doi: 10.1016/j.neuroling.2013.09.002
- Paap, K. R., Myuz, H. A., Anders, R. T., Bockelman, M. F., Mikulinsky, R., and Sawi, O. M. (2016). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *J. Cognit. Psychol.* 29, 89–112. doi: 10.1080/20445911.2016.1248436
- Pelham, S. D., and Abrams, L. (2014). Cognitive advantages and disadvantages in early and late bilinguals. *J. Exp. Psychol. Learn. Mem. Cognit.* 40, 313–325. doi: 10.1037/a0035224
- Portocarrero, J. S., Burright, R. G., and Donovan, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Arch. Clin. Neuropsychol.* 22, 415–422. doi: 10.1016/j.acn.2007.01.015
- Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Biling. Lang. Cognit.* 13, 253–262. doi: 10.1017/S1366728909990526
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Available at: <http://www.r-project.org/>
- Ransdell, S. E., and Fischler, I. (1987). Memory in a monolingual mode: when are bilinguals at a disadvantage? *J. Mem. Lang.* 26, 392–405. doi: 10.1016/0749-596X(87)90098-2
- Ransdell, S. E., and Fischler, I. (1989). Effects of concreteness and task context on recall of prose among bilingual and monolingual speakers. *J. Mem. Lang.* 28, 278–291. doi: 10.1016/0749-596X(89)90034-X
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59–108. doi: 10.1037/0033-295X.85.2.59
- Raven, J. C. (1990). *Advanced Progressive Matrices: Sets I, II*. Oxford: Oxford Psychologists Press.
- Ross, J., and Melinger, A. (2017). Bilingual advantage, bidialectal advantage or neither? Comparing performance across three tests of executive function in middle childhood. *Dev. Sci.* 20:e12405. doi: 10.1111/desc.12405
- Runqvist, E., Gollan, T. H., Costa, A., and Ferreira, V. S. (2013). A disadvantage in bilingual sentence production modulated by syntactic frequency and similarity across languages. *Cognition* 129, 256–263. doi: 10.1016/j.cognition.2013.07.008
- Salvatierra, J. L., and Rosselli, M. (2010). The effect of bilingualism and age on inhibitory control. *Int. J. Biling.* 15, 26–37. doi: 10.1177/1367006910371021
- Schwartz, A. I., Kroll, J. F., and Diaz, M. (2007). Reading words in Spanish and English: mapping orthography to phonology in two languages. *Lang. Cognit. Process.* 22, 106–129. doi: 10.1080/01690960500463920
- Schwibbe, M., Röder, K., Schwibbe, G., Borchardt, M., and Geiken-Pophanken, G. (1981). Zum emotionalen Gehalt von Substantiven, Adjektiven und Verben [The emotional contents of nouns, adjectives, and verbs]. *Z. Exp. Angew. Psychol.* 28, 486–501.
- Shipstead, Z., Redick, T. S., and Engle, R. W. (2012). Is working memory training effective? *Psychol. Bull.* 138, 628–654. doi: 10.1037/a0027473
- Simons, G. F., and Fennig, C. D. (2018). *Ethnologue: Languages of the World*, 21st Edn. Dallas, TX: Texas SIL Int.
- Singh, N., and Mishra, R. K. (2013). Second language proficiency modulates conflict-monitoring in an oculomotor Stroop task: evidence from Hindi-English bilinguals. *Front. Psychol.* 4:322. doi: 10.3389/fpsyg.2013.00322
- Soveri, A., Rodríguez-Fornells, A., and Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a within group analysis approach. *Front. Psychol.* 2:183. doi: 10.3389/fpsyg.2011.00183
- Surraín, S., and Luk, G. (2017). Describing bilinguals: a systematic review of labels and descriptions used in the literature between 2005–2015. *Biling. Lang. Cognit.* 1–15. doi: 10.1017/S1366728917000682
- Unsworth, N., Redick, T. S., McMillan, B. D., Hambrick, D. Z., Kane, M. J., and Engle, R. W. (2015). Is playing video games related to cognitive abilities? *Psychol. Sci.* 26, 759–774. doi: 10.1177/0956797615570367
- US Census Bureau (2015). *Detailed Languages Spoken at Home and Ability to Speak English for the Population 5 years and Over: 2009-2013*. Available at: <https://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html>
- Valian, V. (2015). Bilingualism and cognition. *Biling. Lang. Cognit.* 18, 3–24. doi: 10.1017/S1366728914000522
- Van Assche, E., Drieghe, D., Duyck, W., Welsaert, M., and Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *J. Mem. Lang.* 64, 88–107. doi: 10.1016/j.jml.2010.08.006
- Vega-Mendoza, M., West, H., Sorace, A., and Bak, T. H. (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition* 137, 40–46. doi: 10.1016/j.cognition.2014.12.008
- Verreyt, N., Woumans, E., Vandelanotte, D., Szmalec, A., and Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Biling. Lang. Cognit.* 19, 181–190. doi: 10.1017/S1366728914000352
- von Bastian, C. C., Locher, A., and Ruffin, M. (2013). Tatool: a Java-based open-source programming framework for psychological studies. *Behav. Res. Methods* 45, 108–115. doi: 10.3758/s13428-012-0224-y
- von Bastian, C. C., and Oberauer, K. (2013). Distinct transfer effects of training different facets of working memory capacity. *J. Mem. Lang.* 69, 36–58. doi: 10.1016/j.jml.2013.02.002
- von Bastian, C. C., Souza, A. S., and Gade, M. (2016). No evidence for bilingual cognitive advantages: a test of four hypotheses. *J. Exp. Psychol. Gen.* 145, 246–258. doi: 10.1037/xge0000120.supp
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., and van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: the case of psi: comment on bem (2011). *J. Pers. Soc. Psychol.* 100, 426–432. doi: 10.1037/a0022790
- Werlen, I. (2009). *Verbinden Oder Trennen? Schweizer Mehrsprachigkeit im Widerspruch Zwischen Einheit und Vielfalt. Figurationen*. Köln: Böhlau, 103–125. doi: 10.7788/figurationen.2009.10.12.103
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol. (Amst.)* 41, 67–85. doi: 10.1016/0001-6918(77)90012-9
- Wiseheart, M., Viswanathan, M., and Bialystok, E. (2016). Flexibility in task switching by monolinguals and bilinguals. *Biling. Lang. Cognit.* 19, 141–146. doi: 10.1017/S1366728914000273
- Woumans, E., Ceuleers, E., Van der Linden, L., Szmalec, A., and Duyck, W. (2015). Verbal and nonverbal cognitive control in bilinguals and interpreters. *J. Exp. Psychol. Learn. Mem. Cognit.* 41, 1579–1586. doi: 10.1037/xlm0000107
- Yang, H., Hartanto, A., and Yang, S. (2016). The complex nature of bilinguals' language usage modulates task-switching outcomes. *Front. Psychol.* 7:560. doi: 10.3389/fpsyg.2016.00560
- Yow, W. Q., and Li, X. (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: why variations in bilingual experiences matter. *Front. Psychol.* 6:164. doi: 10.3389/fpsyg.2015.00164

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Oschwald, Schättin, von Bastian and Souza. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Impact of Language Experience on Attention to Faces in Infancy: Evidence From Unimodal and Bimodal Bilingual Infants

Evelyne Mercure^{1*}, Isabel Quiroz², Laura Goldberg¹, Harriet Bowden-Howl^{1,3}, Kimberley Coulson^{1,4}, Teodora Gliga², Roberto Filippi⁵, Peter Bright⁶, Mark H. Johnson^{2,7} and Mairéad MacSweeney¹

¹ Institute of Cognitive Neuroscience, University College London, London, United Kingdom, ² Centre for Brain and Cognitive Development, Birkbeck, University of London, London, United Kingdom, ³ School of Psychology, University of Plymouth, Plymouth, United Kingdom, ⁴ Department of Psychology and Sports Sciences, University of Hertfordshire, Hatfield, United Kingdom, ⁵ Institute of Education, University College London, London, United Kingdom, ⁶ Department of Psychology, Anglia Ruskin University, Cambridge, United Kingdom, ⁷ Department of Psychology, University of Cambridge, Cambridge, United Kingdom

OPEN ACCESS

Edited by:

Francesca Marina Bosco,
Università degli Studi di Torino, Italy

Reviewed by:

Arturo Hernandez,
University of Houston, United States
Valerio Santangelo,
University of Perugia, Italy

*Correspondence:

Evelyne Mercure
e.mercure@ucl.ac.uk

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 29 June 2018

Accepted: 20 September 2018

Published: 16 October 2018

Citation:

Mercure E, Quiroz I, Goldberg L, Bowden-Howl H, Coulson K, Gliga T, Filippi R, Bright P, Johnson MH and MacSweeney M (2018) Impact of Language Experience on Attention to Faces in Infancy: Evidence From Unimodal and Bimodal Bilingual Infants. *Front. Psychol.* 9:1943. doi: 10.3389/fpsyg.2018.01943

Faces capture and maintain infants' attention more than other visual stimuli. The present study addresses the impact of early language experience on attention to faces in infancy. It was hypothesized that infants learning two spoken languages (unimodal bilinguals) and hearing infants of Deaf mothers learning British Sign Language and spoken English (bimodal bilinguals) would show enhanced attention to faces compared to monolinguals. The comparison between unimodal and bimodal bilinguals allowed differentiation of the effects of learning two languages, from the effects of increased visual communication in hearing infants of Deaf mothers. Data are presented for two independent samples of infants: Sample 1 included 49 infants between 7 and 10 months (26 monolinguals and 23 unimodal bilinguals), and Sample 2 included 87 infants between 4 and 8 months (32 monolinguals, 25 unimodal bilinguals, and 30 bimodal bilingual infants with a Deaf mother). Eye-tracking was used to analyze infants' visual scanning of complex arrays including a face and four other stimulus categories. Infants from 4 to 10 months (all groups combined) directed their attention to faces faster than to non-face stimuli (i.e., attention capture), directed more fixations to, and looked longer at faces than non-face stimuli (i.e., attention maintenance). Unimodal bilinguals demonstrated increased attention capture and attention maintenance by faces compared to monolinguals. Contrary to predictions, bimodal bilinguals did not differ from monolinguals in attention capture and maintenance by face stimuli. These results are discussed in relation to the language experience of each group and the close association between face processing and language development in social communication.

Keywords: infants, bilingualism, Deaf, sign language, face processing, eye-tracking, bimodal bilingualism, visual attention

INTRODUCTION

From the first days of life, infants attend preferentially to faces and face-like stimuli (Johnson et al., 1991; Valenza et al., 1996; Farroni et al., 2005). These early biases in attention to faces are likely to maximize face experience and social interactions from the very beginning of postnatal life, allowing infants to rapidly develop complex face processing skills.

In older infants, faces continue to capture and maintain attention more than other visual stimuli. Indeed, it has been observed that 6-month-olds direct their first saccade to faces more often than predicted by chance in a complex array comprising a face and multiple visual objects. Increased attention capture by faces compared to objects was also observed in the same experimental design in 7- and 14-month-olds (Elsabbagh et al., 2013). However, increased attention capture by face stimuli was not observed in a similar, but black and white, experimental design in 3- and 6-month-olds (Di Giorgio et al., 2012) or in a color presentation of a face and a toy in 4-to-8-month-olds (DeNicola et al., 2013). Faces are also scanned more extensively than other visual stimuli, attracting a larger number of fixations and increased looking time in 6-month-old infants (Gliga et al., 2009; Di Giorgio et al., 2012) and in 4-to-8-month-olds (DeNicola et al., 2013), but not in 3-month-old infants (Di Giorgio et al., 2012). An increase in attention to faces between 3 and 9 months was observed in a more natural setting where infants watched a cartoon animation (Frank et al., 2009). Interestingly faces had a greater tendency to capture and sustain attention in infants at-risk for autism than infants at low-risk for autism, irrespective of whether these infants were later diagnosed with autism or not (Elsabbagh et al., 2013). Moreover, longer looking time at face stimuli at 7 months was associated with poorer performance in face recognition in 3-year-old infants at-risk for autism (de Klerk et al., 2014). These results are contrary to the idea that autism evolves from an initial lack of attention or interest in social stimuli early in life, but rather suggest complex interactions between social and attentional mechanisms in early development. The level of attention to faces reflects the infant's interest and processing needs, and higher attention may sometimes associate with processing difficulties.

Although face processing and language acquisition have been traditionally studied in parallel, a few previous studies have suggested that early bilingualism may impact face processing mechanisms in infancy. Different face scanning patterns have been observed for monolingual and bilingual infants when presented with talking faces (Lewkowicz and Hansen-Tift, 2012; Pons et al., 2015). At 4 months, bilinguals show increased attention to the mouth compared to monolinguals. While monolinguals show a preference for looking at the eyes of a talking face, bilinguals show no preference for the mouth or eyes at that age. A strong preference for looking at the mouth of talking faces later develops in monolinguals and bilinguals, and can be observed in both groups at 8 months. At 12 months, monolinguals show preferential looking to the mouth for faces talking in a non-native language, while no preference for the eyes or mouth is observed for the native language. In contrast, 12-month-old bilinguals show a preference for the mouth of faces talking in both native and non-native languages. Moreover, bilingual 8-month-olds are better than monolingual infants of the same age at distinguishing two different languages when silently articulated (Weikum et al., 2007; Sebastián-Gallés et al., 2012), and bilingual infants from 4 to 8 months tend to spend more time looking at talking faces than monolinguals (Mercure et al., 2018). Increased attention to the mouth was also observed

for faces displaying non-linguistic emotional movements in 8-month-old bilinguals compared to monolinguals (Ayneto and Sebastian-Galles, 2017), suggesting that bilingualism influences face scanning patterns beyond the context of speech processing. In adulthood, early bilinguals may not demonstrate the classic "other race effect" (Kandel et al., 2016) that is robustly observed in monolinguals (Meissner and Brigham, 2001). These results suggest an impact of early bilingualism on face scanning and face processing.

Unimodal bilinguals acquire two or more spoken languages simultaneously. In other words, these infants acquire two linguistic codes (two sets of sounds, two lexicons, two sets of grammatical rules) and learn to keep them apart, while experiencing a reduced exposure to each of these codes compared to monolinguals (Werker, 2012; Costa and Sebastián-Gallés, 2014). Even though this process is extremely complex, bilingual infants usually reach the milestones of early language development at the same age as monolinguals, including canonical babbling, first word production, and first word combinations (Werker, 2012; Costa and Sebastián-Gallés, 2014). This complex process appears to be made possible by some adaptations in speech and language processing including an increased sensitivity to visual articulation (Sebastián-Gallés et al., 2012) and an increased visual attention to the mouth of talking faces (Pons et al., 2015). Bilingual infants may develop a strategy of orienting to faces faster than monolinguals and scanning them more extensively than monolinguals, which would allow them to make optimal use of articulation cues potentially displayed by these faces. This strategy appears to generalize to contexts in which no speech is present, such as for faces displaying non-linguistic emotional movements (Ayneto and Sebastian-Galles, 2017). This study tests the hypothesis that, compared to monolingual infants, bilingual infants exposed to two spoken languages from birth will demonstrate increased attention capture and attention maintenance for faces in the absence of speech or movement. Attention to faces has never been studied for static faces in bilingual infants. This would suggest that early language experience can impact on attention allocation mechanisms for social stimuli, even in the absence of speech and movement.

A second group of interest in the current study were hearing infants with Deaf mothers. These infants are likely to differ in attention to faces as a result of differences in language and communicative experience. If a Deaf mother uses a sign language, such as British Sign Language (BSL) as her preferred mode of communication, her infant is likely to experience two languages in different modalities. These infants are exposed to a signed language processed mainly in the visual modality (e.g., BSL), and a spoken language processed mainly in the auditory modality (e.g., spoken English). For this reason, they are often referred to as "bimodal bilinguals," as opposed to "unimodal bilinguals" who are exposed to two spoken languages. Bimodal bilinguals achieve the early linguistic milestones in each of their languages at the same time as children learning two spoken languages (Petitto et al., 2001; Hofmann and Chilla, 2015). Like unimodal bilinguals, bimodal bilinguals may achieve this more complex task by increasing their attention to faces. Congruent with this idea, using

eye-tracking, we have previously reported that bimodal bilingual infants spend longer looking at talking faces than monolingual infants (Mercure et al., 2018). Moreover, because infants with Deaf mothers often experience visual forms of communication, visual attention is key to their communicative experience with their mother and other Deaf people in their environment. Sign language communication requires visual attention to the signer and attention to the face appears to be crucial. When presented with sign language, 4- and 14-month-old infants with and without experience of sign language share their visual attention between the signer's face and hands, but generally spend longer looking at the face than the hands area (Palmer et al., 2012). Similarly, adult signers focus the largest proportion of their visual attention to the face, and not the hands, when perceiving sign language communication (Muir and Richardson, 2005; De Filippo and Lansing, 2006; Emmorey et al., 2008). This increased attention to the face during sign language communication is the hypothesized mechanism for an observed enhancement of certain aspects of face processing in Deaf and hearing signers compared to non-signers (Bettger et al., 1997; McCullough and Emmorey, 1997; Emmorey, 2001; Stoll et al., 2017). Due to the crucial importance of visual attention for sign language communication, Deaf mothers have been observed to use various strategies to obtain visual attention from their child, such as moving in their child's existing focus of attention (Woll and Kyle, 1989). These patterns of interaction with the mother and with other Deaf communication partners may lead to increased visual attention to the mother, and especially her face (Palmer et al., 2012), in infants of Deaf mothers. Whether their particular experience of communication in the visual modality has an impact on their attention to static faces has never been studied before. We hypothesize that, because of the increased complexity of learning two languages and the increased importance of visual attention in their communication with their Deaf mother, bimodal bilingual infants would demonstrate enhanced attention capture and maintenance for static faces compared to monolinguals and possibly greater than unimodal bilinguals.

The deployment of selective attention in adulthood is not only influenced by perceptual properties of the object (e.g., luminance, contrast, movement), but also by strategies, rewards, and the significance that objects have gained through experience (Chelazzi et al., 2013). Since language experience influences the significance of the face cues in social communication, it is also likely to influence attention to faces. The present study addresses this question by comparing three groups of infants with different language experience. The comparison of two groups of bilinguals – unimodal and bimodal bilinguals – allows distinguishing effects that are caused by learning two languages, from those that are linked to bimodal bilinguals' unique experience of communication in the visual modality. Visual scanning of complex arrays was studied in two independent samples of infants, following an existing experimental protocol (Gliga et al., 2009; Elsabbagh et al., 2013). Sample 1 compared monolinguals and unimodal bilinguals between 7 and 10 months. Sample 2 compared three groups of 4-to-8-month-old infants with different language experience: monolinguals, unimodal bilinguals, and bimodal bilinguals. It was hypothesized

that, compared to monolingual infants, unimodal and bimodal bilinguals would show enhanced attention capture and attention maintenance by faces when they are presented within a complex visual array. It was also predicted that bimodal bilinguals may show this effect to a greater degree than unimodal bilinguals.

MATERIALS AND METHODS

Participants

Sample 1

A total of 49 hearing infants between 7 and 10 months contributed data. A further seven infants participated in the study but were excluded due to equipment malfunction or failure to calibrate ($n = 6$), or experimenter error ($n = 1$). Infants were from two groups with different language experience: 26 monolingual infants with hearing parents (16 girls, mean age = 8.7 months), 23 unimodal bilingual infants with hearing parents (6 girls, mean age = 8.4 months). Age did not differ significantly between groups [$F(1) = 2.0$; $p = 0.159$; $\eta^2 = 0.042$]. Monolingual infants were only exposed to English. Both parents were hearing and only used one language. Unimodal bilinguals were frequently and regularly exposed to English and one or more additional spoken language(s). The combination of languages varied between infants. Exposure to each language was estimated by using an English adaptation (Byers-Heinlein, 2009) of the language exposure questionnaire designed by Bosch and Sebastián-Gallés (1997). Unimodal bilinguals were exposed to English on average 52% of the time (standard deviation = 24).

Sample 2

A total of 88 hearing infants between 4 and 8 months contributed data. A further seven infants participated in the study but were excluded due to equipment malfunction or failure to calibrate ($n = 3$), withdrawal ($n = 1$), or failure to reach looking time criteria ($n = 3$; see section "Data Analyses"). Infants were from three groups with different language experience: 32 monolingual infants with hearing parents (16 girls, mean age = 6.2 months), 25 unimodal bilingual infants with hearing parents (eight girls, mean age = 6.2 months), and 31 bimodal bilingual infants with a Deaf mother (18 girls; mean age = 6.4 months). Age did not differ between groups [$F(2) = 0.354$; $p = 0.703$; $\eta^2 = 0.008$]. Monolingual infants were only exposed to English. Both parents were hearing and only used one language. Unimodal bilinguals were frequently and regularly exposed to English and one or more additional spoken language(s). The combination of languages varied between infants. All infants in this group had a hearing bilingual/multilingual mother. 18 unimodal bilingual infants also had a bilingual/multilingual father, and seven had a monolingual father. None reported hearing deficits in any immediate family members. Unimodal bilinguals were exposed to English on average 46% of the time (standard deviation = 23; Byers-Heinlein, 2009). Bimodal bilinguals were frequently and regularly exposed to BSL and English. All infants in this group had a Deaf mother using BSL as her preferred mode of communication; 27 bimodal bilinguals also had a second severely/profoundly D/deaf parent, three had a second parent who was hearing or had mild hearing

loss, and one had a single Deaf mother. Bimodal bilinguals were exposed to English on average 40% of the time (standard deviation = 21; Byers-Heinlein, 2009). There was no difference in language exposure to English between the two groups of bilinguals ($p = 0.311$).

Infants with hearing parents (Sample 1 and 2) were contacted from the Birkbeck Babylab database of volunteers recruited from advertisements at mum-and-baby groups, parenting websites and publications. Bimodal bilinguals (Sample 2) were recruited through social media and websites specifically aimed at the Deaf community. Most infants were born at term (37–42 weeks gestation), except for five infants born slightly before term (34–36 weeks) (four monolinguals and one unimodal bilingual: for these infants, a corrected age was used). Parents reported no hearing problems (except for one infant's mother reporting glue ear) or vision problems (except for one infant's mother reporting a suspected squint), and no serious mental or physical conditions (except for one infant who had undergone heart surgery). Deaf families were geographically spread across the whole of Great Britain, while infants with hearing parents came mostly from London and surrounding areas. Travel expenses were reimbursed, and a baby t-shirt and certificate of participation were offered to families. This study was carried out in accordance with the recommendations of UCL and Birkbeck Research Ethics Committees. All parents gave written informed consent prior to participation, after explanations of the study in English or BSL depending on the parents' preferred mode of communication by fluent members of the research team. The protocol was approved by the UCL and Birkbeck Research Ethics Committees and conforms to the Declaration of Helsinki.

Procedure

Infants from Sample 1 were invited to participate in a larger *Bilingual Babies* research protocol, which began with three eye-tracking tasks presented in TobiiStudio (the “attention to faces” task reported here, as well as tasks investigating audiovisual speech perception and eye gaze perception), followed by seven short eye-tracking tasks on a different experimental set up. The whole protocol usually required between 1 and 1.5 h per infant, including resting, napping, and feeding time. Infants from Sample 2 were invited to participate in the larger *Speak and Sign* research protocol, including a functional near infrared spectroscopy task (investigating brain activation in response to infant-directed spoken and sign language), the same three eye-tracking tasks on TobiiStudio described for Sample 1 and behavioral measures (the Mullen Scales of Early Learning and videos of parent–child interaction). The whole protocol usually required between 1.5 and 3 h per infant, including resting, napping, and feeding time. Only data from the “attention to faces” task are reported in the present article. The stimuli and procedures for this task were identical for both samples.

During the “attention to faces” task, infants sat on their parent's lap in a dimly lit room about 60 cm away from a TobiiT120 eye-tracker (17-in diameter, screen refresh rate 60 Hz, ET sampling rate of 60 Hz, spatial accuracy < 1°). Infant gaze position was calibrated with colorful animations using a five-point routine. Each infant's gaze and behavior was monitored

throughout the study via camera and Tobii Studio LiveViewer. The experimenter occasionally shook a rattle behind the screen to attract the infant's attention.

Stimuli

Five different slides were presented for 10 s each (Gliga et al., 2009; Elsabbagh et al., 2013). In each slide, five color images belonging to five object categories were presented: faces, phase-scrambled faces, birds, cars, and phones (see **Figure 1**). Each individual image was presented only once and the position of each category in the slide was randomized. Images were all of comparable size and presented at an equal distance from the center of the screen. When viewed from a 55 cm distance, the images had an eccentricity of 9.3° and covered an area of approximately 5.2° × 7.3°. Differences in color and luminosity were minimized. Visual saliency (the sensory prominence of an object compared to its background) has been observed to influence visual attention selection mechanisms in adults (Santangelo, 2015), children (Cavallina et al., 2018), and infants (Althaus and Mareschal, 2012). The stimulus categories used in the present study did not differ in terms of visual saliency (Elsabbagh et al., 2013). Faces all had direct gaze and happy expression. There were three female faces and two male faces of different ethnic origins. Scrambled faces were created from each face by randomizing the phase spectra while maintaining the original outer face contour, with the amplitude and color spectra remaining constant. These “attention to faces” slides were interleaved with blocks from other studies.

Data Analysis

Data were excluded in trials where infants looked at the entire slide for less than 1 s. Only infants with at least three good trials were included in the analyses. These criteria are identical to the ones used by Elsabbagh et al. (2013). Five rectangular regions of interest corresponding to the five categories of objects on each slide were defined in Tobii Studio. Three measures were extracted



FIGURE 1 | Sample stimulus slides. Face was obscured for publication purposes only.

for each category of objects and averaged for all included trials: fixation latency (the time difference between the beginning of the trial and the beginning of the first fixation to each region of interest), fixation count (the number of fixations within each region of interest), and total fixation duration (the total time spent fixating within each region of interest during the trial period of 10 s). As we did not have any specific hypotheses regarding group differences in attention to birds, cars, phones, and scrambled faces stimuli, the measures for all these stimuli were averaged to create a Non-Face stimulus category. However, any significant Face vs. Non-Face effect was followed by planned comparisons for individual contrasts between Face and each stimulus category to clarify the stability of the effect across control conditions.

RESULTS

Fixation Latency

Sample 1

The latency between the beginning of each trial and the beginning of the first fixation to Face and Non-Face stimuli was analyzed with a stimulus (2) \times group (2) ANOVA (see **Figure 2**). A significant effect of stimulus was found [$F(1,47) = 86.1$; $p < 0.001$; $\eta^2 = 0.647$], with Faces attracting infants' attention faster than other stimulus categories. Individual comparisons of Faces to each stimulus category (birds, cars, phones, and scrambled faces) revealed highly significant effects (all $p < 0.001$). There were no interaction of stimulus \times group, but a borderline group effect [$F(1) = 3.2$; $p = 0.058$; $\eta^2 = 0.075$], suggested that unimodal bilinguals tended to orient to both stimulus categories faster than monolinguals. The group effect was significant for Face stimuli [$F(1) = 4.2$; $p = 0.045$; $\eta^2 = 0.083$], but not for Non-Face stimuli [$F(1) = 0.4$; $p = 0.542$; $\eta^2 = 0.008$].

Sample 2

Fixation latency was analyzed with a stimulus (2) \times group (3) ANOVA (see **Figure 2**). A significant effect of stimulus was found [$F(1,85) = 124.7$; $p < 0.001$; $\eta^2 = 0.595$], with infants orienting faster to Face than to Non-Face stimuli [individual contrasts all $p < 0.001$]. There were no main effect of group [$F(2) = 1.1$; $p = 0.342$; $\eta^2 = 0.025$] or interaction of group \times stimulus [$F(2,85) = 0.6$; $p = 0.572$; $\eta^2 = 0.013$]. Group effects were not significant on Face fixation latency [$F(2) = 1.6$; $p = 0.204$; $\eta^2 = 0.037$; *post hoc t*-tests: monolinguals vs. unimodal bilinguals: $p = 0.391$; monolinguals vs. bimodal bilinguals: $p > 0.999$; unimodal vs. bimodal bilinguals: $p = 0.312$].

Pooled Analyses

Data from monolinguals and unimodal bilinguals of both studies were pooled together and Face fixation latencies were analyzed in a group (2) \times sample (2) ANOVA (see **Figure 2**). Bimodal bilinguals were excluded as they were only present in Sample 2. There was a significant group effect [$F(1) = 6.2$; $p = 0.014$; $\eta^2 = 0.057$]. Overall unimodal bilinguals oriented to faces faster than monolinguals. There were no effect of sample [$F(1) = 0.1$; $p = 0.741$; $\eta^2 = 0.001$], and no interaction of sample \times group

[$F(1) = 0.2$; $p = 0.664$; $\eta^2 = 0.002$]. The same ANOVA for Non-Face fixations revealed no group effect [$F(1) = 0.6$; $p = 0.440$; $\eta^2 = 0.006$], no sample effect [$F(1) = 2.1$; $p = 0.145$; $\eta^2 = 0.021$] and no interaction of group \times sample [$F(1) < 0.1$; $p = 0.964$; $\eta^2 < 0.001$].

Fixation Count

Sample 1

The number of fixations that infants directed to Faces and Non-Faces was analyzed in a stimulus (2) \times group (2) ANOVA (see **Figure 3**). A significant effect of stimulus was found [$F(1,47) = 188.2$; $p < 0.001$; $\eta^2 = 0.800$]. Faces attracted more fixations than any of the other object categories (all $p < 0.001$). There were no main effect of group [$F(1) = 0.6$; $p = 0.443$; $\eta^2 = 0.013$] or stimulus \times group interaction [$F(1,47) = 0.2$; $p = 0.634$; $\eta^2 = 0.005$]. A main effect of group was not significant when only Face stimuli were considered [$F(1) = 0.4$; $p = 0.511$; $\eta^2 = 0.009$].

Sample 2

The number of fixations was analyzed in a stimulus (2) \times group (3) ANOVA (see **Figure 3**). There was a significant effect of stimulus [$F(1,85) = 235.9$; $p < 0.001$; $\eta^2 = 0.735$] with faces attracting more fixations than any of the other objects (all individual contrasts $p < 0.001$). There was a significant interaction of stimulus \times group [$F(2,85) = 3.4$; $p = 0.037$; $\eta^2 = 0.075$], but no significant main effect of group [$F(2) = 3.0$; $p = 0.055$; $\eta^2 = 0.066$]. Unimodal bilinguals tended to direct more fixations to faces than the other groups [group effect on Face fixation: $F(2) = 3.4$; $p = 0.038$; $\eta^2 = 0.074$; *post hoc t*-tests: monolinguals vs. unimodal bilinguals: $p = 0.075$; monolinguals vs. bimodal bilinguals: $p > 0.999$; unimodal vs. bimodal bilinguals: $p = 0.067$]. Groups did not differ in terms of fixation to Non-Face stimuli [$F(2) = 0.2$; $p = 0.699$; $\eta^2 = 0.008$].

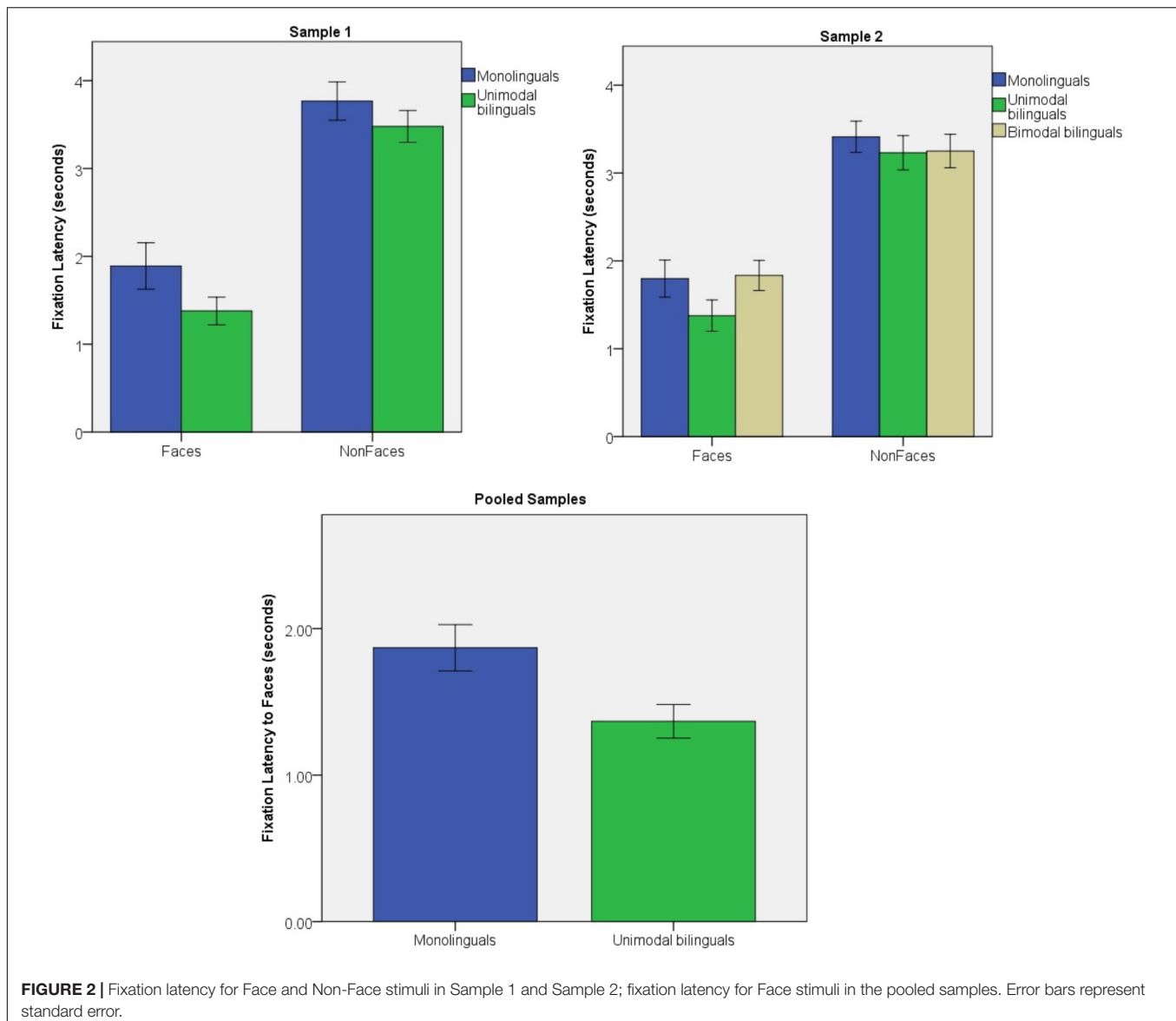
Pooled Samples

After excluding bimodal bilinguals, the number of Face fixations was analyzed for monolinguals and unimodal bilinguals in a group (2) \times sample (2) ANOVA (see **Figure 3**). Unimodal bilinguals directed significantly more fixations to Faces than monolinguals [$F(1) = 4.7$; $p = 0.032$; $\eta^2 = 0.044$]. There were no effect of sample [$F(1) = 0.2$; $p = 0.636$; $\eta^2 = 0.002$] and no interaction of sample \times group [$F(1) = 1.7$; $p = 0.201$; $\eta^2 = 0.016$].

Total Fixation Duration

Sample 1

The total amount of time fixating Faces and Non-Faces over the whole trial was analyzed in a stimulus (2) \times group (2) ANOVA (see **Figure 4**). A significant effect of stimulus was found [$F(1,47) = 135.6$; $p < 0.001$; $\eta^2 = 0.743$]. Infants looked at faces for longer than any of the other object categories (all $p < 0.001$). There were no main effect of group [$F(1) = 0.4$; $p = 0.513$; $\eta^2 = 0.009$] or stimulus \times group interaction [$F(1, 47) = 0.2$; $p = 0.622$; $\eta^2 = 0.005$]. The main effect of group was not significant when only Face stimuli were considered [$F(1) = 0.3$; $p = 0.562$; $\eta^2 = 0.007$].



Sample 2

Total fixation duration was analyzed in a stimulus (2) \times group (3) ANOVA (see **Figure 4**). There was a significant effect of stimulus [$F(1,85) = 283.2$; $p < 0.001$; $\eta^2 = 0.769$] with faces being fixated for longer than any of the other objects (all individual contrasts $p < 0.001$). The main effect of group [$F(2) = 1.8$; $p = 0.175$; $\eta^2 = 0.040$], and the stimulus \times group interaction [$F(2,85) = 2.2$; $p = 0.113$; $\eta^2 = 0.050$] was not significant. Group effects were not significant on Face fixation duration [$F(2) = 2.1$; $p = 0.131$; $\eta^2 = 0.047$; *post hoc t*-tests: monolinguals vs. unimodal bilinguals: $p = 0.304$; monolinguals vs. bimodal bilinguals: $p > 0.999$; unimodal vs. bimodal bilinguals: $p = 0.178$].

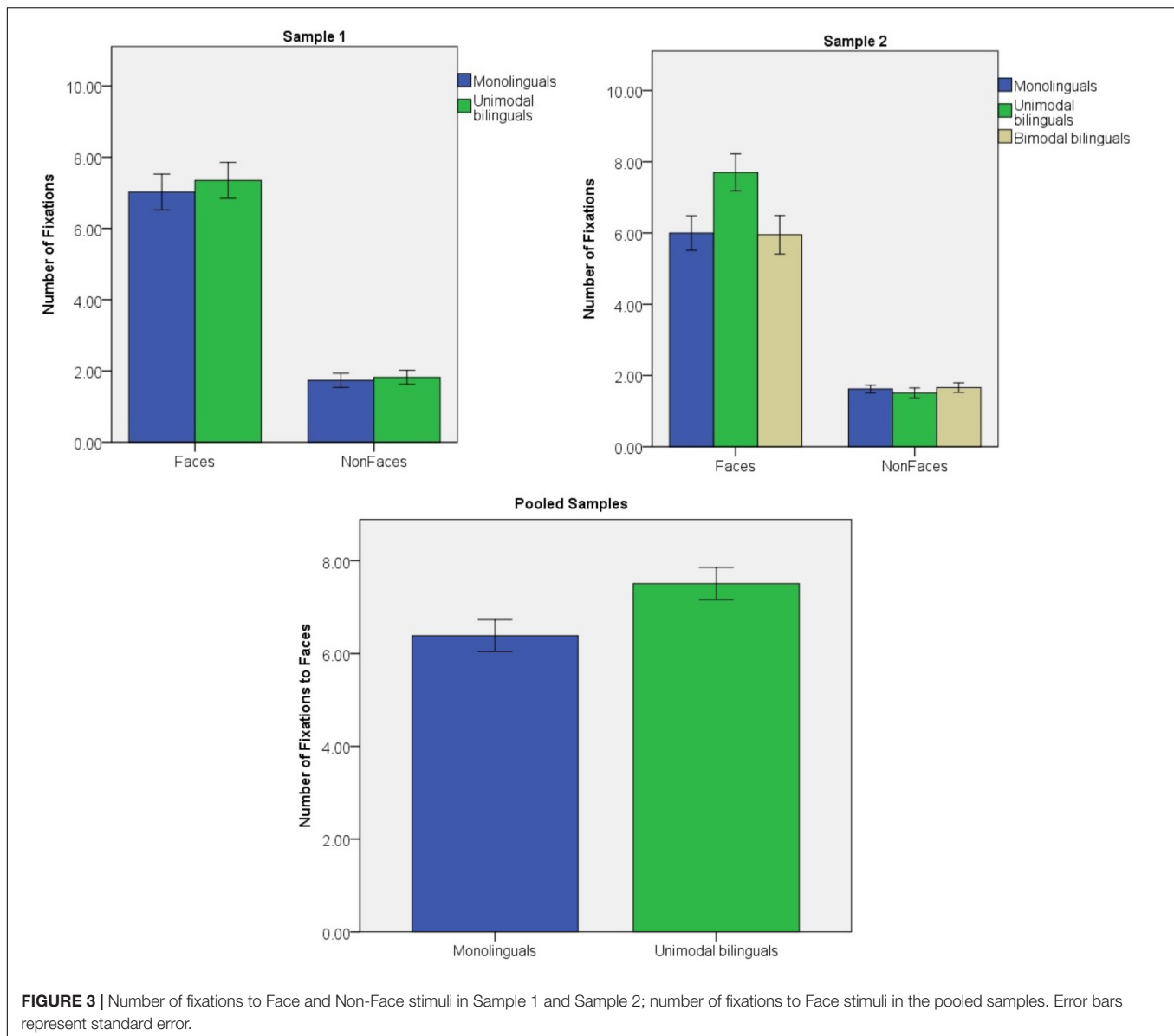
Pooled Samples

After excluding bimodal bilinguals, total fixation duration to Face stimuli were analyzed for monolinguals and unimodal bilinguals in a group (2) \times sample (2) ANOVA (see **Figure 4**).

Unimodal bilinguals tended to spend longer looking at faces than monolinguals, but this difference was not significant [$F(1) = 3.9$; $p = 0.136$; $\eta^2 = 0.022$]. There were no effect of sample [$F(1) = 0.2$; $p = 0.647$; $\eta^2 = 0.002$] and no interaction of sample \times group [$F(1) = 0.4$; $p = 0.522$; $\eta^2 = 0.004$].

DISCUSSION

The present study assessed the influence of early language experience on the development of attention to faces in infancy. Previous literature suggests that faces capture and/or maintain infants' visual attention more than other stimuli (Gliga et al., 2009; Di Giorgio et al., 2012; DeNicola et al., 2013; Elsabbagh et al., 2013). The present findings are consistent with this literature. Indeed, it was observed that infants from 4 to 10 months orient to faces in a complex visual array faster than they

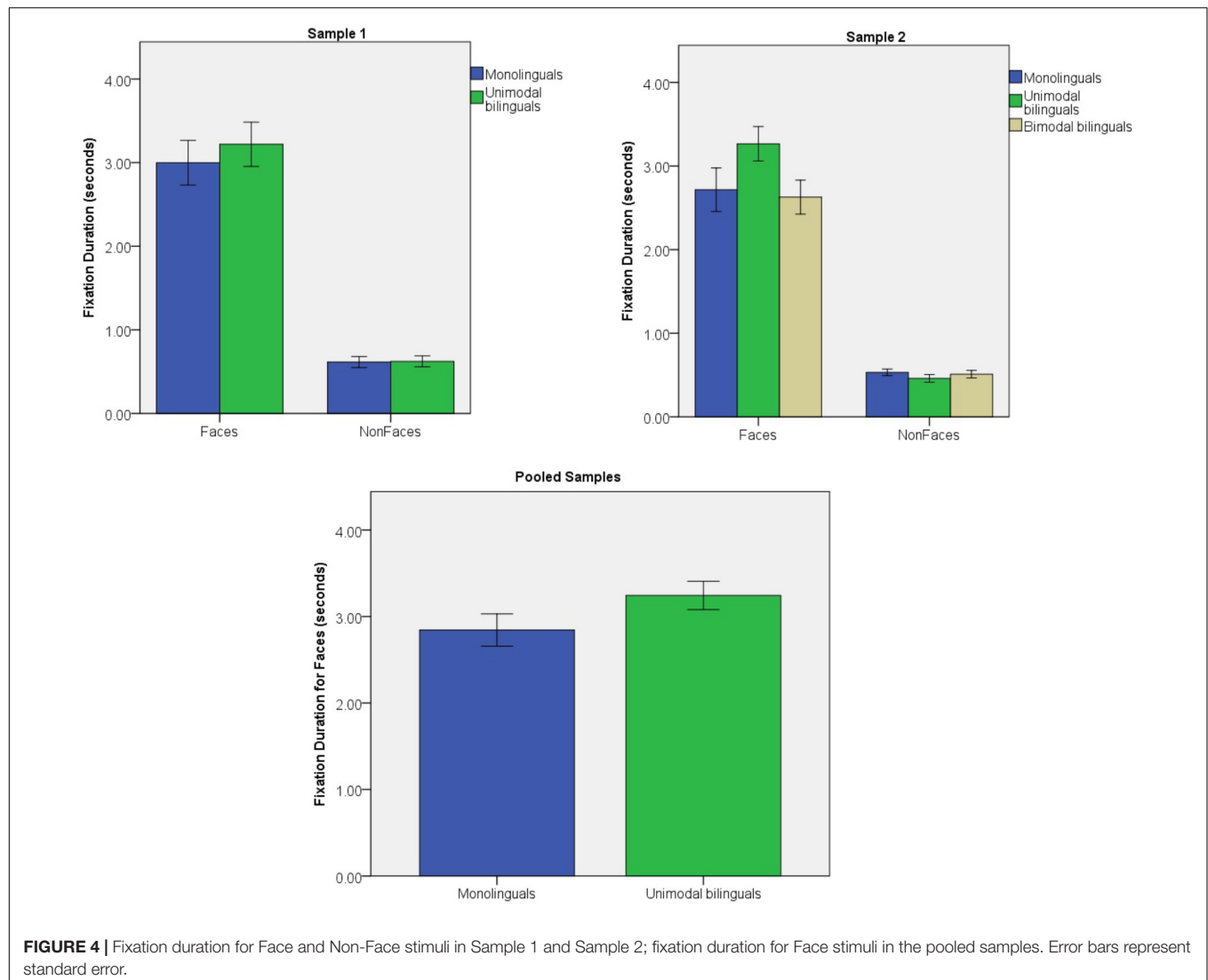


orient to objects or abstract patterns. Infants also directed more fixations at faces than other visual stimuli. They also fixated faces for longer than other visual stimuli.

It was predicted that unimodal bilingual infants would demonstrate increased attention capture and maintenance by face stimuli compared to monolinguals. Consistent with this hypothesis, it was observed that unimodal bilinguals between 4 and 10 months were generally faster at orienting to faces compared to monolinguals. This effect was significant in the pooled samples, and appeared to be more reliable in older infants as it approached significance in Sample 1 (7-to-10-month-olds), but not in Sample 2 (4-to-8-month-olds). Unimodal bilinguals also directed more fixations to faces than monolingual infants of the same age. These effects appeared to be more reliable in younger infants, as it was significant in the younger sample, but not the older sample alone. However, no reliable group

differences could be observed in the amount of time infants spent fixating faces. Taken together, these results suggest that unimodal bilinguals direct their attention to faces faster than monolinguals (especially older infants) and that they scan faces more extensively than monolinguals (especially younger infants).

The second hypothesis was that bimodal bilingual infants with Deaf mothers would demonstrate increased attention capture and maintenance by face stimuli compared to monolinguals and potentially unimodal bilinguals. Like unimodal bilinguals, bimodal bilinguals learn two languages, but unlike unimodal bilinguals, they also have a unique experience of communication in the visual modality with their Deaf mother and potentially other Deaf communication partners. Visual attention is key to communication between a Deaf mother and her infant and it has been observed that Deaf mothers deploy strategies to obtain visual attention from their infants (Woll and Kyle, 1989).



It was hypothesized that these strategies would act to maximize attention to faces in bimodal bilingual infants, and predicted that increased attention capture and maintenance by face stimuli would be apparent even when presented with static faces. Contrary to hypothesis, bimodal bilinguals did not differ from

monolinguals in terms of attention to faces. Bimodal bilinguals oriented to faces faster than to objects and they scanned faces more extensively than objects. However, the magnitude of these effects did not differ from those of monolinguals. It was previously observed that bimodal bilinguals demonstrated

TABLE 1 | *p*-value of the group differences for Face stimuli for each measure and each experimental sample.

	Monolinguals vs. unimodal bilinguals	Monolinguals vs. bimodal bilinguals	Unimodal vs. bimodal bilinguals
Fixation latency	Sample 1: $p = 0.045^*$ Sample 2: $p = 0.391$ Pooled samples: $p = 0.014^*$	$p > 0.999$	$p = 0.312$
Fixation count	Sample 1: $p = 0.511$ Sample 2: $p = 0.075$ Pooled samples: $p = 0.032^*$	$p > 0.999$	$p = 0.067$
Fixation duration	Sample 1: $p = 0.562$ Sample 2: $p = 0.304$ Pooled samples: $p = 0.136$	$p > 0.999$	$p = 0.178$

increased looking time to talking faces in comparison to monolinguals (Mercure et al., 2018). However, the present results suggest that this effect does not translate to static faces within a complex array. Unlike unimodal bilinguals, bimodal bilinguals do not have to differentiate two spoken languages or learn two systems of speech sounds. The languages that they learn use different sensory modalities and are therefore more easily discriminable. When presented with an unfamiliar static face, bimodal bilinguals do not know whether this is the face of someone using spoken or sign language. The face, but also the hands can be used as cues to discriminate between these languages. For this reason, the presence of a static face without spoken or sign language production may not lead to increased attention to faces in bimodal bilinguals as it does for unimodal bilinguals. However, if the face begins to produce speech, increased attention to the face would be observed in bimodal bilinguals as a strategy to process a language modality in which the infant has less experience than monolinguals (Mercure et al., 2018).

It is important to note that the first hypothesis was tested on data pooled from two independent samples of infants that were collected at two different time points, with a total of 58 monolingual and 48 unimodal bilingual infants. In contrast, the second hypothesis was tested on a single sample of infants collected at one time point, including 32 monolingual and 31 bimodal bilingual infants. Differences in visual attention to faces between monolinguals and unimodal bilinguals were more reliable in the pooled samples than in either of the individual samples. This suggests that there was individual variability in these measures and that analyses benefited from increased sample sizes. Nevertheless, for each of the measures that showed significant group differences in the pooled samples (fixation duration and fixation count), a significant or borderline effect was also observed on one of the individual samples, with a smaller sample size. Due to difficulty in recruiting bimodal bilinguals, it was not possible to recruit a second sample from this special population. However, inspection of **Figures 2–4** suggests that attention capture and attention maintenance was highly similar in monolingual and bimodal bilingual infants. Moreover, the *p*-values of the pairwise contrasts between monolingual and bimodal bilingual infants on each of the three measures taken on Face stimuli were larger than 0.999 (see **Table 1**). It is therefore unlikely that significant group differences between monolinguals and bimodal bilinguals would be present in a larger sample size.

In adulthood, it has been observed that visual search performance is influenced by rewards associated with each target (Kristjánsson et al., 2010), and selective attention is greatly influenced by the significance that objects have gained through experience (Chelazzi et al., 2013). Unimodal bilingual infants learn that visual cues of articulation are useful to distinguish spoken languages. This is reflected in their increased attention to the mouth of talking faces (Pons et al., 2015), and their increased ability at distinguishing languages based on visual articulation (Sebastián-Gallés et al., 2012). The current data suggest that a unimodal bilingual experience in infancy may reinforce an increased allocation of visual attention to faces, and that this effect could generalize to still faces. Increased visual attention to still

faces would allow unimodal bilinguals to take advantage of visual cues of articulation to discriminate different spoken languages if the face was to begin producing speech. It was observed in the present study that unimodal bilinguals scanned faces more extensively than monolinguals and this effect was more pronounced in the younger infant sample (4–8 months). In older infants, this strategy might be modified to orient faster to faces and to engage in extensive scanning for moving/talking faces, but not for still faces. This more sophisticated strategy would allow them to take advantage of visual cues of articulation of talking faces, but would free the infant's attention to explore other stimuli in the case that articulation cues are not available (for example, in still faces). Congruent with this idea, a faster orientation to faces was observed in unimodal bilinguals compared to monolinguals, and was more pronounced in the older infant sample (7–10 months).

This study demonstrates an impact of language experience on the early development of attention to faces in infancy. The increased complexity of learning two *spoken* languages was found to increase attention capture and maintenance for still faces. These visual strategies may be adaptive to maximize the use of potential visual cues of articulation to allow the discrimination of two spoken languages. Bimodal bilingualism and the experience of communication in the visual modality with a Deaf mother do not appear to impact attention to unfamiliar still faces. Increased attention to faces for bimodal bilinguals compared to monolinguals may be restricted to talking faces in this group (Mercure et al., 2018). Our data suggest that there are complex interactions in the development of face processing and language learning in the context of social communication in infancy.

AUTHOR CONTRIBUTIONS

The original idea was conceived by EM, with input from MM, MJ, RF, PB, and TG. The task was designed by TG. Sample 1 was recruited and tested by IQ. Sample 2 was recruited and tested by EM, LG, HB-H, and KC. The data were analyzed by EM with advice from TG, MM, PB, RF, and MJ. The manuscript was written by EM with input from RF, PB, TG, MJ, and MM.

FUNDING

This study was funded by an ESRC Future Research Leader fellowship to EM (ES/K001329/1) and by a British Academy/Leverhulme grant to RF and PB (SG162171). MJ was supported by the UK Medical Research Council (G0701484), and MM by a Wellcome Trust Fellowship (100229/Z/12/Z).

ACKNOWLEDGMENTS

The authors would like to thank all the parents and carers who contributed to this study, Laura Pirazzoli and Catherine Weston for their help with data collection, Prof. Bencie Woll for helpful comments on this dataset, and Prof. Annette Karmiloff-Smith who inspired our research.

REFERENCES

- Althaus, N., and Mareschal, D. (2012). Using saliency maps to separate competing processes in infant visual cognition. *Child Dev.* 83, 1122–1128. doi: 10.1111/j.1467-8624.2012.01766.x
- Ayneto, A., and Sebastian-Galles, N. (2017). The influence of bilingualism on the preference for the mouth region of dynamic faces. *Dev. Sci.* 20:e12446. doi: 10.1111/desc.12446
- Bettger, J. G., Emmorey, K., McCullough, S. H., and Bellugi, U. (1997). Enhanced facial discrimination: effects of experience with American Sign Language. *J. Deaf Stud. Deaf Educ.* 2, 223–233. doi: 10.1093/oxfordjournals.deafed.a014328
- Bosch, L., and Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition* 65, 33–69. doi: 10.1016/S0010-0277(97)00040-1
- Byers-Heinlein, K. (2009). Bilingualism Questionnaire Usage Guidelines. Personal communication.
- Cavallina, C., Puccio, G., Capurso, M., Bremner, A. J., and Santangelo, V. (2018). Cognitive development attenuates audiovisual distraction and promotes the selection of task-relevant perceptual saliency during visual search on complex scenes. *Cognition* 180, 91–98. doi: 10.1016/j.cognition.2018.07.003
- Chelazzi, L., Perlato, A., Santandrea, E., and Della Libera, C. (2013). Rewards teach visual selective attention. *Vision Res.* 85, 58–72. doi: 10.1016/j.visres.2012.12.005
- Costa, A., and Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nat. Rev. Neurosci.* 15, 336–345. doi: 10.1038/nrn3709
- De Filippo, C. L., and Lansing, C. R. (2006). Eye fixations of deaf and hearing observers in simultaneous communication perception. *Ear Hear.* 27, 331–352. doi: 10.1097/01.aud.0000226248.45263.ad
- de Klerk, C. C. J. M., Gliga, T., Charman, T., Johnson, M. H., and The BASIS team. (2014). Face engagement during infancy predicts later face recognition ability in younger siblings of children with autism. *Dev. Sci.* 17, 596–611. doi: 10.1111/desc.12141
- DeNicola, C. A., Holt, N. A., Lambert, A. J., and Cashon, C. H. (2013). Attention-orienting and attention-holding effects of faces on 4-to 8-month-old infants. *Int. J. Behav. Dev.* 37, 143–147. doi: 10.1177/0165025412474751
- Di Giorgio, E., Turati, C., Altoè, G., and Simion, F. (2012). Face detection in complex visual displays: an eye-tracking study with 3-and 6-month-old infants and adults. *J. Exp. Child Psychol.* 113, 66–77. doi: 10.1016/j.jecp.2012.04.012
- Elsabbagh, M., Gliga, T., Pickles, A., Hudry, K., Charman, T., Johnson, M. H., et al. (2013). The development of face orienting mechanisms in infants at-risk for autism. *Behav. Brain Res.* 251, 147–154. doi: 10.1016/j.bbr.2012.07.030
- Emmorey, K. (2001). *Language, Cognition, and the Brain: Insights from Sign Language Research*. London: Psychology Press.
- Emmorey, K., Thompson, R., and Colvin, R. (2008). Eye gaze during comprehension of American Sign Language by native and beginning signers. *J. Deaf Stud. Deaf Educ.* 14, 237–243. doi: 10.1093/deafed/enn037
- Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., and Csibra, G. (2005). Newborns' preference for face-relevant stimuli: effects of contrast polarity. *Proc. Natl. Acad. Sci. U.S.A.* 102, 17245–17250. doi: 10.1073/pnas.0502205102
- Frank, M. C., Vul, E., and Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition* 110, 160–170.
- Gliga, T., Elsabbagh, M., Andravizou, A., and Johnson, M. (2009). Faces attract infants' attention in complex displays. *Infancy* 14, 550–562.
- Hofmann, K., and Chilla, S. (2015). Bimodal bilingual language development of hearing children of deaf parents. *Eur. J. Spec. Needs Educ.* 30, 30–46. doi: 10.1080/08856257.2014.943563
- Johnson, M. H., Dziurawiec, S., Ellis, H., and Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition* 40, 1–19. doi: 10.1016/0010-0277(91)90045-6
- Kandel, S., Burfin, S., Meary, D., Ruiz-Tada, E., Costa, A., and Pascalis, O. (2016). The impact of early bilingualism on face recognition processes. *Front. Psychol.* 7:1080. doi: 10.3389/fpsyg.2016.01080
- Kristjánsson, Á., Sigurjónsdóttir, Ó., and Driver, J. (2010). Fortune and reversals of fortune in visual search: reward contingencies for pop-out targets affect search efficiency and target repetition effects. *Atten. Percept. Psychophys.* 72, 1229–1236. doi: 10.3758/APP.72.5.1229
- Lewkowicz, D. J., and Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1431–1436. doi: 10.1073/pnas.1114783109
- McCullough, S., and Emmorey, K. (1997). Face processing by deaf ASL signers: evidence for expertise in distinguishing local features. *J. Deaf Stud. Deaf Educ.* 2, 212–222. doi: 10.1093/oxfordjournals.deafed.a014327
- Meissner, C. A., and Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychol. Public Policy Law* 7, 3–35. doi: 10.1037/1076-8971.7.1.3
- Mercure, E., Kushnerenko, E., Goldberg, L., Bowden-Howl, H., Coulson, K., Johnson, M. H., et al. (2018). Language experience influences audiovisual speech integration in unimodal and bimodal bilingual infants. *Dev. Sci.* doi: 10.1111/desc.12701 [Epub ahead of print].
- Muir, L. J., and Richardson, I. E. (2005). Perception of sign language and its application to visual communications for deaf people. *J. Deaf Stud. Deaf Educ.* 10, 390–401. doi: 10.1093/deafed/eni037
- Palmer, S. B., Fais, L., Golinkoff, R. M., and Werker, J. F. (2012). Perceptual narrowing of linguistic sign occurs in the 1st year of life. *Child Dev.* 83, 543–553. doi: 10.1111/j.1467-8624.2011.01715.x
- Petitot, L. A., Katerelos, M., Levy, B. G., Gauna, K., Tétreault, K., and Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: implications for the mechanisms underlying early bilingual language acquisition. *J. Child Lang.* 28, 453–496. doi: 10.1017/S0305000901004718
- Pons, F., Bosch, L., and Lewkowicz, D. J. (2015). Bilingualism modulates infants' selective attention to the mouth of a talking face. *Psychol. Sci.* 26, 490–498. doi: 10.1177/0956797614568320
- Santangelo, V. (2015). Forced to remember: when memory is biased by salient information. *Behav. Brain Res.* 283, 1–10. doi: 10.1016/j.bbr.2015.01.013
- Sebastián-Gallés, N., Albareda-Castellot, B., Weikum, W. M., and Werker, J. F. (2012). A bilingual advantage in visual language discrimination in infancy. *Psychol. Sci.* 23, 994–999. doi: 10.1177/0956797612436817
- Stoll, C., Palluel-Germain, R., Caldara, R., Lao, J., Dye, M. W., Aptel, F., et al. (2017). Face recognition is shaped by the use of sign language. *J. Deaf Stud. Deaf Educ.* 23, 62–70. doi: 10.1093/deafed/enx034
- Valenza, E., Simion, F., Cassia, V. M., and Umiltà, C. (1996). Face preference at birth. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 892–903. doi: 10.1037/0096-1523.22.4.892
- Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., and Werker, J. F. (2007). Visual language discrimination in infancy. *Science* 316:1159. doi: 10.1126/science.1137686
- Werker, J. (2012). Perceptual foundations of bilingual acquisition in infancy. *Ann. N. Y. Acad. Sci.* 1251, 50–61. doi: 10.1111/j.1749-6632.2012.06484.x
- Woll, B., and Kyle, J. G. (1989). "Communication and language development in children of deaf parents," in *The Social and Cognitive Aspects of Normal and Atypical Language Development*, eds S. von Tetzchner, L. S. Siegel, and L. Smith (Berlin: Springer), 129–144.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Mercure, Quiroz, Goldberg, Bowden-Howl, Coulson, Gliga, Filippi, Bright, Johnson and MacSweeney. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Measuring the Timing of the Bilingual Advantage

Sara Incera*

Multilingual Laboratory, Department of Psychology, Eastern Kentucky University, Richmond, KY, United States

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Judith F. Kroll,
Pennsylvania State University,
United States
Evelyne Mercure,
University College London,
United Kingdom
Megan Zirnstein,
University of California, Riverside,
United States, in collaboration with
reviewer JK

*Correspondence:

Sara Incera
sara.incera@eku.edu

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 27 June 2018

Accepted: 27 September 2018

Published: 16 October 2018

Citation:

Incera S (2018) Measuring
the Timing of the Bilingual Advantage.
Front. Psychol. 9:1983.
doi: 10.3389/fpsyg.2018.01983

Empirical evidence has supported the idea that the bilingual advantage is a question of nuanced differences between bilinguals and monolinguals. In this article, I review findings from studies using eye tracking, mouse tracking, and event-related potentials (ERPs) which are particularly suited to measure time. Understanding the timing of the processes underlying executive function is crucial in evaluating the intricacies of the bilingual mind. Furthermore, I provide recommendations on how to best use these timing techniques to compare bilinguals and monolinguals. Temporal differences can characterize ongoing discussions of the bilingual advantage and help explain conflicting findings. Methodological and analytical innovations to better investigate the timing of the cognitive processes at play will inform a wide range of areas in cognitive science.

Keywords: bilingualism, timing, eye tracking, mouse tracking, event-related potentials

INTRODUCTION

More than half of the world's population is bilingual (Grosjean, 2010). Studying the cognitive processes (e.g., executive function, conflict monitoring) underlying the bilingual mind is an important topic. The bilingual advantage refers to the idea that being bilingual is linked to cognitive benefits (for a review see Bialystok, 2017). However, there are researchers that have challenged this idea (Paap and Greenberg, 2013; de Bruin et al., 2015; Paap, 2015). In light of the debate over the bilingual advantage, there is a need for a more nuanced explanation of the consequences of bilingualism. It is crucial to take into account information regarding who the bilinguals and monolinguals are (Luk and Bialystok, 2013), the types of experimental tasks implemented, the particular cognitive resources that may be critical to bilingualism (Takahesu Tabori et al., 2018), and the contexts in which bilinguals learned and normally use their languages (Green and Abutalebi, 2013). In addition to all of these variables, and possibly interacting with many of them, researchers need to consider the timing of the cognitive processes underlying participants' responses.

The focus of the present paper is the timing of the cognitive differences between bilinguals and monolinguals. By timing I refer to the first one second (1,000 ms) of participants' responses. Even though an important endeavor for researchers is to investigate the bilingual advantage over years or decades (Filippi et al., 2018; Incera and McLennan, 2018a), a review of those studies is beyond the scope of the current article. In addition, practice effects (Green and Abutalebi, 2013) and stimulus onset asynchrony manipulations (Martín et al., 2010) are likely to influence the bilingual advantage. However, investigations that do not measure participants' responses as they unfold over time are beyond the focus of this review. When talking about timing in the present paper I am always referring to the unfolding of participants' responses in milliseconds (ms). Using high temporal resolution techniques such as eye tracking, mouse tracking, and event-related brain potentials (ERPs), it is possible to analyze how each participant responds over time. Studying participants'

responses using time-sensitive techniques can guide the debate over the bilingual advantage by providing information about the timing of the cognitive processes at play.

Many researchers investigating the bilingual advantage have used experimental tasks in which the main outcome variable is reaction times (RTs). Typically, the dependent variable is the amount of time that participants take to complete a specific task, such as pressing a button after being exposed to visual or auditory stimuli. In this paper, I review studies that compare bilinguals and monolinguals using techniques that measure participants' responses over time. Furthermore, I put forward methodological recommendations (see **Table 1**) that I believe will improve our understanding of the timing of the bilingual advantage. The goal of these suggestions is to better compare across studies using high temporal resolution measures. Triangulating across these techniques can generate new research questions and provide novel insights. A better understanding of the timing of the cognitive processes underlying executive function can help uncover nuanced differences between bilinguals and monolinguals.

EYE TRACKING

Eye tracking has been available as a research tool since the 1970s (Cooper, 1974), but eye tracking did not become a mainstream methodology in spoken language research until the 1990s

(Tanenhaus et al., 1995; Tanenhaus and Spivey-Knowlton, 1996; Allopenna et al., 1998). Using the eye-tracking methodology, it is possible to measure “the probability of fixating a particular object as a function of time” (Tanenhaus and Spivey-Knowlton, 1996, p. 584). Researchers can analyze the total number of fixations on a specific area of the screen, or the proportion of fixations on areas of interest compared to control areas. Furthermore, it is typical to calculate the average number of fixations every 100 ms. Traditionally, eye-tracking figures include “time” on the x-axis and “proportion of fixations” on the y-axis (e.g., Allopenna et al., 1998). Most researchers report the first second (1,000 ms) of participants' responses from target onset and represent the different conditions (e.g., fixations to each object) as separate lines. This way of representing the results has also been used when reporting mouse-tracking and ERP data, which makes this method a convenient way to compare results across methodologies.

Many of the eye-tracking studies with bilingual populations have focused on reading (Libben and Titone, 2009; Pivneva et al., 2014; Cop et al., 2017; Enkin et al., 2017; Indrarathne and Kormos, 2018) or auditory processing (Spivey and Marian, 1999; Marian and Spivey, 2003; Blumenfeld and Marian, 2007; Bartolotti and Marian, 2012; Ito et al., 2018). However, there are a few studies that have used eye tracking to test the bilingual advantage hypothesis for inhibitory control (Bialystok et al., 2006; Blumenfeld and Marian, 2011; Mercier et al., 2014; Blumenfeld et al., 2016).

Bialystok et al. (2006) measured executive control using an antisaccade task, an experimental paradigm in which response suppression is required to resist moving the eyes toward the briefly exposed target. These researchers performed two studies, each with 96 participants (24 monolingual young adults, 24 bilingual young adults, 24 monolingual older adults, and 24 bilingual older adults) recruited from their university research pool in Toronto. Bialystok et al. (2006) found no effects of aging or bilingualism when the eye-tracking task was presented in isolation (Study 1). However, they found a bilingual advantage that increased with age when the same visual display was coupled with keypress responses (Study 2). The authors explained this pattern by stating: “Saccadic eye movements are more rapid (150–350 ms) than button-pressing responses (350–650 ms) and are arguably more automatic and less amenable to higher level cognitive control” (Bialystok et al., 2006, p. 1352).

The fact that effects can emerge in button press but not in eye tracking is not limited to the bilingual advantage. For example, long-term repetition priming effects (responding to a word faster when you have heard that word in a previous block of trials) are very robust in button-press tasks but do not emerge in eye-tracking tasks. To my knowledge, no published study has reported long-term repetition priming in proportion of fixations over time. It follows that triangulation across methodologies is crucial toward gaining a better understanding of the nature of the effects found in such experiments. These apparently contradictory results are puzzling, but can be an opportunity to refine our theories. Using the same stimuli across different techniques researchers can explore what aspects of the task are driving the results.

TABLE 1 | Summary of the recommendations covered in this article.

Design	Include a time-sensitive measure in your investigation Triangulate across measures and compare the results Collaborate with researchers who are proficient using other methods Control for variables that influence timing (language exposure/proficiency) Add a baseline condition before starting your experimental task
Analysis	Pre-register separate hypotheses for each time-sensitive measure Use factor analyses to combine different dependent variables Use factor analyses to combine different independent variables Decide whether you want to treat “time” as categorical or continuous Treat bilingualism as a continuous variable Focus on the interactions instead of the main effect of bilingualism Include trial presentation order as a control variable Report the raw (in ms) instead of the normalized mouse trajectories Report the time window of the effects instead of a single point in time Report stimulus-locked instead of response-locked ERP responses Create a section in the results to integrate across findings Develop statistical procedures to pinpoint the timing of the effects
Visualization	Create time figures that can be compared across measures Plot at least the first 1,000 ms of the responses (more if RTs are longer) Place time on the x-axis and the dependent variable on the y-axis Present different conditions as different lines Include lines for bilinguals and monolinguals within the same figure

Blumenfeld and Marian (2011) asked bilingual and monolingual participants to listen to words in English (their native language). For each trial participants had to identify the target word among four pictures, one of which was a similar-sounding within-language competitor (e.g., hamper/hammer). In the next trial the previously inhibited competitor picture became the target, a clever way to measure negative priming. In addition, participants responded to a version of the Stroop task in which they had to indicate the direction of an arrow. The arrow direction and arrow location could be congruent (leftward-facing arrow located on the left) or incongruent (leftward-facing arrow located on the right). These researchers reported a bilingual advantage in inhibitory control related to timing: "...bilinguals may return to a baseline activation state faster after inhibiting irrelevant information. In fact, the better bilinguals were at resolving Stroop interference, the less residual competitor inhibition they showed" (p. 11). Furthermore, they extended these findings to older adults: "...bilingual groups showed quicker target deactivation, reflecting more lifespan changes in activation for monolinguals than bilinguals" (Blumenfeld et al., 2016, p. 8). According to Blumenfeld and Marian (2011), the timing of inhibition (i.e., the time participants take to activate/deactivate a particular target) could be an important way in which bilinguals and monolinguals differ.

Mercier et al. (2014) monitored the eye movements of English monolinguals and French-English bilinguals while they listened to words in English. The non-target pictures included a within-language competitor, a between-language competitor, and a filler. Participants also responded to a battery of inhibitory control tasks. Mercier et al. (2014) reported a delayed onset of within-language competition for native French participants with low English exposure when compared to native English participants and to native French participants with high English exposure. According to these results, the timing of participant's responses not only differs between bilinguals and monolinguals, timing differs between bilingual groups with unequal levels of language exposure. If you test bilingual participants in English, those with more experience using English will respond faster than those with less experience using English.

While these studies have made tentative conclusions about time and have supported the idea that the timing of bilinguals and monolinguals differs, the reporting of the results is heavily focused on overall responses. As it is typical in the literature, researchers report overall patterns across several hundreds of milliseconds. Furthermore, it is common to create a separate graph for each group (bilingual/monolingual) and then show the patterns for the different conditions (target/within-language competitor/cross-language competitor/filler). While this approach is very useful to understand lexical activation, it might fall short to understand bilingual effects. To better evaluate group differences researchers need to compare the unfolding patterns of bilinguals and monolinguals by plotting them within the same figure. This approach will make it possible to measure the time at which the responses of bilingual and monolingual participants diverge.

MOUSE TRACKING

Mouse tracking is a tool that allows researchers to measure the unfolding of cognitive processes by recording participants' computer mouse trajectories (Spivey et al., 2005). Since the landmark PNAS article, "Continuous attraction toward phonological competitors" (Spivey et al., 2005), researchers have applied the mouse-tracking paradigm to a wide range of cognitive tasks. In 2009, the open source software MouseTracker became publicly available (Freeman and Ambady, 2010), making the technology accessible to a larger number of researchers. More recently, Kieslich and Henninger (2017) developed Mousetrap, an OpenSesame plugin that facilitates the combination of mouse tracking with other techniques such as eye tracking. Within the open science framework, researchers are building online communities to increase the exchange of validated experimental tasks across teams, an approach that increases replicability. Furthermore, Mousetrap directly connects to the statistical programming language R, a feature designed to streamline data analysis (Kieslich and Henninger, 2017).

Mouse-tracking measures have been implemented with bilingual populations (Bartolotti and Marian, 2012; Incera and McLennan, 2016, 2018a,b). In 2016, my co-author and I reported the results of a Stroop task in which English-Spanish bilinguals, English-Other bilinguals (a group that included a wide range of language backgrounds), and English monolinguals responded to Spanish and English color words (Incera and McLennan, 2016). We found that initiation times (the time it takes to start moving the mouse) were longer for the English-Spanish bilinguals, followed by the English-Other, and the English monolinguals. However, the overall trajectory was more efficient (straighter/faster) for those who took longer to start moving the mouse. In light of these results, we argued that bilinguals are qualitatively (as opposed to quantitatively) different from monolinguals. We proposed that this pattern of results indicates that bilinguals are experts at managing information (Incera and McLennan, 2016).

Results from our study provided initial support for the *Bilingual Expertise Hypothesis*, the idea that bilinguals are experts at managing information. The expertise pattern (i.e., longer initiation times coupled with more efficient responses) has been recently replicated in a study in which English monolinguals and Chinese-English bilinguals were compared using the Flanker, Simon, and Spatial Stroop tasks (Damian et al., 2018). Furthermore, this pattern also emerged in a Master's Thesis about attentional switching that compared bilingually exposed infants to their monolingual counterparts (Kakvan, 2017). Just as experts in a variety of domains (e.g., baseball) have a slower initiation of response followed by more efficient performance (Shank and Haywood, 1987; Incera and McLennan, 2016), bilinguals across different tasks show this expertise pattern.

The *Bilingual Expertise Hypothesis* can also be connected to the literature regarding the long term consequences of language experience. According to the *Adaptive Control Hypothesis* (Green and Abutalebi, 2013), language control processes adapt to the recurrent demands placed on them by the interactional context. One of the ways in which this adaptation might occur is that

bilinguals become experts at managing their languages. If that is the case, changes due to language exposure will not simply result in participants becoming “faster” or “slower” at responding to a particular task. Instead, language exposure could qualitatively alter the unfolding of participants’ responses. Furthermore, changes across the lifespan that influence cognitive processes could also interact with the expertise pattern. For example, older adults might take longer to initiate mouse movements regardless of their language background, an aging pattern that could obscure expertise effects in older groups. The short and long term consequences of bilingualism are likely to interact, resulting in a variety of patterns that researchers need to disentangle.

It is important to acknowledge that the expertise pattern not always emerges when comparing bilinguals and monolinguals in a mouse-tracking task. In a recent study, my co-author and I used a similar Stroop task to investigate bilingualism across the lifespan and did not find differences in initiation times (Incera and McLennan, 2018a). There are several differences between our 2016 and our 2018 study that could explain these apparently contradictory findings. First, in the 2016 study we presented four response alternatives in the screen (RED YELLOW – BLUE GREEN), while in the 2018 study there were only two (RED – GREEN). The working memory capacity necessary to keep in mind four (as opposed to two) responses could have enhanced the expertise pattern. Second, in the 2016 study Spanish and English words were presented randomly, while in the 2018 only English words were presented. Being in bilingual mode might be more likely to result in the emergence of the expertise pattern, a possibility supported by the fact that in the original experiment the expertise pattern was more pronounced in the English-Spanish bilinguals than the English-Other bilinguals. These results point to the idea that task characteristics are likely to influence the unfolding of participants’ responses.

Another interesting aspect of the Incera and McLennan’s (2018a) study is that, contrary to previous research (Bialystok et al., 2004, 2008; Blumenfeld et al., 2016), no Bilingualism by Age interaction emerged. Instead, our results suggest that after controlling for baseline performance the bilingual advantage remains stable across the lifespan. Consequently, it is important to control for baseline motor differences between groups. Choices like the distance or size of the target can alter the mouse trajectory (Walker et al., 1997). Controlling for differences in motor movements is particularly important in quasi-experimental approaches—when comparing participants that cannot be randomly assigned to groups. To evaluate the influence of personal variables (e.g., bilingualism, age), it is necessary to distinguish effects at the motor level from those arising at the cognitive level. To do so, I strongly encourage researchers to add a baseline measure to their studies (see Incera and McLennan, 2018a, for an example of a baseline task).

Another important consideration to be mindful of when analyzing mouse-tracking data is the abundance of dependent variables. MouseTracker (Freeman and Ambady, 2010) provides numerous overall variables that summarize the trajectory using a single number: initiation time, reaction time, maximum deviation, area under the curve, maximum deviation time, x -flips, and y -flips. Based on preliminary analyses of the data collected

in my lab, most of these variables tend to load onto two factors: (1) how straight are the mouse movements? (area under the curve, maximum deviation, x flips) and (2) how fast are the mouse movements? (initiation time, reaction time, maximum deviation time). Additional factor analyses are necessary to properly evaluate whether these two factors remain stable across different populations and tasks. Moreover, factor analysis is a powerful methodology to summarize across a wide range of independent variables traditionally used in bilingual research (Marian et al., 2007; Anderson et al., 2018a,b).

The key advantage of mouse tracking is that this paradigm provides measures that unfold over time: x -coordinates, y -coordinates, velocity, acceleration, and angle. The most commonly reported dependent variable—and closest equivalent to proportion of fixations—is x -coordinates over time. When looking at the mouse trajectories (Incera and McLennan, 2018a, Figure 2), it is possible to observe that the difference in x -coordinates (separation of the lines) between bilinguals and monolinguals emerges around 500 ms after stimulus onset. These results follow those of Bialystok et al. (2006) eye-tracking study in that the bilingual advantage may be evident only later on in the response. If we want to represent the mouse trajectories in line with the eye-tracking figures, we should put time on the x -axis, and x -coordinates on the y -axis. Alternatively, it is possible to represent these trajectories to closely mimic the visual display of the actual experiment. To mimic the visual display, we need to flip the figure by putting time on the y -axis and the dependent variable (x -coordinates) on the x -axis. The latter approach (time: y -axis) is more visually appealing, but the former (time: x -axis) might be better aligned with the way data from eye-tracking and event-related potentials are often represented.

EVENT-RELATED POTENTIALS

Event-related brain potentials provide detailed information about timing (see Moreno et al., 2008, for an overview of ERPs in the study of bilingual language processing). Several research teams have investigated bilingual populations using ERPs (Liu and Perfetti, 2003; Moreno and Kutas, 2005; Ojima et al., 2005; Kotz, 2009; Van Heuven and Dijkstra, 2010; Garcia-Sierra et al., 2011; Martin et al., 2013; Grundy et al., 2017; Zirnstein et al., 2018). Researchers have used this methodology to specifically test the bilingual advantage by measuring the effects of learning a second language on brain activation (Sullivan et al., 2014; Moreno and Lee, 2015) and by comparing bilinguals’ and monolinguals’ levels of executive control (Kousaie and Phillips, 2012, 2016; Kuipers and Thierry, 2013; Coderre and Van Heuven, 2014; Moreno et al., 2014; Heidlmayr et al., 2015; Grundy et al., 2017; Zirnstein et al., 2018). In this review, I focus on studies that used the Stroop task to investigate how the cognitive processes underlying the bilingual advantage unfold over time (Kousaie and Phillips, 2012, 2016; Coderre and Van Heuven, 2014; Heidlmayr et al., 2015).

In the Stroop task (Stroop, 1935) participants need to avoid reading the word and instead report the color of the stimuli in front of them (e.g., answering “green” to the stimuli BLUE written

in green font). The Stroop effect refers to the difference between the incongruent (BLUE in green) and the congruent (BLUE in blue) conditions. The Stroop task has been used in numerous studies to investigate the timing of conflict resolution (Liotti et al., 2000; Badzakova-Trajkov et al., 2009). In monolingual participants, researchers have found an effect between 400 and 450 ms (Liotti et al., 2000) or between 370 and 480 ms (Badzakova-Trajkov et al., 2009); this negative interference effect has been associated with the N400. According to Badzakova-Trajkov et al. (2009), in the Stroop task the N400 emerges in the anterior cingulate region, and it is likely to reflect the identification and resolution of conflict between reading the word and naming the color. The N400 is also an important ERP component in the bilingual literature (Kerkhofs et al., 2006; Midgley et al., 2009).

Heidlmayr et al. (2015) compared French-German bilinguals to French monolinguals in an adapted version of the Stroop task. In addition to congruent, incongruent, and control conditions participants had to respond to a negative priming condition (the color inhibited in the previous trial becomes the target color in the new trial). In line with eye-tracking and mouse-tracking studies that speculated that the bilingual advantage might only become evident relatively late during processing (Bialystok et al., 2006; Incera and McLennan, 2018a), Heidlmayr et al. (2015) found reduced ERP effects in bilinguals' responses to the Stroop task in the N400 and in late time windows (540–700 ms). These researchers found a bilingual advantage in the N400 Stroop effect over the posterior scalp, associated with the anterior cingulate cortex. Heidlmayr et al. (2015) did not find group differences in early components (e.g., N200, P300), but the N400 Stroop effect was reduced in bilinguals when compared to monolinguals.

Kousaie and Phillips (2012, 2016) used ERPs to compare high proficient English-French bilinguals to English monolinguals in the Stroop, Simon, and Flanker tasks (Kousaie and Phillips, 2012, 2016). In the Stroop task, the P300 peaked earlier for young bilinguals than young monolinguals (Kousaie and Phillips, 2012) and the N200 peaked earlier for old bilinguals than old monolinguals (Kousaie and Phillips, 2016). It is important to highlight that Kousaie and Phillips defined the N200 between 220 and 360 ms, and the P300 between 300 and 500 ms (which technically includes the N400). When looking at the waveforms of their Stroop task (Kousaie and Phillips, 2012, Figure 2), it becomes obvious that the bilingual and monolingual lines diverge during both the P300 and the N400. In order to better compare the time-course of the bilingual advantage across studies, researchers need to report the specific time period during which bilingual and monolingual groups differ.

Coderre and Van Heuven (2014) used ERPs to compare a group of Chinese-English bilinguals to a group of English monolinguals in a version of the Stroop task in which stimulus onset asynchronies (SOAs) were manipulated (the word and the color were not always presented at the same time). Coderre and Van Heuven (2014) found a significant negative effect at Cz and Pz between 350 and 550 ms in the monolingual group and the bilingual group when tested in their native language. However, when bilinguals were tested in their second language the effect

was delayed (see Mercier et al., 2014, for equivalent findings in eye tracking). It is important to highlight that the time window reported by Coderre and Van Heuven (2014) (350 – 550 ms) incorporates the previously discussed P300 (Kousaie and Phillips, 2012) and N400 (Heidlmayr et al., 2015) components.

In order to compare across studies it is important to better determine how many milliseconds after stimulus onset a particular process is expected to emerge. A helpful approach to avoid large time-windows is to report peak latencies. For example, Coderre and Van Heuven (2014) reported that the bilingual L2 incongruent effect (529 ms) peaked later than the bilingual L1 (459 ms) and the monolingual (434 ms) incongruent effect. In order to report peak latencies, the ERP averages need to be time-locked to the moment in time in which the stimulus was presented. Researchers need to carefully consider the theoretical implications of reporting stimulus-locked (time-locked to the moment in time in which the stimulus was presented) or response-locked (time-locked to the response of the participant) ERP averages. In order to compare ERP responses to eye-tracking and mouse-tracking responses, I recommend reporting responses locked to the moment in time in which the stimulus was presented.

Crucially, Coderre and Van Heuven (2014) reported that in the –400 ms SOA, the bilingual L1 experienced a significantly later Stroop effect compared to monolinguals. This delayed onset of conflict processing in bilinguals could be indicative of enhanced inhibitory control. Coderre and Van Heuven (2014) discuss these findings in line with the dual control theory (De Pisapia and Braver, 2006; Braver et al., 2009). According to Braver and Colleagues (2009), there are “two mechanisms of cognitive control: one a “late correction” reactive response engaged to resolve conflict once it has occurred; and one a proactive “early selection” strategy engaged to pre-emptively reduce control demands for when conflict occurs.” (Coderre and Van Heuven, 2014, p. 13). This dual control theory aligns with the predictions derived from the *Bilingual Expertise Hypothesis*. First, the proactive “early selection” strategy could be the reason why bilinguals take longer to start moving the mouse. Second, the “late correction” reactive response relates to how bilinguals respond faster later on. Differences between bilinguals and monolinguals could emerge from alternative ways of processing information through these two mechanisms of cognitive control.

INTEGRATION

Triangulating eye-tracking, mouse-tracking, and ERP measures can be tremendously useful in painting a clearer picture of the timing of the bilingual advantage. When trying to evaluate the timing of a particular task across different techniques it becomes obvious that there are numerous gaps in the literature. However, the few studies that have focused on timing point to the conclusion that investigating the unfolding of participants' responses can help improve our understanding of the differences between bilinguals and monolinguals. In order to move forward it is important to (1) use the same sample and task across different techniques, (2) use the same task and technique across different

samples, and (3) use the same technique and sample across different tasks. The type of task being used, and the cognitive processes engaged in that particular task, are likely to influence the timing of participants' responses. Only by triangulating across samples, tasks, and techniques it will be possible to understand the timing of the cognitive processes driving these effects.

Pioneer researchers have already made efforts to integrate eye tracking and mouse tracking in their work with bilinguals. Bartolotti and Marian (2012) reported eye-tracking and mouse-tracking data collected within the same task. These researchers trained bilingual and monolingual participants in an artificial language to be able to compare them. Participants listened to spoken words and had to choose from pairs of drawings in the screen (Bartolotti and Marian, 2012). According to their eye-tracking data, bilingual and monolingual participants experienced similar early activation of the native-language competitor (200 ms after word onset) but bilinguals resolved the competition faster than monolinguals (700 ms vs. 1400 ms). While Bartolotti and Marian (2012) used the mouse-tracking results to discuss how bilinguals and monolinguals differ in the way in which they manage competition, they did not report specific timing information derived from the mouse trajectories.

Bartolotti and Marian (2012) reported the normalized, as opposed to the raw, mouse trajectories (this distinction relates to the previously mentioned way of plotting ERP data by using stimulus-locked vs. response-locked averages). The normalized mouse trajectories standardize participants' responses by dividing each trajectory in 100 bins. These bins include longer time windows for slower participants (e.g., 50 ms per bin for someone who took 5000 ms to respond) and shorter time windows for faster participants (e.g., 10 ms per bin for someone who took 1000 ms to respond). Normalized trajectories can be useful to answer questions like: what was the position of the mouse half way through the response? However, raw mouse trajectories are necessary to answer questions like: how many milliseconds after stimulus onset does the bilingual advantage emerge? Researchers can only examine the average time at which a particular effect emerges using raw trajectories (e.g., x-coordinates over time).

In addition to measuring participants' responses to the same task using different techniques, it is important to analyze the data in an integrated way. Researchers tend to report results from different dependent variables in separate sections. I recommend creating a paragraph within the results section in which the outcomes from different techniques can be integrated (similar to the "General Discussion" when reporting several experiments). It would be helpful to plot the eye- and mouse-tracking data in a single plot, and to discuss the similarities and differences of the timing across these techniques. Importantly, the way in which the data from these different methodologies converge can be as informative as the way in which they differ.

SUGGESTIONS

Combining time-sensitive techniques can be extremely useful when trying to understand the time course of the cognitive processes underlying executive function. However, it is important

to keep in mind that using different methodologies can pose technical challenges and increase the complexity of the statistical analyses. Team collaborations, in which different researchers are experts in a variety of methodologies, can be highly effective in overcoming these difficulties. Furthermore, it is important to pre-register specific hypotheses for each technique, in particular when differences between these methodologies are likely to emerge. Triangulating across techniques can substantially increase the number of dependent variables, so researchers need to clearly distinguish between confirmatory and exploratory analyses.

Numerous analytical innovations have been proposed in an effort to shed new light on the discussion surrounding the bilingual advantage (Woumans and Duyck, 2015; Calvo et al., 2016). Useful methodological advances like multiverse analysis—performing all analyses across the whole set of alternatively processed data sets corresponding to a large set of reasonable scenarios (Steege et al., 2016)—are being implemented to investigate whether arbitrary analytical choices can influence the effects of language usage on executive function (Poarch et al., 2018). Since it is virtually impossible to perfectly match bilinguals and monolinguals (Filippi et al., 2018), it is important to control for baseline levels of performance and to focus on the group by condition interactions—as opposed to the main effect of bilingualism (Incera and McLennan, 2018a). In addition, including trial presentation order as a control variable (Mercier et al., 2014) can help eliminate noise and improve the quality of the analysis.

When using statistical analyses to investigate responses over time it is crucial to properly model the covariance structure. When data points are collected over time, it is logical to assume that measures of the same participant are correlated. Data points that are closer together tend to correlate more than data points that are farther apart, which challenges the assumption of random error. Therefore, time analyses must address the issue of covariation between time points. Ignoring the covariance structure when modeling time can lead to erroneous inferences (Littell et al., 2000; Lui et al., 2012). According to Littell et al. (2000) the choice of the covariance structure can have important effects on tests and estimates of fixed effects. Lui et al. (2012) argued that researchers need to empirically consider what type of error structure best fits the data. To do so, they recommend using AIC and BIC in the selection of a proper residual covariance structure. The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are tools to compare statistical models in order to choose the best fit for a given set of data. Covariation can be a problem when analyzing timing data, researchers need to ensure they are choosing models with the right covariance structure.

A key question that researchers focusing on timing need to consider is whether "time" should be treated as a categorical (Kousaie and Phillips, 2012, 2016; Mercier et al., 2014) or as a continuous (Blumenfeld et al., 2016; Incera and McLennan, 2016, 2018a) variable (for a discussion of the statistical implications of this choice see Lui et al., 2012). The advantage of treating time as a categorical variable is that you can use specific time windows (e.g., P300, N400) to compare across studies. In addition, this approach simplifies the statistical analyses and allows for clearer

a priori predictions. However, focusing on 100 ms time bins is a crude approach when the goal is to better understand the timing of the effects. Researchers have argued against the practice of categorizing continuous variables (MacCallum et al., 2002) and in favor of treating time (and bilingualism) as continuous variables (Incera and McLennan, 2018a). Approaches like growth curve analysis (Mirman, 2016), latent growth curve analysis (Ferrer et al., 2008), and piecemeal growth curve analysis (Calet et al., 2015), can be useful when treating time as a continuous variable. These methodologies take into account the overall pattern of the trajectory instead of focusing on arbitrary time windows.

Temporal differences are often easy to visualize in figures, but relatively difficult to pinpoint with our current statistical methods. For example, in a mouse-tracking study in which a group of Spanish-English bilinguals participated in a Stroop task with Spanish and English color words (Incera and McLennan, 2018b), my co-author and I reported that within-language interference (English words with English response alternatives) emerged 80 ms earlier than between-language interference (Spanish words with English response alternatives). It is obvious that if we had used 100 ms time-windows we would have missed this 80 ms time difference. Instead, we performed 50 within-participants *t*-tests (one every 20 ms) for the first 1,000 ms of the mouse trajectories. To maintain the overall Type-I error rate below 0.05, we used Monte Carlo simulations to calculate the minimum threshold of contiguous *t*-tests that had to be significant in order to consider the effect real (for a detailed explanation of this approach, see Dale et al., 2007; Yamamoto et al., 2016). Using this threshold, we observed that interference emerged 420 ms after stimulus onset in the within-language condition and 500 ms after stimulus onset in the between-language condition, which led us to conclude that the difference is 80 ms.

To my knowledge, there is no clear path to test whether this 80 ms temporal difference is a real effect above and beyond random chance. One approach could be to perform 50 ANOVAs, but establishing thresholds using Monte Carlo simulations would become increasingly difficult. Another approach could be to use growth curve analysis. However, it is not clear how researchers can use this technique (a tool that was created to evaluate the overall pattern of the trajectory) to pinpoint the moment at which two trajectories diverge. Even piecemeal growth curve analysis can be limited when the goal is to evaluate timing because researchers tend to use theoretical reasons (not empirical analyses) to select the time periods for the different growth

patterns. As such, developing new statistical approaches that researchers can use to specify the moment at which a particular cognitive process influences participants' responses (e.g., an analysis of the point of divergence between two trajectories or the inflection point within a single trajectory) is an important endeavor likely to inform other areas of psychological science.

CONCLUSION

While data on the timing of the bilingual advantage are scarce, the empirical evidence available suggests that the effects of language experience unfold differently in the bilingual mind than in the monolingual mind. Bilinguals may be more efficient processors than monolinguals, but those effects may only be evident at certain points in time, and may differ across different samples and tasks. Understanding the timing of these effects can help explain why and how bilinguals process information differently. Therefore, it is crucial to take advantage of temporally sensitive methodologies such as eye tracking, mouse tracking, and ERPs, in order to better understand the bilingual advantage.

Investigating the timing of the bilingual advantage has the potential to stimulate new research questions and provide novel insights. Focusing only on the final outcome of participants' responses can lead to inconclusive results because of subtle time differences in the unfolding of the underlying cognitive processes. In addition to many other important aspects of the bilingual experience (e.g., sample characteristics, task characteristics), researchers need to consider the timing of the cognitive processes at play. Methodological and analytical innovations to better investigate the timing of the bilingual advantage have the potential to inform a wide range of areas in psychological science.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

Thank you to Conor T. McLennan, Maria J. Donaldson-Misener, Adam L. Lawson, and Lisa M. Stronsick for feedback on previous versions of this manuscript.

REFERENCES

- Alloppenna, P. D., Magnuson, J. S., and Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: evidence for continuous mapping models. *J. Mem. Lang.* 38, 419–439. doi: 10.1006/jmla.1997.2558
- Anderson, J. A. E., Hawrylewicz, K., and Bialystok, E. (2018a). Who is bilingual? Snapshots across the lifespan. *Biling. Lang. Cogn.* 1–12. doi: 10.1017/S1366728918000950
- Anderson, J., Mak, L., Chahi, A., and Bialystok, E. (2018b). The language and social background questionnaire: assessing degree of bilingualism in a diverse population. *Behav. Res. Methods* 50, 250–263. doi: 10.3758/s13428-017-0867-9
- Badzakova-Trajkov, G., Barnett, K. J., Waldie, K. E., and Kirk, I. J. (2009). An ERP investigation of the Stroop task: the role of the cingulate in attentional allocation and conflict resolution. *Brain Res.* 1253, 139–148. doi: 10.1016/j.brainres.2008.11.069
- Bartolotti, J., and Marian, V. (2012). Language learning and control in monolinguals and bilinguals. *Cogn. Sci.* 36, 1129–1147. doi: 10.1111/j.1551-6709.2012.01243.x
- Bialystok, E. (2017). The bilingual adaptation: how minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099
- Bialystok, E., Craik, F., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 859–873. doi: 10.1037/0278-7393.34.4.859

- Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290
- Bialystok, E., Craik, F. I. M., and Ryan, J. (2006). Executive control in a modified antisaccade task: effects of aging and bilingualism. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1341–1354. doi: 10.1037/0278-7393.32.6.1341
- Blumenfeld, H., and Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: examining proficiency and lexical status using eye-tracking. *Lang. Cogn. Process.* 22, 633–660. doi: 10.1080/01690960601000746
- Blumenfeld, H., and Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition* 118, 245–257. doi: 10.3174/ajnr.A1256
- Blumenfeld, H., Schroeder, S. R., Bobb, S. C., Freeman, M. R., and Marian, V. (2016). Auditory word recognition across the lifespan: links between linguistic and nonlinguistic inhibitory control in bilinguals and monolinguals. *Linguist. Approaches Biling.* 6, 119–146. doi: 10.1021/acschemneuro.5b00094
- Braver, T. S., Paxton, J. L., Locke, H. S., and Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7351–7356. doi: 10.1073/pnas.0808187106
- Calet, N., Gutiérrez-Palma, N., Simpson, I. C., González-Trujillo, M. C., and Defior, S. (2015). Suprasegmental phonology development and reading acquisition: a longitudinal study. *Sci. Stud. Read.* 19, 51–71. doi: 10.1080/1088438.2014.976342
- Calvo, N., García, A. M., Manoloff, L., and Ibáñez, A. (2016). Bilingualism and cognitive reserve: a critical overview and a plea for methodological innovations. *Front. Aging Neurosci.* 7:249. doi: 10.3389/fnagi.2015.00249
- Coderre, E. L., and Van Heuven, W. J. B. (2014). Electrophysiological explorations of the bilingual advantage: evidence from a Stroop task. *PLoS One* 9:e103424. doi: 10.1371/journal.pone.0103424
- Cooper, R. (1974). The control of eye fixation by the meaning of spoken language: a new methodology for the real-time investigation of speech perception, memory, and language. *Cogn. Psychol.* 6, 84–107. doi: 10.1016/0010-0285(74)90005-X
- Cop, U., Dirix, N., Drieghe, D., and Duyck, W. (2017). Presenting GECO: an eyetracking corpus of monolingual and bilingual sentence reading. *Behav. Res. Methods* 49, 602–615. doi: 10.3758/s13428-016-0734-0
- Dale, R., Kehoe, C., and Spivey, M. J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Mem. Cogn.* 35, 15–28. doi: 10.3758/BF03195938
- Damian, M. F., Ye, W., Oh, M., and Yang, S. (2018). Bilinguals as “experts”? Comparing performance of mono- to bilingual individuals via a mousetracking paradigm. *Biling. Lang. Cogn.* 1–18. doi: 10.1017/S1366728918000901
- de Bruin, A., Treccani, B., and Della Sala, S. (2015). The connection is in the data: we should consider them all. *Psychol. Sci.* 26, 947–949. doi: 10.1177/0956797615583443
- De Pisapia, N., and Braver, T. S. (2006). A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. *Neurocomputing* 69, 1322–1326. doi: 10.1016/j.neucom.2005.12.100
- Enkin, E., Nicol, J., Brooks, Z., and Zavaleta, K. L. (2017). Reading in Spanish as a second language: an eye-tracking study. *Read. Matrix* 17, 1–15.
- Ferrer, E., Balluerka, N., and Widaman, K. F. (2008). Factorial invariance and the specification of second-order latent growth models. *Methodology* 4, 22–36. doi: 10.1027/1614-2241.4.1.22
- Filippi, R., D’Souza, D., and Bright, P. (2018). A developmental approach to bilingual research: the effects of multi-language experience from early infancy to old age. *Int. J. Biling.* 1–28. doi: 10.1177/1367006917749061
- Freeman, J. B., and Ambady, N. (2010). MouseTracker: software for studying real-time mental processing using a computer mouse-tracking method. *Behav. Res. Methods* 42, 226–241. doi: 10.3758/BRM.42.1.226
- García-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., et al. (2011). Bilingual language learning: an ERP study relating early brain responses to speech, language input, and later word production. *J. Phon.* 39, 546–557. doi: 10.1016/j.wocn.2011.07.002
- Green, D. W., and Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377
- Grosjean, F. (2010). *Bilingual: Life and Reality*. Cambridge, MA: Harvard University Press. doi: 10.4159/9780674056459
- Grundy, J. G., Anderson, J. A. E., and Bialystok, E. (2017). Bilinguals have more complex EEG brain signals in occipital regions than monolinguals. *Neuroimage* 159, 280–288. doi: 10.1016/j.neuroimage.2017.07.063
- Heidlmayr, K., Hemforth, B., Moutier, S., and Isel, F. (2015). Neurodynamics of executive control processes in bilinguals: evidence from ERP and source reconstruction analyses. *Front. Psychol.* 6:821. doi: 10.3389/fpsyg.2015.00821
- Incera, S., and McLennan, C. T. (2016). Mouse tracking reveals that bilinguals behave like experts. *Biling. Lang. Cogn.* 19, 610–620. doi: 10.1017/S1366728915000218
- Incera, S., and McLennan, C. T. (2018a). Bilingualism and age are continuous variables that influence executive function. *Aging Neuropsychol. Cogn.* 25, 443–463. doi: 10.1080/13825585.2017.1319902
- Incera, S., and McLennan, C. T. (2018b). The time course of within and between-language interference in bilinguals. *Int. J. Biling.* 22, 88–99. doi: 10.1177/1367006916644688
- Indrarathne, B., and Kormos, J. (2018). The role of working memory in processing L2 input: insights from eye-tracking. *Bilin. Lang. Cogn.* 21, 355–374. doi: 10.1017/S1366728917000098
- Ito, A., Pickering, M. J., and Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: a visual world eye-tracking study. *J. Mem. Lang.* 98, 1–11. doi: 10.1016/j.jml.2017.09.002
- Kakvan, M. (2017). *Attentional Switching in Infants Exposed to Bilingual- Versus Monolingual Environment*. Available at: <http://hdl.handle.net/10315/33562>
- Kerkhofs, R., Dijkstra, T., Chwilla, D. J., and De Bruijn, E. R. A. (2006). Testing a model for bilingual semantic priming with interlingual homographs: RT and N400 effects. *Brain Res.* 1068, 170–183. doi: 10.1016/j.brainres.2005.10.087
- Kieslich, P. J., and Henninger, F. (2017). Mousetrap: an integrated, open-source mouse-tracking package. *Behav. Res. Methods* 49, 1652–1667. doi: 10.3758/s13428-017-0900-z
- Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain Lang.* 109, 68–74. doi: 10.1016/j.bandl.2008.06.002
- Kousaie, S., and Phillips, N. A. (2012). Conflict monitoring and resolution: are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain Res.* 1446, 71–90. doi: 10.1016/j.brainres.2012.01.052
- Kousaie, S., and Phillips, N. A. (2016). A behavioural and electrophysiological investigation of the effect of bilingualism on aging and cognitive control. *Neuropsychologia* 94, 23–35. doi: 10.1016/j.neuropsychologia.2016.11.013
- Kuipers, J. R., and Thierry, G. (2013). ERP-pupil size correlations reveal how bilingualism enhances cognitive flexibility. *Cortex* 49, 2853–2860. doi: 10.1016/j.cortex.2013.01.012
- Libben, M. R., and Titone, D. A. (2009). Bilingual lexical access in context: evidence from eye movements during reading. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 381–390. doi: 10.1037/a0014875
- Liotti, M., Woldorff, M. G., Perez, R., and Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia* 38, 701–711. doi: 10.1016/S0028-3932(99)00106-2
- Littell, R. C., Pendergast, J., and Natarajan, R. (2000). Modelling covariance structure in the analysis of repeated measures data. *Stat. Med.* 19, 1793–1819. doi: 10.1002/1097-0258(20000715)19:13<1793::AID-SIM482>3.0.CO;2-Q
- Liu, Y., and Perfetti, C. A. (2003). The time course of brain activity in reading English and Chinese: an ERP study of Chinese bilinguals. *Hum. Brain Mapp.* 18, 167–175. doi: 10.1002/hbm.10090
- Lui, S., Rovine, M. J., and Molenaar, P. C. M. (2012). Selecting a linear mixed model for longitudinal data: repeated measures analysis of variance, covariance pattern model, and growth curve approaches. *Psychol. Methods* 17, 15–30. doi: 10.1037/a0026971
- Luk, G., and Bialystok, E. (2013). Bilingualism is not a categorical variable: interaction between language proficiency and usage. *J. Cogn. Psychol.* 25, 605–621.
- MacCallum, R. C., Zhang, S., Preacher, K. J., and Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychol. Methods* 7, 19–40. doi: 10.1037/1082-989X.7.1.19
- Marian, V., Blumenfeld, H., and Kaushanskaya, M. (2007). Bilinguals and multilinguals. *Hear. Res.* 50, 940–967.
- Marian, V., and Spivey, M. (2003). Competing activation in bilingual language processing: within- and between-language competition. *Biling. Lang. Cogn.* 6, 97–115. doi: 10.1017/S1366728903001068

- Martin, C. D., Strijkers, K., Santesteban, M., Escera, C., Hartsuiker, R. J., and Costa, A. (2013). The impact of early bilingualism on controlling a language learned late: an ERP study. *Front. Psychol.* 4:815. doi: 10.3389/fpsyg.2013.00815
- Martín, M. C., Macizo, P., and Bajo, T. (2010). Time course of inhibitory processes in bilingual language processing. *Br. J. Psychol.* 101, 679–693. doi: 10.1348/000712609X480571
- Mercier, J., Pivneva, I., and Titone, D. (2014). Individual differences in inhibitory control relate to bilingual spoken word processing. *Biling. Lang. Cogn.* 17, 89–117. doi: 10.1017/S1366728913000084
- Midgley, K. J., Holcomb, P. J., and Grainger, J. (2009). Language effects in second language learners and proficient bilinguals investigated with event-related potentials. *J. Neurolinguistics* 22, 281–300.
- Mirman, D. (2016). *Growth Curve Analysis and Visualization Using R*. Boca Raton, FL: CRC Press.
- Moreno, E., and Kutas, M. (2005). Processing semantic anomalies in two languages: an electrophysiological exploration in both languages of Spanish-English bilinguals. *Cogn. Brain Res.* 22, 205–220. doi: 10.1016/j.cogbrainres.2004.08.010
- Moreno, E. M., Rodríguez-Fornells, A., and Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *J. Neurolinguistics* 21, 477–508. doi: 10.1016/j.jneuroling.2008.01.003
- Moreno, S., and Lee, Y. (2015). Short-term second language and music training induces lasting functional brain changes in early childhood. *Child Dev.* 82, 394–406. doi: 10.1111/cdev.12297
- Moreno, S., Wodniecka, Z., Tays, W., Alain, C., and Bialystok, E. (2014). Inhibitory control in bilinguals and musicians: event related potential (ERP) evidence for experience-specific effects. *PLoS One* 9:e94169. doi: 10.1371/journal.pone.0094169
- Ojima, S., Nakata, H., and Kakigi, R. (2005). An ERP study of second language learning after childhood: effects of proficiency. *J. Cogn. Neurosci.* 17, 1212–1228. doi: 10.1162/0898929055002436
- Paap, K. R. (2015). The neuroanatomy of bilingualism: will winds of change lift the fog? *Lang. Cogn. Neurosci.* 3798, 1–4. doi: 10.1080/23273798.2015.1082607
- Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Pivneva, I., Mercier, J., and Titone, D. (2014). Executive control modulates cross-language lexical activation during L2 reading: evidence from eye movements. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 787–796. doi: 10.1037/a0035583
- Poarch, G. J., Vanhove, J., and Berthele, R. (2018). The effect of bidialectalism on executive function. *Int. J. Biling.* 1–17. doi: 10.1177/1367006918763132
- Shank, M. D., and Haywood, K. M. (1987). Eye movements while viewing a baseball pitch. *Percept. Mot. Skills* 64(3 Pt 2), 1191–1197. doi: 10.2466/pms.1987.64.3c.1191
- Spivey, M. J., Grosjean, M., and Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proc. Natl. Acad. Sci. U.S.A.* 102, 10393–10398. doi: 10.1073/pnas.0503903102
- Spivey, M. J., and Marian, V. (1999). Cross talk between native and second languages: partial Activation of an Irrelevant Lexicon. *Psychol. Sci.* 10, 281–284. doi: 10.1111/1467-9280.00151
- Steegen, S., Tuerlinckx, F., Gelman, A., and Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspect. Psychol. Sci.* 11, 702–712. doi: 10.1177/1745691616658637
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651
- Sullivan, M., Janus, M., Moreno, S., Astheimer, L., and Bialystok, E. (2014). Early stage second-language learning improves executive control: evidence from ERP. *Brain Lang.* 139, 84–98. doi: 10.1016/j.bandl.2014.10.004
- Takahesu Tabori, A. A., Mech, E. N., and Atagi, N. (2018). Exploiting language variation to better understand the cognitive consequences of bilingualism. *Front. Psychol.* 9:1686. doi: 10.3389/fpsyg.2018.01686
- Tanenhaus, M. K., and Spivey-Knowlton, M. J. (1996). Eye-Tracking. *Lang. Cogn. Process.* 11, 583–588. doi: 10.1080/016909696386971
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., Sedivy, J. C., Everhard, K. M., Sedivy, J. C., et al. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* 268, 1632–1634. doi: 10.1126/science.7777863
- Van Heuven, W. J. B., and Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Res. Rev.* 64, 104–122. doi: 10.1016/j.brainresrev.2010.03.002
- Walker, N., Philbin, D. A., and Fisk, A. D. (1997). Age-Related differences in movement control: adjusting submovement structure to optimize performance. *J. Gerontol. Seri. B Psychol. Sci. Soc. Sci.* 52, 40–53.
- Woumans, E., and Duyck, W. (2015). The bilingual advantage debate: moving toward different methods for verifying its existence. *Cortex* 73, 356–357. doi: 10.1016/j.cortex.2015.07.012
- Yamamoto, N., Incera, S., and McLennan, C. T. (2016). A reverse stroop task with mouse tracking. *Front. Psychol.* 7:670. doi: 10.3389/fpsyg.2016.00670
- Zirnsstein, M., van Hell, J. G., and Kroll, J. F. (2018). Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition* 176, 87–106. doi: 10.1016/j.cognition.2018.03.001

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer EM and handling editor declared their shared affiliation at the time of the review.

Copyright © 2018 Incera. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Importance of Socioeconomic Status as a Modulator of the Bilingual Advantage in Cognitive Ability

Kamila Naeem¹, Roberto Filippi¹, Eva Periche-Tomas¹, Andriani Papageorgiou¹ and Peter Bright^{2*}

¹ Institute of Education, University College London, London, United Kingdom, ² Department of Psychology, Anglia Ruskin University, Cambridge, United Kingdom

OPEN ACCESS

Edited by:

Anatoliy V. Kharkhurin,
American University of Sharjah,
United Arab Emirates

Reviewed by:

Francesca Martina Branzi,
University of Manchester,
United Kingdom
Miriam Gade,
Medical School Berlin, Germany

*Correspondence:

Peter Bright
peter.bright@anglia.ac.uk

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 25 June 2018

Accepted: 06 September 2018

Published: 26 September 2018

Citation:

Naeem K, Filippi R, Periche-Tomas E,
Papageorgiou A and Bright P (2018)
The Importance of Socioeconomic
Status as a Modulator of the Bilingual
Advantage in Cognitive Ability.
Front. Psychol. 9:1818.
doi: 10.3389/fpsyg.2018.01818

Between-group variability in socioeconomic status (SES) has been identified as a potentially important contributory factor in studies reporting cognitive advantages in bilinguals over monolinguals (the so called “bilingual advantage”). The present study addresses the potential importance of this alternative explanatory variable in a study of low and high SES bilingual and monolingual performance on the Simon task and the Tower of London (TOL) task. Results indicated an overall bilingual response time advantage on the Simon task, despite equivalent error rates. Socioeconomic status was an important modulator in this effect, with evidence that bilingualism may be particularly important in promoting speed of processing advantages in low status individuals but have little impact in high status individuals. However, there was a *monolingual* advantage on the TOL test of executive planning ability. Together, our findings run counter to the central assertion of the bilingual advantage account, that the process of multi-language acquisition confers a broad cognitive advantage in executive function. We discuss these findings in the context of SES as an important modulator in published studies advocating a bilingual cognitive advantage.

Keywords: bilingual advantage, socioeconomic status, executive function, demographics, Simon task, Tower of London

INTRODUCTION

According to recent estimates, more of the world’s population today is bilingual or multilingual than monolingual (Grosjean, 2010; Paradis et al., 2011). The dominant belief amongst academics until the 1960s was that second language learning had detrimental effects on cognitive development, particularly verbal IQ (e.g., Saer, 1923), and second language learning was discouraged (Hakuta and Diaz, 1985). This view was gradually overturned following the publication of a large scale study of middle-class monolingual and balanced-bilingual children attending French primary schools in Canada (Peal and Lambert, 1962). On the basis of their findings, these authors claimed that bilinguals typically show better mental flexibility, superior concept formation and higher IQ. In particular, their work indicated that bilingualism can confer general cognitive advantages which are not restricted to linguistic processing. Nevertheless, socioeconomic status (SES) was inadequately addressed as a possible alternative explanatory variable distinguishing the monolingual and bilingual groups, and the possibility that any bilingual advantage might be explained by such uncontrolled variables has become an important debate.

The likelihood that bilingual environments place disproportionately challenging demands on the developing brain is intuitively attractive and plausible if we accept the claim that cognitive resources must be allocated to the inhibition of one language while thinking or communicating in the other. The argument follows that these additional inhibitory demands underpin the development of our cognitive resources in such a way that would not typically be observed in monolingual contexts. Confirmatory evidence focused on and highlighted a bilingual advantage in inhibitory control (e.g., Bialystok, 1982, 2001), a position substantiated by a wealth of published evidence (e.g., Frye et al., 1995; Zelazo et al., 1996; Bialystok, 1999; Zelazo et al., 2003).

Much of the evidence for the bilingual advantage is based on performance on the Simon task, in which participants respond to the color of a stimulus, ignoring its position on the computer display. Typically, the stimulus is either green or blue and can be presented on the right or left of central fixation. In congruent trials the correct response (left or right) is aligned with its spatial position, but in incongruent trials the stimulus color/response mapping is crossed such that presentation on the left requires a right motor response and presentations on the right requires a left motor response. Reaction times are generally shorter for congruent trials than incongruent ones (known as the Simon effect), but this disparity is typically smaller for bilinguals than for monolinguals (Bialystok, 2006). Age has been found to influence the size of the bilingual advantage on this test, with evidence that the effect is particularly strong in older adults, indicating that lifelong experience of managing two languages may attenuate the age-related decline in inhibitory processing (Bialystok et al., 2004). Furthermore, the advantage was observed not only on the incongruent trials, suggesting that bilingualism may confer cognitive enhancement beyond inhibitory control *per se* and generalize to executive function more generally. These results have been replicated in subsequent research (Bialystok, 2006; Bialystok et al., 2006; Costa et al., 2008; Martin-Rhee and Bialystok, 2008), which has allowed refinement and clarification of the bilingual advantage to encompass monitoring (Costa et al., 2009), task switching (Prior and Gollan, 2011), and working memory (Luo et al., 2013; Kerrigan et al., 2017). Such work has encouraged reconceptualization of the bilingual advantage in terms of conflict monitoring (Costa et al., 2009) and general mental flexibility (Kroll and Bialystok, 2013). Synthesizing the findings from 31 studies, Hilchey and Klein (2011) concluded that, rather than reflecting advantages in inhibitory control, the bilingual advantage is better characterized as a domain-general “global advantage” in monitoring conflict and regulating task demands, and this explains the faster overall response times on both congruent and incongruent trials in conflict resolution tasks.

In the last decade, much of the research favoring the bilingual advantage has come under increasing scrutiny, with claims of poor experimental control, particularly with respect to matching of potential confounding variables across monolingual and bilingual groups. In particular, authors have claimed that alternative explanatory variables for intergroup differences, such as SES and other demographic or cultural factors, have not

been systematically considered within study designs (for large-scale reviews see Paap et al., 2015; Lehtonen et al., 2018). Compounding this issue, evidence has been presented which indicates a lack of convergent validity across different tests developed to measure the same specific cognitive mechanisms thought to underpin the bilingual advantage (e.g., inhibitory control; Paap and Greenberg, 2013). Furthermore, another recent meta-analysis raises the complication that evidence for a bilingual advantage has been overplayed in the literature because of the tendency for journals to favor positive rather than null effects (i.e., publication bias; De Bruin et al., 2015b).

Whether the process of acquiring a second language confers a genuine cognitive advantage remains fiercely debated in the literature, and the counterclaim that factors independent of multi-language acquisition, such as SES, offer more plausible and parsimonious explanations for group differences in test performance is increasingly reported (e.g., Morton and Harper, 2007; Paap and Greenberg, 2013; Antón et al., 2014; Duñabeitia et al., 2014; Gathercole et al., 2014; Paap et al., 2015; Von Bastian et al., 2016; D’Souza et al., 2018; Goldsmith and Morton, 2018). In one of the earliest studies to question modern conceptualization of the bilingual advantage, Morton and Harper (2007) single out SES as a particularly important variable. They found a clear cognitive control advantage in children from high SES families relative to those from low SES families, but no evidence for performance differences between bilingual and monolingual children from the same socioeconomic backgrounds (see also Noble et al., 2005). Participants with low SES, particularly in young adult populations, are underrepresented in the literature on the bilingual advantage. However, one recent study has focused on non-linguistic executive control in Greek-Albanian young adult bilinguals from underprivileged social contexts, finding no bilingual advantage in interference control (Vivas et al., 2017).

In the present study we examine the effects of bilingualism and multilingualism on executive function in low and high SES, age-matched participants, addressing whether cognitive performance in young adults from underprivileged, low SES backgrounds might be disproportionately sensitive to factors associated with multilanguage acquisition. Given the interest of SES as an alternative explanatory variable for the bilingual cognitive advantage, we established this for each participant using stringent measurement criteria. The low SES bilingual group was composed of first-generation immigrants, half of whom had refugee and/or asylum seeker status. We employed two widely used tests in the literature on bilingual cognition and executive function, the Simon task and the Tower of London (TOL) task. To the extent that bilingualism, regardless of SES, confers an advantage in response inhibition, we predicted that bilinguals would perform disproportionately well on incongruent (conflict condition) relative to congruent (non-conflict condition) trials on the Simon task. Conversely, if bilingualism is associated with a more global cognitive monitoring advantage, they should perform proportionately better on both congruent and incongruent trials, relative to monolinguals. We also predicted that, if the bilingual advantage extends to planning and sustained cognitive control of behavior

toward a goal, bilinguals should perform better on the TOL task.

MATERIALS AND METHODS

Participants

The participants consisted of 90 adults aged between 18 and 30 years at the time of testing, of whom 45 were monolingual and 45 bilingual. Within each of these language groups, 20 had low SES and 25 high economic status, calculated on the basis of employment status and history, education and income. Age was statistically equivalent across language [$F(1, 86) = 0.08$, $p = 0.76$, $\eta^2 = 0.001$] and SES [$F(1, 86) = 0.19$, $p = 0.66$, $\eta^2 = 0.002$] groups and the language by SES interaction effect was negligible [$F(1, 86) = 0.039$, $p = 0.84$, $\eta^2 = 0.000$]. With respect to background cognitive performance, the language groups were equivalent on the Raven's Matrices test of fluid intelligence [$F(1, 86) = 0.095$, $p = 0.76$, $\eta^2 = 0.001$], digit span forward [$F(1, 86) = 0.87$, $p = 0.35$, $\eta^2 = 0.007$] and backwards [$F(1, 86) = 0.05$, $p = 0.82$, $\eta^2 = 0.000$] and although there was a highly significant main effect of SES ($p < 0.001$ in all cases), there were no significant language group by SES interaction effects ($p = 0.76$, $\eta^2 = 0.001$; $p = 0.96$, $\eta^2 = 0.000$; $p = 0.16$, $\eta^2 = 0.019$, respectively).

All low SES participants attended government-funded vocational courses at the same college in a predominantly low socioeconomic area in London, where they were recruited for participation in the current study. Although the monolingual controls were born and educated in the United Kingdom, the low SES bilinguals were immigrants of which half ($n = 10$) reported having refugee or asylum seeker status. The low SES participants were in receipt of financial social support, which was a condition for their participation and group allocation. High SES participants were recruited, using opportunity sampling, from local university and professional sectors in London.

Irrespective of SES (high/low), participants received their education in English, but the bilinguals spoke a language or languages other than English at home, and the majority reported using predominantly English to communicate outside of the home². Among the low SES bilinguals, 18 claimed proficiency in a third language, 8 in a fourth language and 4 in a fifth language. Among the high SES bilinguals, 8 claimed proficiency in a third language and 1 in a fourth language. The monolinguals were not functionally proficient in any language other than English despite foreign language instruction in school.

All participants completed a language history questionnaire adapted from Li et al. (2006) and used in earlier studies by

TABLE 1 | Bilingual participants' language history information.

	Low SES	SD	High SES	SD
Years learning L2	11.75	5.71	16.72	5.41
Self-rated L2 literacy	15.00	3.81	15.80	3.61
Self-rated L2 proficiency	14.50	3.75	17.40	2.38

Sum of scores for L2 Reading and Writing was used to achieve self-rated L2 literacy, and Speaking and Comprehension to achieve self-rated L2 proficiency (each of the four skills were measured on a scale 1–10, where 1 = not literate and 10 = highly literate/proficient).

Filippi et al. (2012, 2015), which gathered language background and biographical information. The questionnaire items included details of employment, education and income to achieve a summary of SES. Additionally, bilingual participants provided language-related information, such as the number of languages acquired, years learning the second language, and individual self-rated competence in each language. Both objective information (e.g., years spent learning the second language) and subjective ratings on reading, writing, speaking and comprehension abilities indicated that all participants categorized as bilingual, irrespective of SES, were highly proficient in at least two languages (Table 1).

This project was reviewed and approved by the UCL Institute of Education Research Ethics Committee. All participants gave informed consent prior to testing.

Tasks

In addition to the tests of background cognitive ability (Raven's Matrices, digit span forward and backwards), all participants were administered the Simon Test and the TOL task:

Simon Task

A computerized version of the Simon task (Simon and Wolf, 1963) was implemented in E-Prime (version 2.0; Schneider et al., 2002, 2007) and administered to all participants to assess inhibitory control based on stimulus-response conflict. The experiment was presented on a laptop computer with a 15.6-inch monitor and a two-button USB keypad connected to the laptop. Each trial began with a fixation cross (+) in the middle of the display that remained visible for 500 ms and was followed by a filled blue or red star (height = 1.7 cm, width = 1.8 cm on screen) displayed 3.9° to the left or right of the fixation point. The goal was to press the corresponding key as quickly as possible according to the color of the star, which was presented for 1000 ms. The blue star was associated with the right index finger key on the keypad, whereas the red star was associated with the left index finger key. Blue and red dots were placed directly above the corresponding keys. Participants rested their index fingers on these keys and were instructed to press the key on the correct side according to the color of the stimulus, regardless of its position on the screen. Trials were defined as congruent if the color stimulus matched the key position (e.g., red star appearing on the left side of the screen required a left key response), and incongruent, when the color stimulus did not match the key position (e.g., red star appearing on the right side of the screen required a left key response). Participants scored

¹For all our reported ANOVA effects we manually calculated the eta-squared statistic (η^2) as a measure of effect size rather than partial eta-squared, automatically provided in the SPSS statistical package. Apart from one-way ANOVAs (where the two values will be identical), η^2 provides a more conservative estimate of the effect size.

²The following second languages were represented in our bilingual group: Amharic, Arabic, Azeri, Bangla, Chinese, Darry, Farsi, French, Fur, Greek, Gujarati, Hindi, Hungarian, Irish, Italian, Koka, Kurdish, Lingala, Malay, Mauritian, Polish, Portuguese, Punjabi, Romanian, Russian, Spanish, Tagalog, Twi, Urdu, Zahu.

one point when they pressed the correct key, with failure to respond within the 1000 ms stimulus presentation time classified as an error. There were in total 36 sequential randomized test trials, 18 congruent and 18 incongruent, with no practice trials. Raw scores were recorded as response times (RTs) and accuracy (proportion correct) for congruent and incongruent trial types.

Tower of London Task

A computerized version of the classic TOL task was administered to assess planning and problem solving (Berg and Byrd, 2002). In this version, the participants had to move colored discs on three pegs of different height to solve 12 problems of increasing difficulty in a fixed number of moves per trial (PEBL software, cf., Mueller and Piper, 2014). The computerized TOL instrument consisted of three piles of different height, the first of which could hold three discs, the second two discs, and the last only one disc. On each trial, the participant was shown a target disc configuration (top panel) and a start configuration (lower panel), each of which displayed three differently colored discs distributed across the three piles. The participant was required to move the discs in the lower panel to match the target configuration using the computer mouse. The number of possible moves was presented on a bar on the side of the screen, which reduced with each complete move. Twelve problems were presented, beginning with those that could be solved in two moves and progressing to those that required five moves. The trial was considered as successful if the solution was correctly submitted within the set number of moves. If the maximum number of moves was reached (irrespective of trial success) that trial terminated and the participant progressed onto the next problem. Scores were recorded as accuracy rates, the number of trials successfully solved, mean first-move latency, calculated as the length of time between the problem presentation and the first move, and mean total trial time.

Design and Procedure

Participants were tested individually. The tasks were presented to all participants in a single session, which lasted between 40 min and 1 h including as many breaks between tasks as the participants wished to take. The order of the tests was as follows: Raven's Progressive Matrices, digit span forward, digit span backward, Simon task, and TOL task. Raw data is provided online in **Supplementary Table 1**.

Materials

All tasks were presented on a laptop computer. Responses for the background measures were recorded by the experimenter on a scoring sheet. Simon and TOL data scores (response times and accuracy) were automatically recorded using E-Prime 2.0 software (Schneider et al., 2007) and stored electronically in a password-protected file. Additionally, the Simon task required the use of a Logitech Gamepad (model F310) and the TOL task was completed using an HP wireless computer mouse (model X3000) to ensure accuracy and ease of navigation.

RESULTS

Simon Task Performance

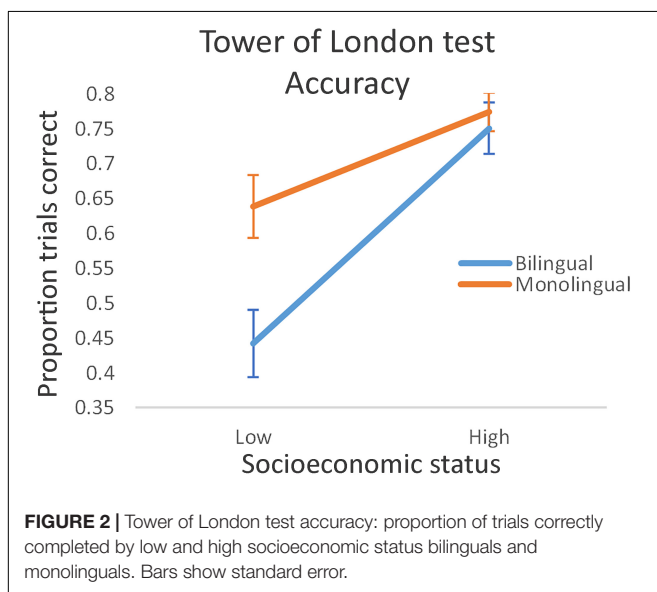
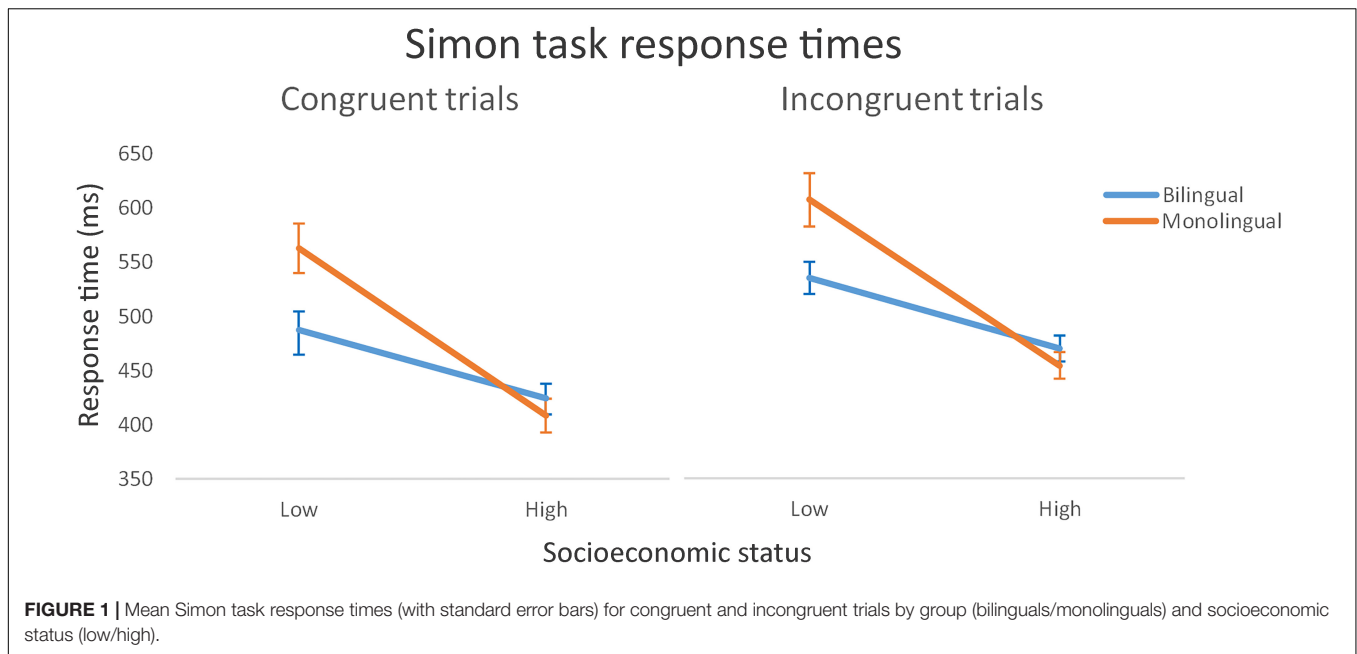
We applied two three-way mixed ANOVA models, one on response times and one on accuracy, with congruency as a within-subjects variable (congruent/incongruent) and language group (monolingual/bilingual) and SES (low/high) as between-subjects variables. The analysis of response times identified a very robust main effect of congruency (i.e., a Simon effect), with longer response times on incongruent trials [$F(1, 86) = 110.71$, $p < 0.001$, $\eta^2 = 0.563$] but negligible congruency \times language group [$F(1, 86) = 0.06$, $p = 0.81$, $\eta^2 = 0.000$], congruency \times SES [$F(1, 86) = 0.02$, $p = 0.88$, $\eta^2 = 0.000$] and congruency \times language group \times SES [$F(1, 86) = 0.02$, $p = 0.9$, $\eta^2 = 0.000$] interaction effects.

There was a marginal main effect of language group [$F(1, 86) = 3.236$, $p = 0.08$, $\eta^2 = 0.022$], with shorter response times in the BL group. The main effect of SES was, however, highly significant [$F(1, 86) = 47.19$, $p < 0.001$, $\eta^2 = 0.326$], with high SES associated with shorter response times. The language group \times SES interaction effect was also significant [$F(1, 86) = 8.17$, $p = 0.005$, $\eta^2 = 0.056$]. The discrepancy in reaction times between low and high SES participants was disproportionately wider in monolinguals, indicating that the importance of SES in driving response times on the Simon task may be greater in monolinguals (**Figure 1**). Of particular interest here was the observation that although high SES MLs and BLs produced statistically equivalent response times [$F(1, 48) = 0.87$, $p = 0.36$, $\eta^2 = 0.018$], low SES MLs produced statistically longer response times than low SES BLs [$F(1, 38) = 7.22$, $p = 0.011$, $\eta^2 = 0.160$].

Mean accuracy performance was at/close to ceiling for congruent trials (0.966) but lower for incongruent trials (0.887). This effect of congruency was highly significant [$F(1, 86) = 33.56$, $p < 0.001$, $\eta^2 = 0.278$] but there were negligible congruency \times language group [$F(1, 86) = 0.02$, $p = 0.89$, $\eta^2 = 0.000$], congruency \times SES [$F(1, 86) = 0.56$, $p = 0.46$, $\eta^2 = 0.005$] and congruency \times language group \times SES [$F(1, 86) = 0.49$, $p = 0.49$, $\eta^2 = 0.004$] interaction effects. Accuracy performance was statistically equivalent across language groups [$F(1, 86) = 0.55$, $p = 0.46$, $\eta^2 = 0.006$] and SES groups [$F(1, 86) = 1.85$, $p = 0.178$, $\eta^2 = 0.021$] and the language group \times SES interaction effect was also non-significant [$F(1, 86) = 1.11$, $p = 0.296$, $\eta^2 = 0.012$].

Tower of London Performance

We applied two-way between groups analysis of variance models on accuracy, planning time and total response time. Each was specified with language group (monolingual/bilingual) and SES (high/low) as the between-subjects variables. There was a significant main effect of language group on accuracy (proportion of trials correct), with monolinguals outperforming bilinguals [$F(1, 86) = 7.87$, $p = 0.006$, $\eta^2 = 0.060$]. There was also a highly significant main effect of SES, with high status conferring the accuracy advantage [$F(1, 86) = 32.33$, $p < 0.001$, $\eta^2 = 0.247$]. The SES \times language group interaction effect was also significant [$F(1, 86) = 4.88$, $p = 0.030$, $\eta^2 = 0.037$]. The difference in performance between low and high SES participants was disproportionately

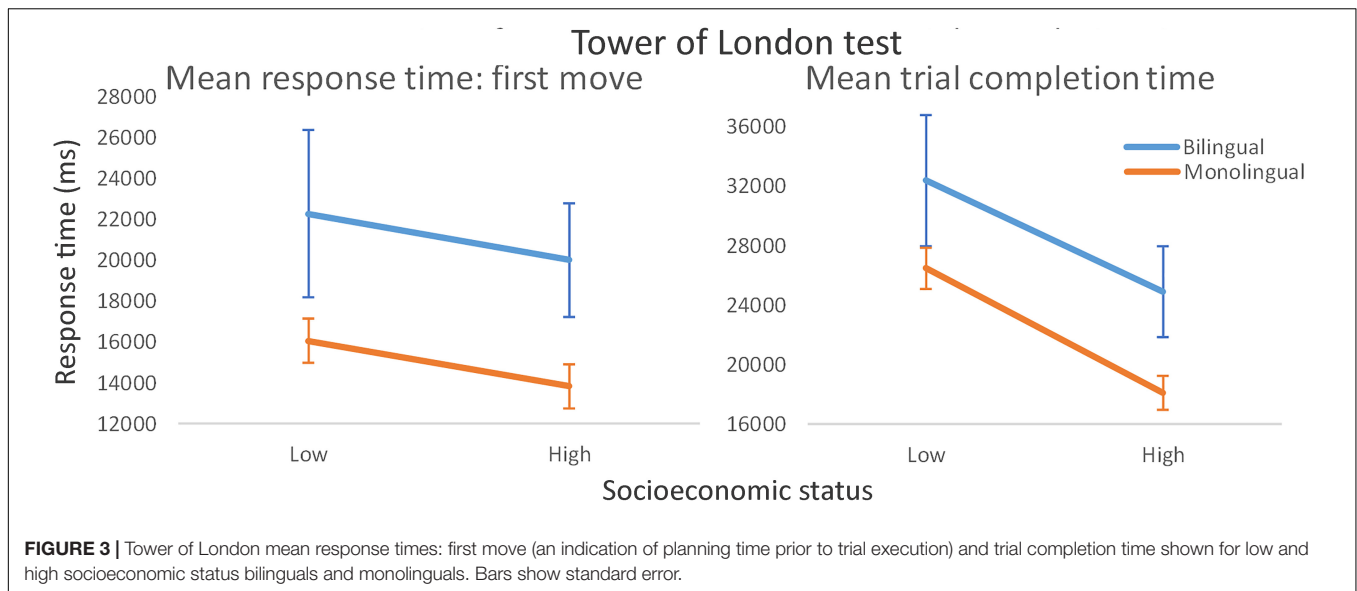


large in bilinguals, with the low status bilingual participants failing to successfully complete more than half the trials on average (see **Figure 2**). Simple effects analysis confirmed the disproportionately poor performance in low SES bilinguals relative to low SES monolinguals [$F(1, 38) = 8.79, p = 0.005, \eta^2 = 0.188$] and statistically equivalent high SES bilingual and monolingual performance [$F(1, 48) = 0.26, p = 0.61, \eta^2 = 0.006$].

Time taken to produce the first move (an indication of solution planning time prior to execution) was compared across language and SES groups. Bilingual participants took significantly longer on this measure [$F(1, 86) = 6.05, p = 0.016, \eta^2 = 0.065$], which when considered in the context of poorer overall accuracy, is clearly inconsistent with claims

that bilingualism confers a broad intellectual advantage. The evidence against bilingual advantage theory is compounded by our observation that monolinguals also produced a shorter mean trial response time across the 12 trials [$F(1, 86) = 5.32, p = 0.024, \eta^2 = 0.053$]. There was a main effect of SES on mean trial completion time, with faster timings produced by high SES participants [$F(1, 86) = 8.31, p = 0.005, \eta^2 = 0.083$] but no main effect for first move response time [$F(1, 86) = 0.80, p = 0.37, \eta^2 = 0.009$]. Language Group \times SES interaction effects were non-significant in both cases ($p > 0.8$; **Figure 3**).

In summary, the bilingual advantage emerged in a marginal overall speed advantage in controlling interference (Simon task performance), but not in higher order cognitive processes involved in planning and problem-solving, as engaged by the TOL task. Socioeconomic status was identified as an important predictor of task performance in both tasks, and there was evidence from the Simon task that bilingualism may offset the response time disadvantage associated with low SES. Nevertheless, in our data, bilingualism conferred a performance disadvantage on the TOL test of planning and problem solving, and the accuracy disadvantage was particularly acute in those bilinguals with low SES, a finding incompatible with claims that multilanguage acquisition is associated with advantages in general mental flexibility and executive function. Although self-rated L2 proficiency was higher in high SES bilinguals [$t(43) = 3.22, p = 0.002$], correlations of proficiency with Simon and TOL task performance were negligible ($p > 0.1$ in all cases), and statistically controlling for this measure did not meaningfully alter the size of our reported effects. Correlational analyses conducted across the Simon and TOL tests revealed consistently small effect sizes, the largest of which ($r = 0.271$), between incongruent response times on the Simon task and total response time on the



TOL test, revealed approximately 7% shared variance in these measures. Although correlations of performance on these tests with non-verbal general ability (as measured by Raven's Matrices) were statistically equivalent across the monolingual and bilingual groups ($p > 0.2$ in all comparisons), these findings suggest that the Simon test and the TOL test may tap different mechanisms of cognitive control, and that these may be differentially influenced by the process of becoming bilingual.

DISCUSSION

The present study examined the effects of bilingualism on cognitive control and higher order executive function in low and high SES young adult bilinguals and monolinguals. We found a monolingual advantage in performance on the TOL task, which is not only incompatible with claims that bilingualism confers a general cognitive advantage in executive function, but infers that it may in fact obstruct the development of planning and goal-directed strategy formation. However, results from the Simon task indicate a marginal bilingual advantage in response times, irrespective of congruency (i.e., whether or not there was a strong demand on response inhibition), but that the advantage is modulated by SES. Our data raise the potentially important and intriguing possibility that multilanguage acquisition may be unimportant in high SES populations, but may help offset the negative impact that impoverished, low socioeconomic conditions have for the development of cognitive mechanisms underpinning information processing. In our study, low-status bilinguals showed significantly faster response times than low-status monolinguals, a pattern that was not replicated in high-status participants.

These findings help clarify the role of SES as a modulating influence on the likelihood that multilanguage acquisition will lead to cognitive advantages. The implication in the context of the size of our observed effects is that SES is the more important variable driving observed advantages, but that multilingual contexts may also be of significant benefit in environments in which access to economic, recreational and educational opportunities are relatively constrained. The advantage, however, appears to be quite specific. Low SES bilinguals performed disproportionately *poorly* on the TOL task (trials correctly completed) but high-status bilingual and monolingual participants' performance was statistically equivalent. Bilinguals were also slower to produce their first move on this test (an indication of the need for longer planning time prior to execution) and to complete each trial. Therefore, a fractionation among components associated with cognitive control was apparent: bilinguals outperformed monolinguals on a task requiring monitoring and responding to compatible and incompatible stimulus-response mappings (the Simon task) and monolinguals outperformed bilinguals on a classic test requiring goal-directed strategic thinking and planning (the TOL test). While these observations cannot strictly be considered a double dissociation (all comparison groups are independent), we find it intriguing that well-matched groups undertaking the same tests under equivalent conditions have presented with a reversal of comparative performance as a function of our primary variable of interest (monolingualism/multilingualism). Our finding that SES influences the effect of multilanguage acquisition on performance in one of these tests but not the other, further complicates our ability to conceptualize the "bilingual advantage."

How should we characterize, separate and distinguish between the cognitive mechanisms associated with the Simon and TOL tasks? Like the Simon task (and the Stroop test), the visual Flanker task incorporates the demand to suppress a prepotent/habitual response tendency (i.e., there is an

incongruent stimulus/response mapping) which is compared with a non-conflicting/congruent response. Costa et al. (2009) employed versions of a flanker task which varied in their monitoring demands in young adult bilingual and monolingual university students and observed an overall bilingual speed advantage in the high-monitoring but not low-monitoring conditions, leading the authors to attribute the advantage to a more effective or efficient monitoring process (rather than, for example, an advantage in inhibitory control). Our data are also inconsistent with the inhibitory control explanation of the bilingual advantage, given that we observed virtually identical trends across monolinguals and bilinguals in both the congruent and incongruent Simon test conditions (a finding robustly supported in a large scale review by Hilchey and Klein, 2011). The model proposed by Costa et al. (2009) attributes the advantage to an enhanced cognitive flexibility for switching between contrasting demands associated with different task conditions (perhaps consistent with the way bilinguals disengage and engage between languages contingent upon social context). The authors further develop their theoretical framework by claiming that this monitoring advantage might incorporate an ongoing evaluation of the likely requirement for active attentional control (e.g., response suppression) given current task demands. That is, the real time processing advantage associated with bilingualism may occur before conflict resolution mechanisms are triggered.

The present findings are, in part, consistent with the kind of bilingual monitoring advantage described by Costa et al. (2009), but indicate that the capacity for bilingualism to confer such an advantage is mitigated by situational conditions associated with SES. On the Simon task, only those with demonstrably low status benefitted from being bilingual, and the fact that, while the disparity in response times between low and high SES was smaller in bilinguals than monolinguals, high SES participants still responded numerically faster. It follows that SES appears to be a more important predictor of cognitive performance (as gauged by this task) than whether or not a person is bilingual. Nevertheless, implications for society of a significant beneficial cognitive impact in low socioeconomic populations are considerable, and we therefore recommend further studies employing a broader range of tasks and larger numbers of trials to examine the replicability of this finding and to further characterize the relationship.

The advantages observed on the Simon task did not transfer to TOL test performance: bilinguals were consistently slower in planning the moves required to match the target disk configuration and in executing those moves, and this was the case irrespective of SES. Compounding this evidence against any bilingual advantage in complex goal-relevant planning was the observation of disproportionately poor accuracy performance in low SES bilinguals. These findings are, in part, consistent with a study of simple and complex Simon task performance which indicated that bilingualism conferred advantages in selective attention specifically in the context of low working memory demand (Salvatierra and Rosselli, 2011). Other studies have reported equivalent monolingual

and bilingual performance on the TOL test (e.g., De Bruin et al., 2015a; Cox et al., 2016) but, to our knowledge, the present study is the first to clearly indicate a disadvantage in a bilingual group. We suggest that the most likely reason for this disparity is that our study is also the first to explicitly recruit participants from the lowest level of SES (like the Cox et al., and de Bruin et al. studies, we observed similar performance in our other (i.e., high SES) monolingual and bilingual groups). Nevertheless, it is also possible that other experience-related factors operating in this group (half of whom were asylum seekers) underpinned the patterns of performance reported here, and more formal assessment of language fluency within and across comparison groups is encouraged. We also note recent evidence that, in young economically disadvantaged bilingual children with low proficiency in both languages, a stronger performance advantage over monolinguals was observed in tasks incorporating higher relative to lower cognitive control demands (Engel de Abreu et al., 2012).

We have recently reported evidence for a bilingual disadvantage in metacognitive processing (Folke et al., 2016), in which we employed a two-alternative-forced-choice task which required participants to determine which of two visually presented circles contained the most dots (with task difficulty systematically manipulated) and then state their confidence in their choice. We found that bilinguals were comparatively less confident on correctly completed trials and more confident on trials completed incorrectly. While purely speculative, one possible explanation for the patterns of TOL accuracy performance in the present study is that the cumulative effect of low SES and bilingualism might underpin comparatively low confidence in ongoing ability on this test, which, in turn, impacts on actual performance. In other words, if accurate monitoring of ongoing performance is not possible (i.e., on tasks in which our subjective assessment of our cognitive performance is poorly calibrated with objective performance) we cannot optimally regulate our knowledge or strategies in the service of goal attainment (see Bright et al., 2018, for further discussion of this theme). The TOL test is considerably more complex than the Simon test, incorporating strategic planning in order to determine moves that will bring the current disk configuration closer to the goal/target configuration, and subgoal conflicts, in which counterintuitive moves away from the goal state are sometimes required. This level of complexity, we would argue, renders performance on this test considerably more likely to be sensitive to the effects of poor metacognitive processing than is the case for the Simon test, which is operationally straightforward (i.e., restricted to processing binary congruent and incongruent stimulus/response mappings).

In summary, our findings are inconsistent with the claim that the process of acquiring a second language confers broad advantages in executive function. Instead, any cognitive advantage appears to relate to basic processing efficiency and is both contingent upon – and of secondary importance to – SES. Furthermore, this advantage may be offset by disadvantages in more complex tasks with stronger strategic and forward planning demands.

We encourage further efforts toward isolating specific cognitive mechanisms that may be modulated positively or negatively through the process of multilanguage acquisition, and to carefully consider the moderating influence of situational, demographic and other factors.

AUTHOR CONTRIBUTIONS

KN collected the data and provided an early draft for this work, under the supervision of RF. RF and PB contributed equally to subsequent theoretical development, data analysis, and write-up for this paper. PB wrote the original submission and post-review revisions. EP-T and AP contributed to additional data collection and the

manuscript editing. All authors read and approved the manuscript.

FUNDING

This work was supported by the Leverhulme Trust UK (RPG-2015-024) and the British Academy (SG162171).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01818/full#supplementary-material>

REFERENCES

- Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., et al. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Front. Psychol.* 5:398. doi: 10.3389/fpsyg.2014.00398
- Berg, K. W., and Byrd, D. L. (2002). The tower of London spatial problem-solving task: enhancing clinical and research implementation. *J. Clin. Exp. Neuropsychol.* 24, 586–604. doi: 10.1076/jcen.24.5.586.1006
- Bialystok, E. (1982). On the relationship between knowing and using linguistic forms. *Appl. Linguist.* 3, 181–206. doi: 10.1093/applin/3.3.181
- Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Dev.* 70, 636–644. doi: 10.1111/1467-8624.00046
- Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy and Cognition*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511605963
- Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Can. J. Exp. Psychol.* 60, 68–79. doi: 10.1037/cjep2006008
- Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290
- Bialystok, E., Craik, F. I., and Ryan, J. (2006). Executive control in a modified antisaccade task: effects of aging and bilingualism. *J. Exp. Psychol. Learn.* 32:1341. doi: 10.1037/0278-7393.32.6.1341
- Bright, P., Ouzia, J., and Filippi, R. (2018). “Metacognition and multilingualism,” in *The Handbook of the Neuroscience of Multilingualism: Blackwell Handbooks in Linguistics Series*, ed. J. W. Schweiter (Hoboken, NJ: Wiley-Blackwell).
- Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86. doi: 10.1016/j.cognition.2006.12.013
- Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don't. *Cognition* 113, 135–149. doi: 10.1016/j.cognition.2009.08.001
- Cox, S. R., Bak, T. H., Allerhand, M., Redmond, P., Starr, J. M., Deary, I. J., et al. (2016). Bilingualism, social cognition and executive functions: a tale of chickens and eggs. *Neuropsychologia* 91, 299–306. doi: 10.1016/j.neuropsychologia.2016.08.029
- De Bruin, A., Bak, T. H., and Della Sala, S. (2015a). Examining the effects of active versus inactive bilingualism on executive control in a carefully matched non-immigrant sample. *J. Mem. Lang.* 85, 15–26. doi: 10.1016/j.jml.2015.07.001
- De Bruin, A., Treccani, B., and Della Sala, S. (2015b). Cognitive advantage in bilingualism: an example of publication bias? *Psychol. Sci.* 26, 99–107. doi: 10.1177/0956797614557866
- D'Souza, A. A., Moradzadeh, L., and Wiseheart, M. (2018). Musical training, bilingualism, and executive function: working memory and inhibitory control. *Cogn. Res. Princ. Implic.* 3:11. doi: 10.1186/s41235-018-0095-6
- Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizao, P., Estévez, A., Fuentes, L. J., et al. (2014). The inhibitory advantage in bilingual children revisited: myth or reality? *Exp. Psychol.* 61, 234–251. doi: 10.1027/1618-3169/a000243
- Engel de Abreu, P. M., Cruz-Santos, A., Tourinho, C. J., Martin, R., and Bialystok, E. (2012). Bilingualism enriches the poor: enhanced cognitive control in low-income minority children. *Psychol. Sci.* 23, 1364–1371. doi: 10.1177/0956797612443836
- Filippi, R., Leech, R., Thomas, M., Green, D., and Dick, F. (2012). A bilingual advantage in controlling language interference during sentence comprehension. *Biling. Lang. Cogn.* 15, 858–872. doi: 10.1017/S1366728911000708
- Filippi, R., Morris, J., Richardson, F., Bright, P., Thomas, M., Karmiloff-Smith, A., et al. (2015). Bilingual children show an advantage in controlling verbal interference during spoken language comprehension. *Biling. Lang. Cogn.* 18, 490–501. doi: 10.1017/S1366728914000686
- Folke, T., Ouzia, J., Bright, P., De Martino, B., and Filippi, R. (2016). A bilingual disadvantage in metacognitive processing. *Cognition* 150, 119–132. doi: 10.1016/j.cognition.2016.02.008
- Frye, D., Zelazo, P. D., and Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cogn. Dev.* 10, 483–527. doi: 10.1016/0885-2014(95)90024-1
- Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Viñas-Guash, N., et al. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Front. Psychol.* 5:11. doi: 10.3389/fpsyg.2014.00011
- Goldsmith, S. F., and Morton, J. B. (2018). Time to disengage from the bilingual advantage hypothesis. *Cognition* 170, 328–329. doi: 10.1016/j.cognition.2017.09.011
- Grosjean, F. (2010). *Bilingual: Life and Reality*. Cambridge, MA: Harvard University Press. doi: 10.4159/9780674056459
- Hakuta, K., and Diaz, R. M. (1985). “The relationship between bilingualism and cognitive ability: a critical discussion and some new longitudinal data,” in *Children's Language*, Vol. 5, ed. K. E. Nelson (Hillsdale, NJ: Erlbaum), 319–344.
- Hilchey, M. D., and Klein, R. M. (2011). Are there bilingual advantages on non-linguistic interference tasks? Implications for the plasticity of executive control processes. *Psychon. Bull. Rev.* 18, 625–658. doi: 10.3758/s13423-011-0116-7
- Kerrigan, L., Thomas, M. S. C., Bright, P., and Filippi, R. (2017). Evidence of an advantage in visual-spatial memory for bilingual compared to monolingual speakers. *Biling. Lang. Cogn.* 20, 602–612. doi: 10.1017/S1366728915000917
- Kroll, J. F., and Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *J. Cogn. Psychol.* 25, 497–514. doi: 10.1080/20445911.2013.799170
- Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., and Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychol. Bull.* 144, 394–425. doi: 10.1037/bul0000142
- Li, P., Sepanski, S., and Zhao, X. (2006). Language history questionnaire: a web-based interface for bilingual research. *Behav. Res. Methods* 38, 202–210. doi: 10.3758/BF03192770

- Luo, L., Craik, F. I. M., Moreno, S., and Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: evidence from aging. *Psychol. Aging* 28, 28–34. doi: 10.1037/a0030875
- Martin-Rhee, M. M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Biling. Lang. Cogn.* 11, 81–93. doi: 10.1017/S1366728907003227
- Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x
- Mueller, S. T., and Piper, B. J. (2014). The psychology experiment building language (pebl) and pebl test battery. *J. Neurosci. Methods* 222, 250–259. doi: 10.1016/j.jneumeth.2013.10.024
- Noble, K. G., Norman, M. F., and Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Dev. Sci.* 8, 74–87. doi: 10.1111/j.1467-7687.2005.00394.x
- Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Paap, K. R., Johnson, H. A., and Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69, 265–278. doi: 10.1016/j.cortex.2015.04.014
- Paradis, J., Genesee, F., and Crago, M. (2011). *Dual Language Development and Disorders: A Handbook on Bilingualism and Second Language Learning*. Baltimore, MD: Paul H. Brookes Publishing.
- Peal, E., and Lambert, W. E. (1962). *The Relation of Bilingualism to Intelligence*. Washington, DC: American Psychological Association.
- Prior, A., and Gollan, T. H. (2011). Good language-switchers are good task-switchers: evidence from Spanish-English and Mandarin-English bilinguals. *J. Int. Neuropsychol. Soc.* 17, 682–691. doi: 10.1017/S1355617711000580
- Saer, D. J. (1923). The effects of bilingualism on intelligence. *Br. J. Psychol.* 14, 25–38.
- Salvatierra, J. L., and Rosselli, M. (2011). The effect of bilingualism and age on inhibitory control. *Int. J. Biling.* 15, 26–37. doi: 10.1177/1367006910371021
- Schneider, W., Eschman, A., and Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh, PA: Psychology Software Tools Inc.
- Schneider, W., Eschman, A., and Zuccolotto, A. (2007). *E-Prime 2.0 Professional*. Pittsburgh, PA: Psychology Software Tools.
- Simon, J. R., and Wolf, J. D. (1963). Choice reaction times as a function of angular stimulus-response correspondence and age. *Ergonomics* 6, 99–105. doi: 10.1080/00140136308930679
- Vivas, A. B., Ladas, A. I., Salvatierra, J. L., and Chrysoschoou, E. (2017). Revisiting the bilingual advantage in attention in low SES Greek-Albanians: does the level of bilingual experience matter? *Lang. Cogn. Neurosci.* 32, 743–756. doi: 10.1080/23273798.2016.1271442
- Von Bastian, C. C., Souza, A. S., and Gade, M. (2016). No evidence for bilingual cognitive advantages: a test of four hypotheses. *J. Exp. Psychol. Gen.* 145:246. doi: 10.1037/xge0000120
- Zelazo, P. D., Frye, D., and Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cogn. Dev.* 11, 37–63. doi: 10.1016/S0885-2014(96)90027-1
- Zelazo, P. D., Müller, U., Frye, D., and Marcovitch, S. (2003). The development of executive function. *Monogr. Soc. Res. Child* 68, 7–8. doi: 10.1111/j.0037-976X.2003.00261.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Naeem, Filippi, Periche-Tomas, Papageorgiou and Bright. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Role of Cognitive Development and Strategic Task Tendencies in the Bilingual Advantage Controversy

Esli Struys^{1,2,3*}, Wouter Duyck⁴ and Evy Woumans^{4*}

¹ Centre for Linguistics, Vrije Universiteit Brussel, Brussels, Belgium, ² Brussels Institute for Applied Linguistics, Vrije Universiteit Brussel, Brussels, Belgium, ³ Center for Neurosciences, Vrije Universiteit Brussel, Brussels, Belgium, ⁴ Department of Experimental Psychology, Ghent University, Ghent, Belgium

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Natsuki Atagi,
University of California, Riverside,
United States
John W. Schwieter,
Wilfrid Laurier University, Canada

*Correspondence:

Esli Struys
estruys@vub.ac.be
Evy Woumans
evy.woumans@ugent.be

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 30 July 2018

Accepted: 04 September 2018

Published: 25 September 2018

Citation:

Struys E, Duyck W and Woumans E
(2018) The Role of Cognitive
Development and Strategic Task
Tendencies in the Bilingual Advantage
Controversy. *Front. Psychol.* 9:1790.
doi: 10.3389/fpsyg.2018.01790

Recent meta-analyses have indicated that the bilingual advantage in cognitive control is not clear-cut. So far, the literature has mainly focussed on behavioral differences and potential differences in strategic task tendencies between monolinguals and bilinguals have been left unexplored. In the present study, two groups of younger and older bilingual Dutch–French children were compared to monolingual controls on a Simon and flanker task. Beside the classical between-group comparison, we also investigated potential differences in strategy choices as indexed by the speed-accuracy trade-off. Whereas we did not find any evidence for an advantage for bilingual over monolingual children, only the bilinguals showed a significant speed-accuracy trade-off across tasks and age groups. Furthermore, in the younger bilingual group, the trade-off effect was only found in the Simon and not the flanker task. These findings suggest that differences in strategy choices can mask variations in performance between bilinguals and monolinguals, and therefore also provide inconsistent findings on the bilingual cognitive control advantage.

Keywords: bilingualism, cognitive control, inhibition, speed-accuracy trade-off, choice strategy

INTRODUCTION

The bilingual advantage in cognitive control assumes that bilinguals outperform monolinguals in conflict tasks, such as the Simon or flanker, due to their continued practice in handling between-language competition (for a recent review, see Zhou and Krott, 2016). These tasks typically contain a mixture of non-conflict (i.e., congruent) and conflict (i.e., incongruent) trials. Performance is consistently slower or less accurate for the latter (for a review study on these effects, see Lu and Proctor, 1995). Despite the general label of an *advantage*, the reported benefits for bilinguals are actually quite diverse (Hilchey and Klein, 2011), and not very consistent across studies: sometimes, they show better performance only on incongruent trials, but not on congruent trials (e.g., Schroeder and Marian, 2012; Marzecova et al., 2013; Pelham and Abrams, 2014); at other times, they outperform monolinguals on overall performance (e.g., Costa et al., 2009; Kapa and Colombo, 2013; Morales et al., 2013). And yet, there are also studies showing a combination of both (Bialystok et al., 2004; Tao et al., 2011; Yang et al., 2011).

Besides the varying manifestation of effects, bilingual benefits have become highly controversial because of repeated failures to replicate this superior performance altogether (e.g., Paap et al., 2015; von Bastian et al., 2016; de Bruin and Della Sala, 2017; Paap, in press). This has even led to the

assertion that there is no coherent evidence for a bilingual advantage in cognitive control (Paap and Greenberg, 2013). Still, the lack of significant differences between groups of monolingual and bilingual participants does not necessarily mean that bilinguals and monolinguals process these cognitive tasks in exactly the same way. There is some evidence that the processes needed for bilingual language control are not the same as those required by monolinguals (e.g., Hernandez et al., 2001), and that these differences have behavioral implications (e.g., Abutalebi et al., 2012). Therefore, it is recommended to abandon the quest for bilingual advantages and instead to focus on the question as to why at least some (but not all) bilinguals tend to process cognitive control tasks differently (but not always better) than monolinguals.

One explanation for this could be related to developmental differences between monolinguals and bilinguals because bilingual *advantages* are not consistently present across the lifespan of a bilingual individual (see Bialystok, 2007). As suggested by Bialystok et al. (2004), it is plausible that enhanced performance on conflict tasks only manifests itself in early childhood when individuals have not yet reached peak performance on these tasks. This in contrast to young adulthood, when performance is at ceiling level and environmental factors have little or no room to increase the efficiency of the processes involved in cognitive control. However, age cannot be the only factor to explain contradictory findings, because even research with children has produced bilingual advantage null effects (see, for instance, Antón et al., 2014).

One other explanation as to why bilingual advantages in cognitive control have only been observed in some but certainly not all studies can be related to the strategic choices made by individuals to carry out these tasks. In any task that involves the registration of response times and accuracy, such as in the interference tasks used to test the bilingual advantage, participants can optimize either speed or accuracy, or any compromise between both. Such conscious or unconscious strategic tendencies will have an effect on performance and this phenomenon is referred to as the speed-accuracy trade-off (Meyer et al., 1988). A tendency for speed may decrease response times at the cost of accuracy rates, whereas a tendency for accuracy may lead to slower response times but higher accuracy rates. This trade-off has been widely tested across various cognitive domains (see, for instance, Mackay, 1982; Forster et al., 2003), and it has been observed in interference tasks, such as the Simon (e.g., Hilchey et al., 2011; Ivanoff et al., 2014; van Wouwe et al., 2014) and flanker task (e.g., Rinkenauer et al., 2004; Wylie et al., 2009; Uemura et al., 2013).

Most studies about bilingual effects on cognitive control only focus on speed but not on accuracy. In a highly critical review article on the bilingual advantage, Paap et al. (2014) report that only 12 out of the 24 reviewed studies found lower response times for bilinguals than monolinguals (Luk et al., 2011; Tao et al., 2011; Salvatierra and Rosselli, 2011; Yang et al., 2011; Abutalebi et al., 2012; de Abreu et al., 2012; Poarch and van Hell, 2012; Schroeder and Marian, 2012; Kapa and Colombo, 2013; Marzecova et al., 2013; Morales et al., 2013; Pelham and Abrams, 2014), while information about the accuracy data is not provided. A separate

analysis on the accuracy data of these 24 studies reveals that only five mention a bilingual advantage in terms of accuracy (Tao et al., 2011; Yang et al., 2011; Marzecova et al., 2013; Morales et al., 2013; Gathercole et al., 2014). This logically implies that the speed and accuracy outcomes did not align in the other studies reporting a bilingual advantage in speed processing and it could also indicate the presence of a speed-accuracy trade-off. One reason why analyses on accuracy are often neglected is because errors are rare in young adults performing cognitive control tasks. Error rates on these tasks are much higher in populations of children under the age of 12 (Bunge et al., 2002), which makes this group perfectly suitable for investigating the developmental aspects of differences in the speed-accuracy trade-off between bilinguals and monolinguals. Moreover, some studies on bilingualism and cognitive control in children have found advantages in response times but not in accuracy (e.g., Martin-Rhee and Bialystok, 2008; Barac and Bialystok, 2012; Poarch and van Hell, 2012), again suggesting a potential speed-accuracy trade-off also in that age group.

The Present Study

This study set out to determine to what extent differences in strategic tendencies toward speed or accuracy between bilinguals and monolinguals explain part of the ongoing controversy surrounding the existence of a bilingual control advantage. It is well-known that the presence of two language systems in the bilingual mind generates conflict at various levels of linguistic analysis (e.g., van Heuven et al., 2008; Moreno et al., 2010; Blanco-Elorrieta and Pyllkanen, 2016) and that bilinguals must develop strategies to cope with this conflict in order to suppress the non-target language system and to activate the target one (e.g., Frenck-Mestre and Pynte, 1997). It has been proposed that domain-general interference tasks (such as the flanker or Simon task) generate conflict that is solved by the same processes as those required for daily bilingual language usage (e.g., Coderre et al., 2016). Strategic choices are not only needed to resolve the conflict generated by the most complex trials, but also to decide how to increase performance on these interference tasks. In general, individuals may optimize either speed or accuracy, which means that they can show faster response times at the cost of higher error rates, or instead be more accurate at a slower pace.

We hypothesize that bilinguals may show different strategies relative to monolinguals, after daily exposure to language conflicts and the need for developing strategies to overcome such conflict. This hypothesis is based on a review of the literature on the bilingual advantage. While some have challenged its existence based on reaction time data (Paap et al., 2014), their case could even be more convincing when error rates or accuracy of processing is considered. In some cases, better performance for bilinguals is only observed when reaction times and not accuracy scores are taken into account. This may be indicative of a selective speed-accuracy trade-off only for bilinguals, suggesting that bilinguals opt for a clear speed strategy when carrying out interference tasks, and this strategic choice may go at the cost of accuracy.

Our study intended to investigate this by assessing the correlation between response time (lower = better) and accuracy

rates (higher = better), possibly showing that faster processing is compensated by lower accuracy. Additionally, we aimed to examine to what extent this speed-accuracy trade-off was related to developmental differences in bilinguals' cognitive control performance. Recent literature on the interaction between bilingualism and cognitive control seems to indicate that bilingual benefits are more frequently found in young children than in young adults, thereby highlighting potential developmental factors affecting this interaction (for a recent review, see Zhou and Krott, 2016). Even within older children and young adults, the cognitive effects of bilingualism seem to dissipate, and this phenomenon can be related to the finding that the age between 6 and 8 years old is critical for rapid development of executive functioning (Best and Miller, 2010). Often, beneficial effects related to bilingualism are reported in children from birth up to the age of six (e.g., Martin-Rhee and Bialystok, 2008; Kovacs and Mehler, 2009; Morales et al., 2013; Crivello et al., 2016; Woumans et al., 2016), but not in children over the age of six (e.g., Martin-Rhee and Bialystok, 2008; Antón et al., 2014; Abdelgafar and Moawad, 2015), which again is indicative of the transition phase of this age group. Therefore, we compared two groups of younger and older children.

Based on previous studies, we anticipated differences between monolinguals and bilinguals in the younger but not in the older age group. In line with the main focus of this article and our first hypothesis, we expected strategic task tendencies to play a role in the development of the bilingual advantage. If it is true that speed-accuracy trade-offs are one of the reasons why bilingual advantages may be very variable, they should be smaller or non-existent in younger compared to older children.

MATERIALS AND METHODS

Participants

Participants were recruited through schools and after-school-care centers in Belgium. Parents received an information letter on the study's procedure and filled out an informed consent when they agreed to let their child take part. In total, we obtained authorisations for a large group of 122 children. There were 59 younger children (6-year-olds), of which 29 were monolingual and 30 bilingual. The older children (11-year-olds) consisted of 31 monolinguals and 32 bilinguals. Mean ages and other demographic variables are reported in **Table 1**. With regard to age, younger monolinguals ($M = 6.7$, $SD = 0.3$) did not differ from younger bilinguals ($M = 6.6$, $SD = 0.3$) ($t < 1.0$, ns). Older monolinguals ($M = 11.5$, $SD = 0.3$) were slightly younger than older bilinguals ($M = 11.8$, $SD = 0.5$) ($t_{118} = -2.91$, $p = 0.004$), hence we analyzed a subset of these two groups, excluding the two youngest monolinguals and the three oldest bilinguals. This left us with two comparable groups of older monolinguals ($M = 11.6$, $SD = 0.3$) and older bilinguals ($M = 11.7$, $SD = 0.3$) ($t_{56} = -1.35$, $p = 0.184$).

The children's language background and socioeconomic status (SES) was assessed through a questionnaire. Parents indicated which languages their child had mastered, at which age they acquired them and how proficient they are in them. The parents

specified the child's language proficiency on a 4-point Likert scale, ranging from 1 (=very low) to 4 (=very high/native). They also confirmed that their child did not have any learning disorders, or language development or comprehension issues. SES was a composite score of the parents' educational levels (elementary, secondary, or higher education) and intelligence was measured through Raven's Progressive Matrices (Raven, 1938; Raven et al., 1998). **Table 1** shows that monolinguals and bilinguals from both age groups were matched for these measures.

Design and Procedure

All children were tested individually and the test battery consisted of an intelligence test (Raven's Matrices) and two control tasks (Simon and flanker). The order of task administration was fixed for all participants: the Simon task came first, followed by the flanker task, to end with the Raven's test. Testing lasted around 30 min per participant. Breaks were allowed between tasks and between experimental blocks during the control tasks. The children were seated at a distance of approximately 60 cm from the screen. Control task stimuli were presented via Tscope software (Stevens et al., 2006) on an IBM-compatible laptop with 15-inch screen, running XP.

Raven's Progressive Matrices

Raven's Matrices is a test of analytic reasoning and is considered to be a good measure of fluid intelligence. This test of intelligence was added to our research design because previous research has shown that acquisition of a second language at a young age may foster intellectual development (Woumans et al., 2016). We administered two versions; the colored (Raven et al., 1998) and the standard version (Raven, 1938). The colored matrices are suited for children aged 5 to 11, whereas the standard matrices are suited for age 11 and older. The former test consists of 36 colored drawings with a missing segment which are equally divided over three sets (A, Ab, B) and ordered in terms of increasing difficulty. Participants are asked to complete the drawings indicating one of the six possible answers. A shortened version of the standard matrices was conducted (Van der Elst et al., 2013) to match the amount of items in the colored version, in which only set B, C, and D of the traditional sets A, B, C, D, and E were employed. In set B, each item had six possible options for completion, in set C and D, each item had eight possible options. Since we used subtests instead of the complete one, raw scores were employed as an estimate of participants' intelligence.

Simon Task

A version of the original task by Simon and Rudell (1967) was implemented. Colored dots appeared either on the left or right side of the screen. Participants were asked to press the left (right) key on the keyboard when a green dot appeared, and the right (left) key when the red dot appeared, and this as quickly and as accurately as possible. Response mapping was counterbalanced across participants according to parity of participant number. Each trial began with a fixation of 600 ms, followed by a clear screen and the stimulus, which lasted until the participant's response or up to 2500 ms. There was a 500 ms blank interval before the next fixation period. The task consisted of 10

TABLE 1 | Demographic data of monolinguals and bilinguals in both age groups.

	Younger children		Older children		Analysis	
	Monolingual	Bilingual	Monolingual	Bilingual	Test	<i>p</i>
<i>N</i>	29	30	29	29		
Male/female Ratio	17/12	13/17	13/16	11/21	$\chi^2(3) = 2.72$	0.437
Age (in years)	6.7 (0.3)	6.6 (0.3)	11.6 (0.3)	11.7 (0.3)	$F_{3,113} = 2301.71$	<0.001
Raven Score	23.7 (3.9)	28.4 (4.4)	24.4 (4.8)	27.9 (3.8)	$F_{3,118} = 9.30$	<0.001
L1 Dutch/French	29/0	30/0	31/0	32/0	–	–
L1 AoA (in years)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	–	–
L1 Proficiency ¹	4.0 (0.0)	3.4 (0.5)	4.0 (0.0)	3.5 (0.5)	$F_{3,113} = 18.55$	<0.001
L2 AoA (in years)	–	0.8 (0.8)	–	0.7 (0.8)	$F_{1,57} < 1.0$	0.618
L2 Proficiency ¹	–	3.1 (0.9)	–	3.4 (0.6)	$F_{1,57} = 2.72$	0.105
SES ²	2.6 (0.5)	2.7 (0.4)	2.6 (0.4)	2.5 (0.5)	$F_{3,113} < 1.0$	0.513

Standard deviations are presented between parentheses. ¹L1 and L2 proficiency were indicated on a 4-point Likert scale, ranging from 1 (=very low proficiency) to 4 (=very high/native proficiency). ²SES was a composite scores of parents' education levels. Three levels were defined: 1 (=elementary), 2 (=secondary), and 3 (=higher).

randomized practice trials and three blocks of 40 randomized experimental trials. Half of all trials presented the colored dot on the same side of the associated response key (congruent trials) and half on the opposite side (incongruent trials).

Flanker Task

A version of the Eriksen flanker task (Eriksen and Eriksen, 1974) was administered, in which five arrows were presented in the center of the screen and participants were asked to indicate the direction (left or right) of the central arrow. The central arrow could either point into the same direction as the four flankers (e.g., < < < < <, congruent trials) or into the other direction (e.g., < < > < <, incongruent trials). Each trial started with a fixation period of 500 ms and was followed by a clear screen and a stimulus presentation of maximum 2500 ms. A blank interval of 500 ms preceded the next trial. The task included 10 practice trials and three blocks of 40 experimental trials each. Half of the trials were incongruent.

TABLE 2 | Reaction times of correct trials (RT – ms) and accuracy scores (ACC – percentages) in the Simon and flanker task split for younger and older monolinguals and bilinguals (standard deviations between parentheses).

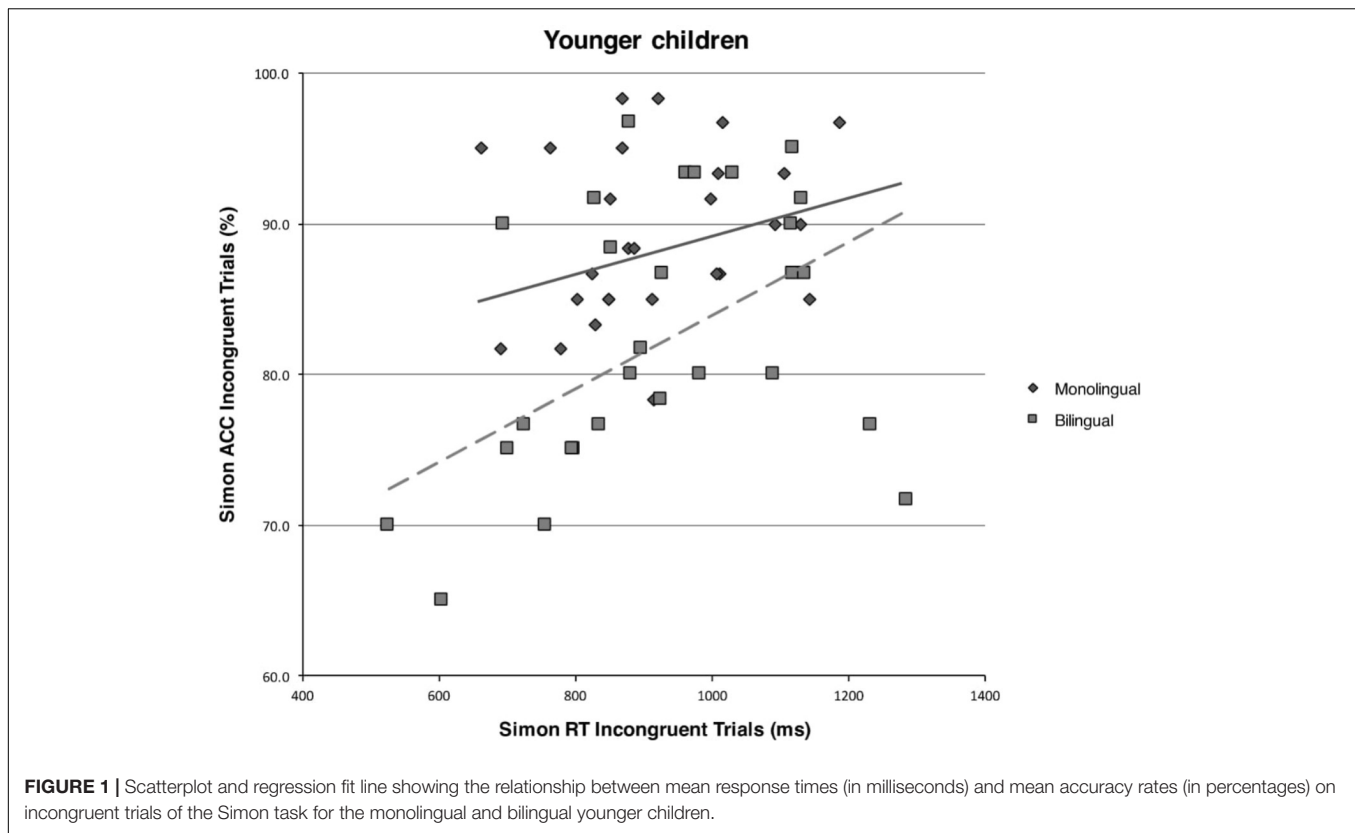
	Younger children		Older children	
	Monolingual	Bilingual	Monolingual	Bilingual
<i>Simon RT</i>				
Congruent	859 (119)	816 (185)	605 (112)	568 (102)
Incongruent	918 (135)	911 (195)	653 (118)	604 (91)
<i>Simon ACC</i>				
Congruent	92.3 (4.7)	89.8 (6.2)	91.4 (7.1)	92.6 (5.1)
Incongruent	88.2 (8.2)	81.8 (9.9)	86.1 (7.8)	88.5 (10.2)
<i>Flanker RT</i>				
Congruent	980 (124)	992 (207)	612 (96)	594 (131)
Incongruent	1241 (200)	1241 (240)	757 (137)	684 (159)
<i>Flanker ACC</i>				
Congruent	92.1 (6.9)	89.3 (9.3)	97.3 (2.1)	95.1 (4.3)
Incongruent	79.6 (14.0)	70.7 (19.9)	88.7 (6.4)	88.4 (7.5)

RESULTS

Cognitive control tasks were analyzed by mean reaction times of correct trials (RT) and accuracy scores (ACC) (see **Table 2**). Outlier RTs were trimmed for individual participants by calculating the mean across all trials and excluding any response deviating by more than 2.5 SD of the mean. This procedure eliminated 2.9% of all Simon data and 2.6% of all flanker data. On the Simon task, data from one younger monolingual and one younger bilingual participant were excluded from further analyses due to performance below chance accuracy level of 60%. On the flanker task, data from 10 younger monolingual and 6 younger bilingual participants were excluded from further analyses for the same reason. This exclusion rate is in line with results from previous studies on cognitive control in young children (e.g., Woumans et al., 2017) and can be explained by our choice to administer the default version of the flanker task (thus not the child-friendly version with fish as stimuli) for the purpose of better comparability with the data from the older children. On the remaining data, 2 (Age Group: Younger, Older) \times 2 (Language Group: Monolingual, Bilingual) \times 2 (Congruency: Congruent, Incongruent) repeated measure ANOVAs were performed to measure the effect of L2 Exposure. Planned comparisons were always employed to disentangle the effects of Age Group and Language Group. When the Levene Statistic was significant, equal variance was not assumed. On the same data, Pearson's correlational analyses between mean response times and mean accuracy rates were conducted to test for speed-accuracy trade-offs. These analyses were first applied to the entire groups of younger and older bilinguals and then to the bilingual and monolingual groups within these two age groups, separately. Statistical significance was corrected for multiple comparisons using a Bonferroni corrected significance level.

Demographics

Analyses revealed that none of the groups differed for male/female ratio or SES (**Table 1**). There was, however, a



difference between younger and older children on Raven scores ($t_{115} = 27.64$, $p < 0.001$), probably due to the fact that raw scores instead of norm scores were used. To our knowledge, no reliable norm scores are available for the subtests that we administered to the participants of the current study (see section “Design and Procedure”). Within the two age groups, none of the Language Groups differed from each other (all t s < 1.0 , ns). Planned comparisons showed that L1 proficiency was, within Age Group, always higher for monolinguals than for bilinguals (Younger: $t_{29} = 6.16$, $p < 0.001$, Older: $t_{28} = 4.53$, $p < 0.001$). Independent samples showed that, across Age Groups, there were no differences between monolinguals and bilinguals on L2 AoA ($t_{57} < 1.0$, $p = 0.618$) and self-reported L2 proficiency ($t_{57} = -1.65$, $p = 0.105$).

Simon Task

Descriptive statistics are summarized in **Table 2**. In the RT analysis, the main effect of Congruency was significant ($F_{1,111} = 147.66$, $p < 0.001$, $\eta_p^2 = 0.571$), indicating faster responses to congruent trials ($M = 711$ ms, $SD = 184$) than to incongruent trials ($M = 770$ ms, $SD = 200$). There was also a main effect of Age Group ($F_{1,111} = 114.66$, $p < 0.001$, $\eta_p^2 = 0.508$) with faster RTs for older children, but no main effect of Language Group ($F_{1,111} = 1.87$, $p = 0.174$, $\eta_p^2 = 0.017$). The two-way interaction between Congruency and Age Group was significant ($F_{1,111} = 12.32$, $p = 0.001$, $\eta_p^2 = 0.100$), revealing a smaller Simon effect for older children ($M = 42$ ms, $SD = 40$) than for younger children ($M = 77$ ms, $SD = 64$). The interaction

between Congruency and Language Group was not significant ($F_{1,111} = 1.39$, $p = 0.240$, $\eta_p^2 = 0.012$), and neither was the one between Age Group and Language Group ($F_{1,111} < 1.0$, ns). Yet, further analyses disclosed a significant three-way interaction between Congruency, Language Group, and Age Group ($F_{1,111} = 6.05$, $p = 0.015$, $\eta_p^2 = 0.052$). Planned comparisons demonstrated a significant difference on the Simon effect for younger monolinguals and bilinguals ($t_{54.25} = -2.16$, $p = 0.036$), with monolinguals displaying a smaller effect, and no significant difference between the older language groups ($t_{55.64} = 1.18$, $p = 0.245$).

In the accuracy analyses, there was a main effect of Congruency ($F_{1,111} = 49.68$, $p < 0.001$, $\eta_p^2 = 0.309$), with higher scores for congruent trials ($M = 91.5\%$, $SD = 5.9$) than for incongruent trials ($M = 86.1\%$, $SD = 9.4$). There was no effect of Age Group ($F_{1,111} = 1.90$, $p = 0.171$, $\eta_p^2 = 0.017$) or Language Group ($F_{1,111} = 1.204$, $p = 0.275$, $\eta_p^2 = 0.011$). There was an Age Group*Language Group interaction ($F_{1,111} = 3.48$, $p = 0.011$, $\eta_p^2 = 0.056$). The difference between younger monolinguals and bilinguals (4.43%) was larger than that between older monolinguals and bilinguals (1.77%). None of the other interactions were significant either (all p s > 0.095).

A Pearson's correlational analysis on the subset of younger monolingual children revealed no significant speed-accuracy trade-off on any of the investigated measures, all p s > 0.017 , the Bonferroni corrected significance level. The one on the subset of younger bilingual children, however, indicated a highly significant speed-accuracy trade-off for incongruent trials

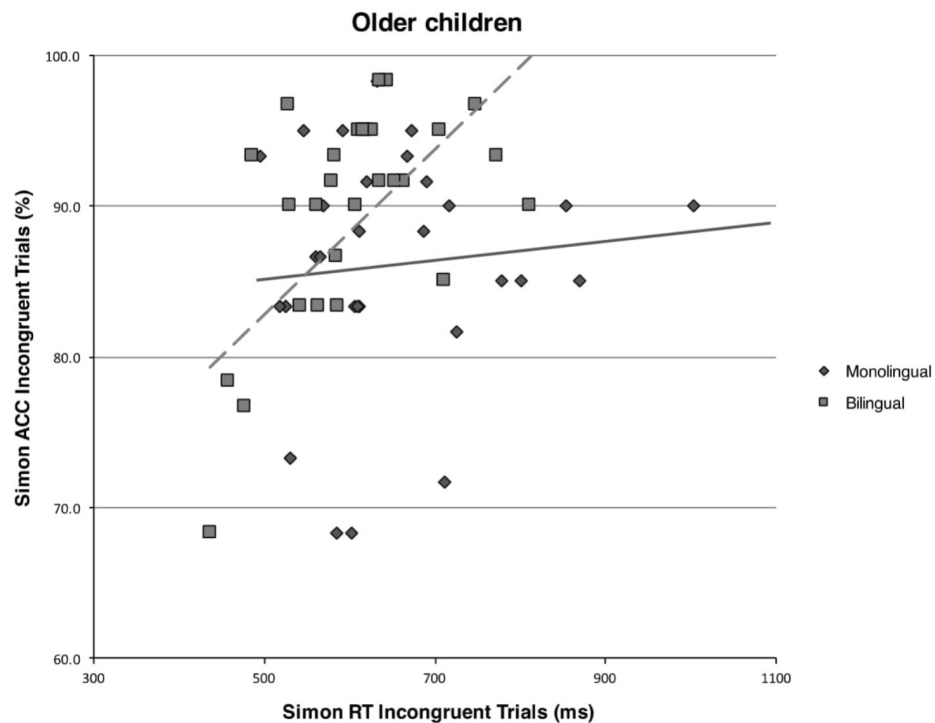


FIGURE 2 | Scatterplot and regression fit line showing the relationship between mean response times (in milliseconds) and mean accuracy rates (in percentages) on incongruent trials of the Simon task for the monolingual and bilingual older children.

($r_{29} = 0.48$, $p = 0.001$) but not for congruent trials or global performance (all p s > 0.017). See **Figure 1** for a graphical representation of the comparison between younger bilingual and monolingual children on the correlation between accuracy rates and response times on incongruent trials of the Simon task.

The same analyses on the subset of older monolingual children also disclosed no significant results (all p s > 0.05). In contrast, analyses on the subset of older bilingual children showed a highly significant speed-accuracy trade-off for global performance ($r_{29} = 0.53$, $p = 0.003$), and for incongruent ($r_{29} = 0.49$, $p = 0.007$) but not congruent trials ($r_{29} = 0.15$, $p = 0.435$). See **Figure 2** for a graphical representation of the comparison between older bilingual and monolingual children on the correlation between accuracy rates and response times on incongruent trials of the Simon task.

Flanker Task

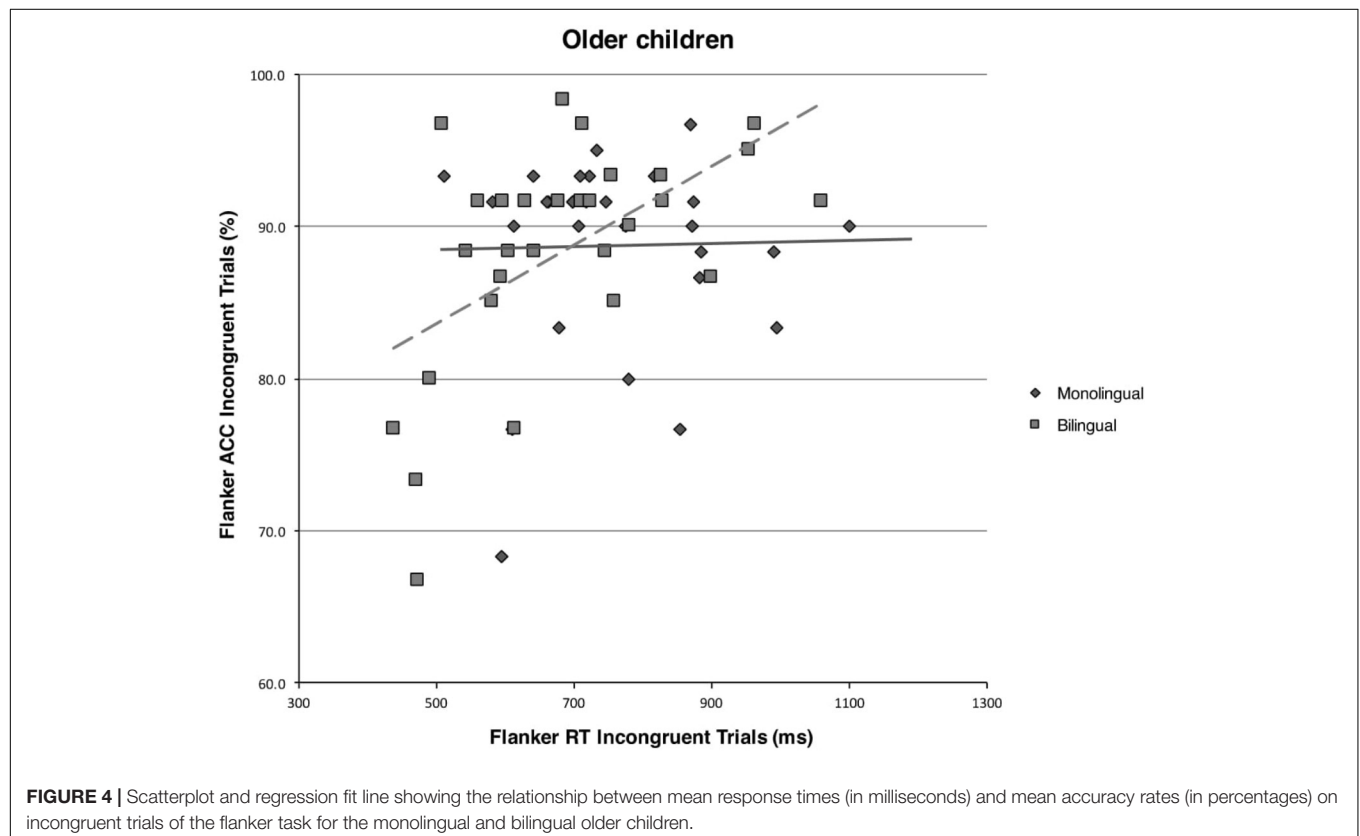
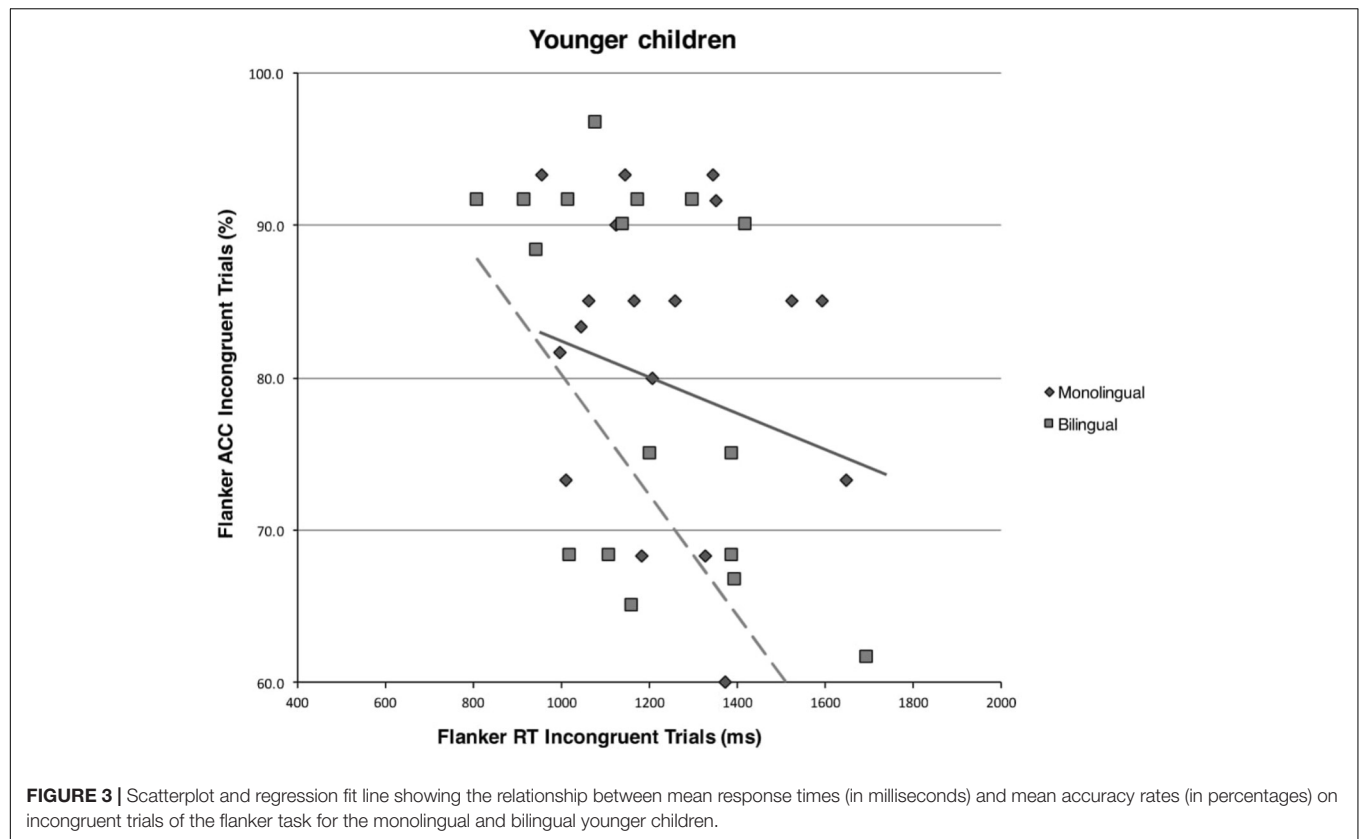
Descriptive statistics are summarized in **Table 2**. For RTs, the main effect of Congruency was significant ($F_{1,97} = 280.44$, $p < 0.001$, $\eta_p^2 = 0.743$), indicating faster responses to congruent trials. There was also a main effect of Age Group ($F_{1,97} = 206.74$, $p < 0.001$, $\eta_p^2 = 0.681$), demonstrating faster RTs for older children, but no effect of Language Group ($F_{1,97} < 1.0$, p ns.). There was, however, a Congruency*Age Group interaction ($F_{1,97} = 38.19$, $p < 0.001$, $\eta_p^2 = 0.282$), with a smaller flanker effect for older children ($M = 118$ ms, $SD = 68$) than for younger children ($M = 255$ ms, $SD = 152$). Although repeated measures analyses exposed no other two-way interaction effects and no

three-way interaction between Congruency, Language Group, and Age Group ($F_{1,97} < 1.0$, p ns.), planned comparisons still signaled a significant difference between older monolinguals and bilinguals on the flanker effect ($t_{55,96} = 3.40$, $p = 0.001$), with a smaller effect for bilinguals ($M = 90$ ms, $SD = 63$) as opposed to monolinguals ($M = 145$ ms, $SD = 61$).

Measuring accuracy, similar results were obtained, with higher scores for congruent trials ($F_{1,97} = 92.07$, $p < 0.001$, $\eta_p^2 = 0.487$) and for older participants ($F_{1,97} = 35.99$, $p < 0.001$, $\eta_p^2 = 0.271$), and for monolinguals ($F_{1,97} = 5.06$, $p < 0.05$). There was also a Congruency*Age Group interaction ($F_{1,97} = 10.75$, $p = 0.001$, $\eta_p^2 = 0.100$), with older children ($M = 7.6\%$, $SD = 5.9$) having a smaller accuracy effect than younger children ($M = 15.5\%$, $SD = 27.3$). No other effects were significant.

Pearson's correlational analyses on the subset of younger monolingual or young bilingual children did not reveal any significant speed-accuracy trade-offs (all p s > 0.017 , the Bonferroni corrected significance level). See **Figure 3** for a graphical representation of the comparison between younger bilingual and monolingual children on the correlation between accuracy rates and response times on incongruent trials of the flanker task.

A Pearson's correlational analysis on the subset of older monolingual children revealed no significant correlations at all (all p s > 0.017). The same analysis on the subset of older bilingual children, however, revealed highly significant speed-accuracy trade-off for global performance ($r_{29} = 0.54$, $p = 0.002$) and for incongruent trials ($r_{29} = 0.55$, $p = 0.002$), but not for congruent



trials ($ps > 0.017$). See **Figure 4** for a graphical representation of the comparison between older bilingual and monolingual children on the correlation between accuracy rates and response times on incongruent trials of the flanker task.

DISCUSSION

The aim of this study was to investigate the role of cognitive development and speed-accuracy trade-offs in the bilingual advantage controversy. Therefore, two groups of children (monolinguals and bilinguals) from two different age categories (younger and older children) were tested on cognitive control performance in two of the most frequently used tasks in the bilingualism literature: the Simon task and the flanker task. In line with previous findings, we only expected group differences between bilinguals and monolinguals in the youngest age group but not in the older one (Bialystok et al., 2004). Nevertheless, we did not merely intend to compare bilinguals to monolinguals in a between-group design, but also determine whether the absence or presence of differences in cognitive control are related to strategic task tendencies (i.e., optimizing either speed or accuracy performance) to resolve conflict. Our expectation was that bilinguals would follow a particular strategy to carry out these tasks, as indicated by a significant speed-accuracy trade-off, while monolinguals would show a more random pattern of behavior. Most crucially, we anticipated a relationship between speed-accuracy trade-off and the bilingual advantage, in the sense that such a trade-off could hide potential group differences.

No Clear-Cut Evidence for a Bilingual Advantage

A first important finding of this study was that there was no clear-cut evidence for a bilingual advantage. On the one hand, we did observe a smaller congruency effect for the older bilinguals on the flanker task; whereas, on the other, we found smaller congruency effects for younger monolinguals on the Simon task and higher accuracy scores for monolinguals in general on the flanker. We could therefore not confirm our first hypothesis that the bilingual advantage would only be found in the youngest and not the oldest group. Our results are, however, in line with recent meta-analyses on the bilingual advantage showing dubious results (de Bruin et al., 2014; Lehtonen et al., 2018). Furthermore, because both global measures of cognitive control (performance on the task as a whole, see, for instance, Costa et al., 2009) and specific measures (performance on incongruent trials only, see, for instance, Marzecova et al., 2013) were not consistently affected by bilingualism, we were unable to distinguish between interpretations of the bilingual advantage in terms of monitoring or inhibition.

Speed-Accuracy Trade-Offs

The major interest of the current study did not lie in the quest for a bilingual advantage, but rather in the investigation of potential differences between bilinguals and monolinguals in strategic task tendencies. In line with our expectations, we found evidence for speed-accuracy trade-offs only for bilinguals and

not monolinguals, and this in the two tasks under scrutiny. These results reveal for the first time a group difference in the strategies underlying the execution of cognitive control tasks. Confronted with the need for conflict resolution in a control task, bilinguals sought to optimize their performance by choosing a clear strategy, either by boosting their response times at the cost of accuracy, or by improving their accuracy rate by slowing down their performance. The monolinguals did not implement a similar strategy, as their performance did not show any relationship between speed and accuracy. We suggest that the cause for this between-group difference is comparable to that of the bilingual advantage, as it may also constitute the combination of training and transfer effects. Bilinguals face the constant need for conflict resolution as they have to manage two language systems, either when they activate the target language in face of interference from the non-target language, or when they switch between languages (e.g., Moreno et al., 2010; Tse and Altarriba, 2012). Compared to other language users, it has been found that bilinguals develop specific strategies to solve these linguistic conflicts (e.g., Frenck-Mestre and Pynte, 1997; Blanco-Elorrieta and Pytkkanen, 2016), and in the domain of language contact at the level of the individual language user, these have been labeled as 'bilingual optimisation strategies' (Muysken, 2013; Indefrey et al., 2017). In the same vein, speed-accuracy trade-offs can be seen as an optimisation strategy intended to boost performance in conflict situations. Interestingly, the implementation of this strategy in bilinguals in the Simon task was only visible for incongruent trials, or those trials for which conflict resolution is needed to attend to the task-relevant dimension in face of competition from a task-irrelevant dimension.

These findings suggest that the optimisation strategies that bilinguals develop when dealing with linguistic conflict may transfer into the non-verbal domain and that they may apply to any situation where a bilingual individual encounters conflict. As such, this training and transfer effect is an elaboration of the theoretical foundations of the bilingual advantage in cognitive control (see Kroll and Bialystok, 2013) as it suggests that a crucial difference between bilinguals and monolinguals regarding cognitive control lies in the strategies bilinguals actively recruit to resolve conflict, even when their response times or accuracy rates do not significantly deviate from those of monolinguals. This observation may have important implications for the bilingual advantage debate. Previously, the quest for bilingual effects in cognitive control was confined to an investigation of potential differences in the speed (or accuracy) of processing, and the absence of these differences led to the assumption that there is no consistent evidence for a bilingual advantage (Paap and Greenberg, 2013; Paap et al., 2014; von Bastian et al., 2016). However, this quest for behavioral advantages could interfere with the different strategies used by bilinguals and monolinguals to carry out these tasks. If bilinguals seek – even unconsciously – to optimize their performance, only one of these two dimensions will be positively affected. Between-group differences in speed-accuracy trade-offs could thus explain why bilingual advantages are observed either in terms of processing speed or accuracy (compare to the studies listed by Paap et al., 2014).

We also propose that differences in strategic task tendencies may mask potential group differences in accuracy or speed. In spite of the between-group differences in speed-accuracy trade-offs, no similar differences were detected when speed and accuracy were analyzed separately. However, our descriptive statistics revealed a tendency of lower response times for the bilinguals and higher accuracy for the monolinguals. In one subgroup (the older children on the flanker task), this even led to a monolingual advantage in accuracy. Within the explanatory framework of strategy choices, we suggest that this is the result of the bilinguals' optimisation strategy to boost response times at cost of lower accuracy. The question may arise why these group differences in speed-accuracy trade-offs have led on only one occasion to group differences in speed or accuracy. One reason for this could be that while the bilinguals as a group make use of optimisation strategies to resolve conflict in control tasks, the choice for a speed or an accuracy strategy may differ between individuals based on their need for interference suppression in daily bilingual language use related to variables such as the differences in proficiency level between L1 and L2, the degree of language switching, and the typological distance between both languages. Only if most or nearly all bilingual participants implement the same strategy to resolve conflict, a clear advantage may be found on that dimension. Previous studies seem to suggest that advantages are more frequently observed in speed than in accuracy, which may reveal a preference for a speed strategy among bilinguals (compare to the studies listed by Paap et al., 2014). However, the design of the current study did not allow us to make any claims on this issue and this is also one of its limitations. We therefore strongly recommend future studies on the bilingual to manipulate the speed and accuracy strategy by explicitly instructing which dimension must be prioritized (Wylie et al., 2009; Uemura et al., 2013). In line with the interpretation of this study's findings, we expect bilinguals to benefit more from these explicit instructions because they have been trained in the usage of optimisation strategies.

Development

The final research question of the current study dealt with the developmental aspects of the bilingual advantage and the potentially interfering role of speed-accuracy trade-offs in the manifestation of this advantage. Compatible with the results for the test population as a whole, an age difference was found between the flanker and the Simon task specifically for the bilingual subgroup. Whereas speed-accuracy trade-offs were observed in both age groups for the Simon task, only the older children showed a correlation between speed and accuracy on the flanker task. These findings were – at least for the Simon task – not in line with our own expectations, as we anticipated a speed-accuracy trade-off in the older but not in the younger children.

A first reason for this may be related to the specific characteristics of each of the two cognitive control tasks, which do not only differ from each other in the mean length of response times (which is significantly higher for the flanker than for the Simon task), but also in the underlying mechanisms of conflict resolution due to compatibility or congruency between stimulus

and response (Kornblum et al., 1990). On an incongruent flanker trial, one (task-relevant) dimension of the stimulus (the direction of the central arrow) conflicts with another (but task-irrelevant) dimension of the same stimulus (the direction of the surrounding arrows). On the other hand, on an incongruent Simon trial, a (task-relevant) dimension of the stimulus (the color of the square) conflicts with a (task-irrelevant) dimension of the response (the location of the response). As a result of these differences, both types of conflict are processed independently (Li et al., 2014) with stimulus–stimulus conflicts (as generated in a flanker task) inducing stronger behavioral effects (Fruhholz et al., 2011) than stimulus–response conflicts (as generated in a Simon task). As it may be more effortful to process a task that induces stronger behavioral effects, it could be that only older children have the ability to make strategic choices on stimulus–stimulus conflicts in the flanker task, whereas the same does not apply to the easier stimulus–response conflicts in the Simon task.

The second reason for the mismatch between the current study's hypotheses and its actual findings is that our expectations regarding the role of development were related to an anticipated bilingual advantage in the younger but not in the older children. As we did not consistently observe such an advantage, the rationale behind developmental differences in speed-accuracy trade-off was no longer present. We therefore assume that the developmental differences between the two tasks were solely caused by the characteristics of the individual tasks instead of any possible relationship with a bilingual advantage.

CONCLUSION

The most important contribution of the current study to the expanding bilingual advantage literature is that cognitive control differences between bilinguals and monolinguals can manifest themselves in strategic task tendencies implemented to resolve conflict, even when consistent performance differences between bilinguals and monolinguals in terms of speed and accuracy are absent. The crucial difference between our two language groups was that only bilingual children showed a consistent pattern of speed-accuracy trade-offs on the flanker and Simon task. Comparable to the theoretical foundations of the bilingual advantage, we have related these differences to a combined training and transfer effect as a result of the specific demands of bilingual language usage. Our findings prompt a nuanced view on the bilingual advantage debate: as we did not find any evidence for performance differences, the term 'advantage' may be a misnomer for what is happening in the bilingual mind (as compared to monolinguals); but at the same time, the variation in implemented strategies to resolve conflict illustrate the impact that constant exposure and usage of two (or more) language systems may have on cognitive processing in the bilingual mind (compare to Woumans et al., 2016).

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Ghent University's Faculty of

Psychology's ethical guidelines, with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Ethical Committee of the Faculty of Psychology at Ghent University.

AUTHOR CONTRIBUTIONS

ES and EW determined the research questions and performed the statistical analysis. EW programmed the tasks and conducted the experiments. WD made suggestions. ES

drafted the manuscript. EW and WD provided critical assessments.

FUNDING

The authors gratefully acknowledge support from LEMMA (Language, Education, and Memory in Multilingualism and Academia), a Concerted Research Actions fund (GOA – BOF13/GOA/032) of Ghent University, and from HOA23, a Research Action funded by the Research Council of the Vrije Universiteit Brussel.

REFERENCES

- Abdelgafar, G. M., and Moawad, R. A. (2015). Executive function differences between bilingual arabic-english and monolingual arabic children. *J. Psycholinguist. Res.* 44, 651–667. doi: 10.1007/s10936-014-9309-3
- Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., et al. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cereb. Cortex* 22, 2076–2086. doi: 10.1093/cercor/bhr287
- Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., et al. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Front. Psychol.* 5:398. doi: 10.3389/fpsyg.2014.00398
- Barac, R., and Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: role of language, cultural background, and education. *Child Dev.* 83, 413–422. doi: 10.1111/j.1467-8624.2011.01707.x
- Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x
- Bialystok, E. (2007). Cognitive effects of bilingualism across the lifespan. *Paper Presented at the 32nd Annual Boston-University Conference on Language Development*. Boston, MA.
- Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290
- Blanco-Elorrieta, E., and Pyllkanen, L. (2016). Bilingual language control in perception versus action: MEG reveals comprehension control mechanisms in anterior cingulate cortex and domain-general control of production in dorsolateral prefrontal cortex. *J. Neurosci.* 36, 290–301. doi: 10.1523/jneurosci.2597-15.2016
- Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., and Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron* 33, 301–311. doi: 10.1016/S0896-6273(01)00583-9
- Coderre, E. L., Smith, J. F., Van Heuven, W. J. B., and Horwitz, B. (2016). The functional overlap of executive control and language processing in bilinguals. *Biling. Lang. Cogn.* 19, 471–488. doi: 10.1017/S1366728915000188
- Costa, A., Hernandez, M., Costa-Faidella, J., and Sebastian-Galles, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don't. *Cognition* 113, 135–149. doi: 10.1016/j.cognition.2009.08.001
- Crivello, C., Kuzlyk, O., Rodrigues, M., Friend, M., Zesiger, P., and Poulin-Dubois, D. (2016). The effects of bilingual growth on toddlers' executive function. *J. Exp. Child Psychol.* 141, 121–132. doi: 10.1016/j.jecp.2015.08.004
- de Abreu, P., Cruz-Santos, A., Tourinho, C. J., Martin, R., and Bialystok, E. (2012). Bilingualism enriches the poor: enhanced cognitive control in low-income minority children. *Psychol. Sci.* 23, 1364–1371. doi: 10.1177/0956797612443836
- de Bruin, A., and Della Sala, S. (2017). "The bilingual advantage debate: Publication biases and the decline effect," in *The Handbook of the Neuroscience of Multilingualism*, ed. J. W. Schwieter (Hoboken, NJ: Wiley-Blackwell), 792–809.
- de Bruin, A., Treccani, B., and Della Sala, S. (2014). Cognitive advantage in bilingualism: an example of publication bias? *Psychol. Sci.* 26, 99–107. doi: 10.1177/0956797614557866
- Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/bf03203267
- Forster, J., Higgins, E. T., and Bianco, A. T. (2003). Speed/accuracy decisions in task performance: built-in trade-off or separate strategic concerns? *Organ. Behav. Hum. Decis. Process.* 90, 148–164. doi: 10.1016/S0749-5978(02)00509-5
- FrenchMestre, C., and Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Q. J. Exp. Psychol. Sect. Hum. Exp. Psychol.* 50, 119–148.
- Fruhholz, S., Godde, B., Finke, M., and Herrmann, M. (2011). Spatio-temporal brain dynamics in a combined stimulus-stimulus and stimulus-response conflict task. *Neuroimage* 54, 622–634. doi: 10.1016/j.neuroimage.2010.07.071
- Gathercole, V. C. M., Thomas, E. M., Kennedy, L., Prys, C., Young, N., Guasch, N. V., et al. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and rnetalinguistic tasks. *Front. Psychol.* 5:11. doi: 10.3389/fpsyg.2014.00011
- Hernandez, A. E., Dapretto, M., Mazziotta, J., and Bookheimer, S. (2001). Language switching and language representation in Spanish-English bilinguals: an fMRI study. *Neuroimage* 14, 510–520. doi: 10.1006/nimg.2001.0810
- Hilchey, M. D., Ivanoff, J., Taylor, T. L., and Klein, R. M. (2011). Visualizing the temporal dynamics of spatial information processing responsible for the Simon effect and its amplification by inhibition of return. *Acta Psychol.* 136, 235–244. doi: 10.1016/j.actpsy.2010.09.003
- Hilchey, M. D., and Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychon. Bull. Rev.* 18, 625–658. doi: 10.3758/s13423-011-0116-7
- Indefrey, P., Sahin, H., and Gullberg, M. (2017). The expression of spatial relationships in Turkish-Dutch bilinguals. *Biling. Lang. Cogn.* 20, 473–493. doi: 10.1017/S1366728915000875
- Ivanoff, J., Blagdon, R., Feener, S., McNeil, M., and Muir, P. H. (2014). On the temporal dynamics of spatial stimulus-response transfer between spatial incompatibility and Simon tasks. *Front. Neurosci.* 8:11. doi: 10.3389/fnins.2014.00243
- Kapa, L. L., and Colombo, J. (2013). Attentional control in early and later bilingual children. *Cogn. Dev.* 28, 233–246. doi: 10.1016/j.cogdev.2013.01.011
- Kornblum, S., Hasbroucq, T., and Osman, A. (1990). Dimensional overlap - Cognitive basis for stimulus-response compatibility - A model and taxonomy. *Psychol. Rev.* 97, 253–270. doi: 10.1037/0033-295X.97.2.253
- Kovacs, A. M., and Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6556–6560. doi: 10.1073/pnas.0811323106
- Kroll, J. F., and Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *J. Cogn. Psychol.* 25, 497–514. doi: 10.1080/20445911.2013.799170
- Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., and Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychol. Bull.* 144, 394–425. doi: 10.1037/bul0000142
- Li, Q., Nan, W. Z., Wang, K., and Liu, X. (2014). Independent processing of stimulus-stimulus and stimulus-response conflicts. *PLoS One* 9:e89249. doi: 10.1371/journal.pone.0089249

- Lu, C. H., and Proctor, R. W. (1995). The influence of irrelevant location information on performance - A review of the Simon and spatial Stroop effects. *Psychon. Bull. Rev.* 2, 174–207. doi: 10.3758/bf03210959
- Luk, G., De Sa, E., and Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Biling. Lang. Cogn.* 14, 588–595. doi: 10.1017/s1366728911000010
- Mackay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychol. Rev.* 89, 483–506. doi: 10.1037/0033-295x.89.5.483
- Martin-Rhee, M. M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Biling. Lang. Cogn.* 11, 81–93. doi: 10.1017/s1366728907003227
- Marzecova, A., Asanowicz, D., Kriva, L., and Wodniecka, Z. (2013). The effects of bilingualism on efficiency and lateralization of attentional networks. *Biling. Lang. Cogn.* 16, 608–623. doi: 10.1017/s1366728912000569
- Meyer, D. E., Osman, A. M., Irwin, D. E., and Kounios, J. (1988). The dynamics of cognition and action - Mental processes inferred from speed accuracy decomposition. *Psychol. Rev.* 95, 183–237. doi: 10.1037/0033-295x.95.2.183
- Morales, J., Calvo, A., and Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *J. Exp. Child Psychol.* 114, 187–202. doi: 10.1016/j.jecp.2012.09.002
- Moreno, S., Bialystok, E., Wodniecka, Z., and Alain, C. (2010). Conflict resolution in sentence processing by bilinguals. *J. Neurolinguist.* 23, 564–579. doi: 10.1016/j.jneuroling.2010.05.002
- Muysken, P. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Biling. Lang. Cogn.* 16, 709–730. doi: 10.1017/s1366728912000727
- Paap, K. R. (in press). “The bilingual advantage debate: quantity and quality of the evidence,” in *The Handbook of the Neuroscience of Multilingualism*, ed. J. W. Schwieter (Hoboken, NJ: Wiley-Blackwell), 755–791.
- Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Paap, K. R., Johnson, H. A., and Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *J. Cogn. Psychol.* 26, 615–639. doi: 10.1080/20445911.2014.944914
- Paap, K. R., Johnson, H. A., and Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69, 265–278. doi: 10.1016/j.cortex.2015.04.014
- Pelham, S. D., and Abrams, L. (2014). Cognitive Advantages and Disadvantages in Early and Late Bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 313–325. doi: 10.1037/a0035224
- Poarch, G. J., and van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: evidence from second-language learners, bilinguals, and trilinguals. *J. Exp. Child Psychol.* 113, 535–551. doi: 10.1016/j.jecp.2012.06.013
- Raven, J. C. (1938). *Progressive Matrices: A Perceptual Test of Intelligence*. London: H. K. Lewis.
- Raven, J. C., Court, J. H., and Raven, J. (1998). *Progressive Coloured Matrices*. Oxford: Oxford Psychologists Press.
- Rinkenauer, G., Osman, A., Ulrich, R., Muller-Gethmann, H., and Mattes, S. (2004). The locus of speed-accuracy trade-off in reaction time: inferences from the lateralized readiness potential. *J. Exp. Psychol. Gen.* 133, 261–282. doi: 10.1037/0096-3445.133.2.261
- Salvatierra, J. L., and Rosselli, M. (2011). The effect of bilingualism and age on inhibitory control. *Int. J. Biling.* 15, 26–37. doi: 10.1177/1367006910371021
- Schroeder, S. R., and Marian, V. (2012). A bilingual advantage for episodic memory in older adults. *J. Cogn. Psychol.* 24, 591–601. doi: 10.1080/20445911.2012.669367
- Simon, J. R., and Rudell, A. P. (1967). Auditory S-R compatibility - Effect of an irrelevant cue on information processing. *J. Appl. Psychol.* 51, 300–304. doi: 10.1037/h0020586
- Stevens, M., Lammertyn, J., Verbruggen, F., and Vandierendonck, A. (2006). Tscope: a C library for programming cognitive experiments on the MS Windows platform. *Behav. Res. Methods* 38, 280–286. doi: 10.3758/bf03192779
- Tao, L., Marzecova, A., Taft, M., Asanowicz, D., and Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Front. Psychol.* 2:123. doi: 10.3389/fpsyg.2011.00123
- Tse, C. S., and Altarriba, J. (2012). The effects of first- and second-language proficiency on conflict resolution and goal maintenance in bilinguals: evidence from reaction time distributional analyses in a Stroop task. *Biling. Lang. Cogn.* 15, 663–676. doi: 10.1017/s1366728912000077
- Uemura, K., Oya, T., and Uchiyama, Y. (2013). Effects of speed and accuracy strategy on choice step execution in response to the flanker interference task. *Hum. Mov. Sci.* 32, 1393–1403. doi: 10.1016/j.humov.2013.07.007
- Van der Elst, W., Ouweland, C., van Rijn, P., Lee, N., Van Boxtel, M., and Jolles, J. (2013). The shortened raven standard progressive matrices: item response theory-based psychometric analyses and normative data. *Assessment* 20, 48–59. doi: 10.1177/1073191111415999
- van Heuven, W. J. B., Schriefers, H., Dijkstra, T., and Hagoort, P. (2008). Language conflict in the bilingual brain. *Cereb. Cortex* 18, 2706–2716. doi: 10.1093/cercor/bhn030
- van Wouwe, N. C., van den Wildenberg, W. P. M., Claassen, D. O., Kanoff, K., Bashore, T. R., and Wylie, S. A. (2014). Speed pressure in conflict situations impedes inhibitory action control in Parkinson's disease. *Biol. Psychol.* 101, 44–60. doi: 10.1016/j.biopsycho.2014.07.0002
- von Bastian, C. C., Souza, A. S., and Gade, M. (2016). No evidence for bilingual cognitive advantages: a test of four hypotheses. *J. Exp. Psychol. Gen.* 145, 246–258. doi: 10.1037/xge0000120
- Woumans, E., Surmont, J., Struys, E., and Duyck, W. (2016). The longitudinal effect of bilingual immersion schooling on cognitive control and intelligence. *Lang. Learn.* 66, 76–91. doi: 10.1111/lang.12171
- Woumans, E., Surmont, J., Struys, E., and Duyck, W. (2017). The Longitudinal Effect of Bilingual Immersion Schooling on Cognitive Control and Intelligence. *Language Learning* 66, 76–91. doi: 10.1111/lang.12171
- Wylie, S. A., van den Wildenberg, W. P. M., Ridderinkhof, K. R., Bashore, T. R., Powell, V. D., Manning, C. A., et al. (2009). The effect of speed-accuracy strategy on response interference control in Parkinson's disease. *Neuropsychologia* 47, 1844–1853. doi: 10.1016/j.neuropsychologia.2009.02.025
- Yang, S. J., Yang, H. J., and Lust, B. (2011). Early childhood bilingualism leads to advances in executive attention: dissociating culture and language. *Biling. Lang. Cogn.* 14, 412–422. doi: 10.1017/s1366728910000611
- Zhou, B. N., and Krott, A. (2016). Data trimming procedure can eliminate bilingual cognitive advantage. *Psychon. Bull. Rev.* 23, 1221–1230. doi: 10.3758/s13423-015-0981-6

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Struys, Duyck and Woumans. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Exploiting Language Variation to Better Understand the Cognitive Consequences of Bilingualism

Andrea A. Takahesu Tabori, Emily N. Mech and Natsuki Atagi*

Department of Psychology, University of California, Riverside, Riverside, CA, United States

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Regina Anders-Jefferson,
San Francisco State University,
United States
Sara Incera,
Eastern Kentucky University,
United States

*Correspondence:

Natsuki Atagi
natagi@ucr.edu

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 01 July 2018

Accepted: 21 August 2018

Published: 07 September 2018

Citation:

Takahesu Tabori AA, Mech EN and
Atagi N (2018) Exploiting Language
Variation to Better Understand
the Cognitive Consequences
of Bilingualism.
Front. Psychol. 9:1686.
doi: 10.3389/fpsyg.2018.01686

Within the past decade, there has been an explosion of research investigating the cognitive consequences of bilingualism. However, a controversy has arisen specifically involving research claiming a “bilingual advantage” in executive function. In this brief review, we re-examine the nature of the “bilingual advantage” and suggest three themes for future research. First, there must be a theoretical account of how specific variation in language experience impacts aspects of executive function and domain general cognition. Second, efforts toward adequately characterizing the participants tested will be critical to interpreting results. Finally, designing studies that employ converging analytical approaches and sensitive methodologies will be important to advance our knowledge of the dynamics between bilingual language experience and cognition.

Keywords: bilingualism, bilingual advantage, individual differences, executive function, cognitive control

INTRODUCTION

A key tenet in research design is parsimony: to design studies that are as simple as possible. However, complex questions and designs are sometimes oversimplified, more so than parsimony requires. For example, the psychological and language sciences have traditionally looked for unifying principles across groups of people, which has led to questions such as “are bilinguals better at cognitive control than monolinguals?” However, this approach leads to group-level analyses with little regard for meaningful variation *within* each group. Rather than treating variation within groups as noise, perhaps we should *start* by studying that variation. Investigating individual differences is not new in the field of Psychology (e.g., Tyler, 1947; Anastasi, 1958), but consistently applying such an approach may provide clarity to the recent controversy about bilingual benefits (Paap and Greenberg, 2013; Paap, 2014; de Bruin et al., 2015; Valian, 2015; von Bastian et al., 2016; but see Baum and Titone, 2014; Fricke et al., in press) and bring about a more nuanced approach to the field. For instance, in much of the published literature on bilingualism, it is difficult to disentangle true null results from those arising from methodological constraints or inadequate comparisons (for a review, see Laine and Lehtonen, in press). This paper identifies three key themes to guide future research in the field. Specifically, we focus on re-examining the notion of the “bilingual advantage,” how we report participant characteristics, and how we might modify research methods and analytical approaches to better account for variability. In particular, we advocate for embracing variability and examining individual differences in bilingual experiences to better understand the cognitive and linguistic consequences of bilingualism. These suggestions take a multidimensional approach on language and have implications for all language researchers.

ON THE “BILINGUAL ADVANTAGE”

The “bilingual advantage” was a phrase first used to describe a result in which bilinguals out-performed monolinguals on tasks of cognitive control, and this advantage was theorized to be driven by the bilinguals’ constant need to manage competition from each language (Bialystok, 2001; Bialystok et al., 2004). This result sparked interest in the potential cognitive benefits of bilingualism, motivating studies that compared bilinguals and monolinguals on tasks of cognitive control (e.g., Costa et al., 2008; Poarch and van Hell, 2012). Although an increasing number of studies examined speakers in varied locations, of varied ages, and with varied tasks, the underlying theory remained largely unchanged and did not grow to encompass how variation may impact outcomes. As a result, research addressing the “bilingual advantage” became dichotomized, both in the experimental groups tested (monolinguals vs. bilinguals) and in the possible outcomes (bilingual advantage vs. no advantage). As such, an expectation that bilingualism would have a main effect on cognitive control performance became commonplace. Dichotomizing the groups tested and the possible outcomes has created controversy whenever studies do not demonstrate advantages for bilinguals relative to monolinguals. However, the problem with this logic is that bilingualism is a multidimensional construct (Luk and Bialystok, 2013), and as such, cannot be treated as a categorical variable. To overcome the controversy surrounding the bilingual advantage, it will be important to understand the mechanisms by which aspects of bilingual language experience (e.g., proficiency, literacy, age of acquisition) give rise to cognitive adaptations.

The Adaptive Control Hypothesis proposed by Green and Abutalebi (2013) provided an initial step toward understanding the relation between bilingual language contexts and cognitive changes. According to this hypothesis, variation in dual language contexts constrain which languages a bilingual can use and the degree to which they can switch between their languages. This contextual constraint is thought to impact cognitive control in distinct ways. Critically, bilingual language contexts are proposed to have specific cognitive outcomes – not just a generalized advantage. Thus, depending on the particular outcome that is investigated, there may or may not be differences between bilinguals and monolinguals. Finding a lack of differences between bilinguals and monolinguals (or between different types of bilinguals) is not inherently a problem, but rather, can be explanatory in its fit into the broader theoretical framework. For example, more than one aspect of bilingual language variation may be responsible for effects of bilingualism on cognitive control: there are mixed results reported in the language switching, language control, and cognitive control literature (e.g., Paap et al., 2014, 2017; Verreyt et al., 2016). Mixed results could be due to a variety of factors such as not measuring and accounting for a critical aspect of language experience or other relevant variables that also affect cognitive control (e.g., age; Kousaie and Phillips, 2017). Thus, in addition to the Adaptive Control Hypothesis, it will be critical to further develop theories that make specific predictions regarding how the variation in bilingual language experience may give rise to

differences in cognitive control or cognition more generally. Without specifying the underlying mechanism, further attempts to investigate bilingual differences may only contribute to, rather than clarify, the controversy.

Pivoting from testing monolinguals vs. bilinguals to answer a yes or no, advantage or no advantage question to one of “bilingual differences” may create greater insight into the mechanisms underlying the consequences of varied language experience (Bak, 2016). Although the primary source of the controversy of the “bilingual advantage” surrounds research investigating the consequences of bilingualism for executive function, the phrase “bilingual advantage” itself is now used widely. In the decade since the initial report (Bialystok, 2001; Bialystok et al., 2004), a virtual explosion of research has arisen claiming bilingual advantages in domains such as visual discrimination and habituation (e.g., Weikum et al., 2007; Sebastián-Gallés et al., 2012; Singh et al., 2015), communicative development (e.g., Fan et al., 2015; Liberman et al., 2017), novel word learning (e.g., Kaushanskaya and Marian, 2009), episodic memory (e.g., Schroeder and Marian, 2012), and phonetic learning (e.g., Antoniou et al., 2015), to name a few. While many of these studies draw connections to the underlying theory relating bilingual language regulation to cognitive control and executive function, more broadly, the results reported are domain-specific and suggest that the controversy surrounding the “bilingual advantage” may be focused too narrowly. While there are many advantages associated with bilingualism, we argue here that it will be important to redefine the way in which we describe such “advantages” to acknowledge the scope of the observed consequences and to promote a more appropriate approach to generalization across studies.

The controversy surrounding the consequences of bilingualism may provide a set of lessons for the field that extend beyond the studies that have been associated with this issue. The lessons emerging from this controversy have relevance not only for those directly investigating the cognitive consequences of bilingualism, but for language science more broadly. Embracing parsimony in the face of complexity may actually lead to oversimplification that slows progress rather than promoting it. Being specific and intentional about the degree to which we generalize terms across domains will clarify similarities and differences between theories of language and cognition. Additionally, there is a fundamental need to embrace variation, appropriately characterize it, and delineate how differences in experience may have consequences for the mind and brain. The following sections identify and suggest initial steps to move toward these goals by providing insight for characterizing our samples and designing our methodology.

ON CHARACTERIZING OUR SAMPLES: BILINGUALS ARE A DIVERSE GROUP

A critical factor that has been largely overlooked in research on the “bilingual advantage” is that bilinguals – as well as monolinguals – are heterogeneous, with a wide range of language backgrounds and experiences. Though a call for more nuanced

characterizations of bilinguals' diverse language experiences is not new (e.g., Green and Abutalebi, 2013; Kroll and Bialystok, 2013; Luk and Bialystok, 2013; Abutalebi and Green, 2016; Bialystok, 2016; Surrain and Luk, in press; Laine and Lehtonen, in press), the focus of much of the published research remains on differences between bilinguals and monolinguals, with little attention paid to who these bilinguals – or even monolinguals, for that matter – are. Understanding speakers' diverse language experiences will allow for a more critical investigation of the consequences of different language experiences for the mind and brain, providing insight into the interactions and moderating variables that may be obscuring group-level differences between bilinguals and monolinguals.

Given that bilingualism is a dynamic, multidimensional variable (e.g., Luk and Bialystok, 2013), detailed information about participant background and experiences – both past and present – is critical. Although participant characteristics such as self-rated proficiency, amount of use, and age of acquisition of each language are often provided (for a review, see Surrain and Luk, in press), in what context speakers learned and used each language in the past is typically left undescribed. However, there is evidence that learning to read in the home language affects literacy skills in other languages (e.g., Shanahan and Escamilla, 2009; Sparrow et al., 2014; Shin et al., 2015), suggesting that biliteracy – and likely the language of schooling – may be relevant dimensions to examine in studies of bilingualism and cognition. Additionally, language brokering (i.e., informal translation) experience has been found to affect language processing (e.g., López et al., 2017; López and Vaid, 2018) and conceptual representations (e.g., López and Vaid, 2016), pointing to the importance of understanding not only how much bilinguals have used each language but also for what purpose they have used each language. Such findings shed light on the need to consider past language experiences when examining a “group” as diverse as bilinguals.

Additionally, evidence for the enduring consequences of early language exposure can be found in research on functionally monolingual speakers who were exposed to a language early in life, but due to life circumstances, lost explicit knowledge of that language, and consequently function exclusively in their second language. One group of such “monolinguals” is international adoptees (IA): those who were exposed to one language as children and later lost all contact with and knowledge of this language after relocating permanently into their country of adoption. A number of studies suggest that despite having no functional knowledge in their first language and having spent the majority of their lives speaking another language, IAs show language processing signatures that are more similar to those of bilingual speakers of their lost language and their second language than those of monolingual speakers of their second language (e.g., Pierce et al., 2014, 2015, 2017). Similarly, research on childhood overhearers (i.e., adults who, as children, overheard speech in a language other than their native language) also suggests that despite not having productive knowledge of the language they overheard, overhearers are able to learn aspects of the phonology of that language better than those who were not childhood overhearers of that language (e.g., Au et al., 2002;

Knightly et al., 2003). Taken together, these findings suggest that despite discontinued use of an early exposed language, there are fundamental changes in language processing that endure into adulthood. Without adequately characterizing speakers' language history, there would be an incomplete picture of the story of how experience with multiple languages impact cognition. Although IAs and overhearers are traditionally considered monolinguals, these studies demonstrate that there is significant variation with second language experience within monolinguals that, if studied, can contribute to our understanding of bilingualism more generally.

Objective measures of speakers' language skills are also needed to not only better characterize bilingual and monolingual samples but also understand the cognitive processes underlying language skill. Objective measures of language proficiency – in addition to self-rated proficiency – should be used and reported to provide a more accurate measure of current language skill (e.g., Tomoschuk et al., in press). Although objective proficiency measures have been found to be correlated with self-rated proficiency (e.g., Marian et al., 2007), the addition of objective measures – especially for aspects of language skill that are particularly relevant for a specific research question – could uncover how bilinguals' diverse language skills may affect cognition as well. For instance, if productive language skills or vocabulary are important aspects of language skill for a particular study, measures such as picture naming tasks (e.g., Multilingual Naming Test; Gollan et al., 2012) and verbal fluency tasks (e.g., Delis et al., 2001) are relatively simple tasks that can be used to objectively measure productive language skills or vocabulary. Additionally, when such objective proficiency measures are combined with cognitive tasks, we can begin to understand what cognitive processes may underlie different language processes (e.g., Zirnstein et al., 2018) – something that is critical to understand in order to uncover the underlying mechanisms of any bilingual differences in cognition. Moreover, objective proficiency measures also better control for cultural differences in self-ratings of language proficiency (e.g., Hoshino and Kroll, 2008; Tomoschuk et al., in press), particularly when comparing multiple groups of bilinguals (e.g., Japanese-English bilinguals vs. Spanish-English bilinguals). Thus, we recommend that future research incorporate both language history questionnaires that capture self-ratings of language proficiency (e.g., LEAP-Q; Marian et al., 2007) as well as objective measures of speakers' language skills to more accurately characterize speakers.

Precise descriptions of speakers' languages, as well as clear definitions for terminology used to describe bilinguals, are also necessary. Although terms such as “native language,” “first language,” and “second language” typically provide information about the order of language acquisition, they are often conflated with other aspects of language skill or status. For instance, these terms may be used to describe a speaker's language dominance (e.g., “native” or “first” language referring to a speaker's most dominant language) or whether a specific language is the majority vs. minority language (e.g., “native” or “first” language referring to the majority language in the community and “second language” referring to the minority language in the community). Bilinguals can also differ in the nature of the two languages

they speak, where some bilinguals' languages differ in phonemic inventories, script, syntactic rules, or even in modality. Moreover, given that some bilinguals have two first languages that were acquired simultaneously (e.g., De Houwer, 1990), and some speakers have first languages that they can no longer speak and/or understand (e.g., Pierce et al., 2014), first vs. second languages can be arbitrary labels for some speakers. Relatedly, there is little consensus on the definition of a "native" language, and even monolinguals can vary widely in the skill they have in their one and only language (e.g., King and Just, 1991; Tanner and Van Hell, 2014). There is also evidence that monolinguals' native language undergoes change when speakers begin to acquire a new language (e.g., Bice and Kroll, 2015), suggesting that the native language is not as stable as once thought. We suggest that future research clearly define terminology used to describe bilinguals and monolinguals so that terms such as "first language" do not conflate order of language acquisition with language skill, status, or other characteristics of bilingual language experience. By both using objective measures of language proficiency and being more precise in our descriptions of speakers' languages, we may be able to understand how diversity in bilinguals' language skill and status are reflected in cognition as well.

Further details about participants' sociolinguistic context would also allow for a deeper understanding of what kinds of speakers were included in a study. Although demographic variables that typically covary with bilingualism – such as socioeconomic status, education, and immigration status – are sometimes reported and/or controlled for in studies (e.g., Morton and Harper, 2007; Carlson and Meltzoff, 2008; Alladi et al., 2013), the context of language use is typically unreported. However, evidence suggests that the larger sociolinguistic context surrounding speakers – both bilingual and monolingual – may affect language and cognition. For instance, a bilingual who speaks a language that is uncommon in their sociolinguistic context would not have as many opportunities to use that language – or switch between their two languages—as a bilingual who speaks a language that is common in their sociolinguistic context. Accordingly, the Adaptive Control Hypothesis (Green and Abutalebi, 2013; Abutalebi and Green, 2016) posits that the ways in which bilinguals use their languages with interlocutors has consequences for language and cognitive control. Indeed, a meta-analysis of studies on the effect of bilingualism on cognition found location-based differences in effect sizes, with effect sizes for studies conducted in Europe being significantly greater than those for studies conducted in the United States and the Middle East (Adesope et al., 2010); such findings suggest that the sociolinguistic contexts within each of these locations may have consequences for the relation between bilingualism and cognition. Moreover, recent evidence from different groups of monolinguals has found that the linguistic diversity of monolinguals' sociolinguistic context impacts infants' social learning (Howard et al., 2014) and preschoolers' language awareness (Atagi, 2018). Altogether, such evidence provides insight into the kinds of language experiences that may be critical when describing research participants. Simply knowing whether individuals are immersed in the first or second language and whether they are proficient or not,

is not sufficient. Although it would be ideal to use methods such as daily diaries and speech recorders (e.g., LENA; Xu et al., 2009) to collect detailed information about speakers' context of language use on a day-to-day basis (which has also been suggested by Laine and Lehtonen, *in press*), these methods are resource-intensive and can be difficult to collect. Thus, minimally, future research should gather information regarding speakers' social networks and communities – along with any available corresponding census data on sociolinguistic context – to better capture speakers' context of language use.

Given the lack of detailed information about participants in the majority of presently published works, it is unsurprising that it is still largely unknown how these different language experiences and skills interact to affect cognition. However, recent research suggests that a complex relation exists between language processing, language regulation, and cognitive control (e.g., Zirnstein et al., 2018). By taking a more nuanced approach to understanding and reporting participants' language backgrounds, we may begin to uncover why and how variability in language background shapes cognition.

ON STUDYING INDIVIDUALS: INDIVIDUAL DIFFERENCES AND CHANGE OVER TIME

A promising direction for the field is to exploit the variability in both current and previous language experiences by examining individual differences – both longitudinally and cross-sectionally – and by conducting more mechanistic studies. As bilingualism is caused by life circumstances rather than experimental ones, bilingualism research has traditionally involved quasi-experimental designs, which is problematic for establishing causality (for a review, see Laine and Lehtonen, *in press*). One way to overcome this problem is by conducting longitudinal studies to control random variation across time in order to isolate the effects of bilingualism due to cumulative language experience. Longitudinal designs have proven to be particularly sensitive to the consequences of bilingualism over the course of development. Santillán and Khurana (2017) followed a large sample of children and used Structural Equation Modeling to predict executive function trajectories starting from the children's entry into the Head Start program until their transition to Kindergarten. The model revealed different trajectories for monolinguals, bilinguals, and learners (i.e., children who were transitioning from monolingual to bilingual classrooms). Children who were bilingual at the beginning of Head Start had the highest executive function performance of the three groups and showed the steepest growth over time. The learners had the lowest performance of all groups but showed more accelerated growth and higher executive function skill at Kindergarten entry compared to their monolingual peers. Longitudinal designs not only reveal that the relation between language and cognition differs across the lifespan, but importantly, they suggest that the effects of bilingual

language experience may impact developmental and learning trajectories.

Although longitudinal designs are especially informative, the expense associated with such a design often precludes its feasibility. One way to overcome this problem is to conduct short-term longitudinal studies or lab-based training studies that expose participants to a second language and to examine the neural or behavioral changes that occur as a result of that exposure (e.g., McLaughlin et al., 2004; Osterhout et al., 2008; Hämäläinen et al., 2017) or to ask what kinds of changes predict successful L2 learning (e.g., Prat et al., 2016). Training studies have also been used to examine how particular bilingual language skills such as language-switching might impact cognitive control (Zhang et al., 2015), providing a causal link for the relationship between aspects of bilingual language experience and executive function. Given the greater experimental control afforded by these approaches, we propose that examining individual differences through learning and training studies will make some important contributions to the field of bilingualism.

CONCLUSION

The controversy involving the “bilingual advantage” has received a great deal of attention in the field with numerous studies addressing the question, and special issues such as this one dedicated to providing productive future directions. In this article, we suggest that much of the controversy in bilingualism research stems from dealing with the variability in bilingual language experiences inappropriately both at theoretical and methodological levels. To study the consequences of knowing multiple languages in its many forms, we must learn to appropriately measure that variation and design studies that can exploit that variation without confounding it with other

factors. First, we suggest that if research findings pose problems for existing accounts, we must actively revise those accounts to accommodate for the complexity of the data. Second, we suggest that sensitively measuring and describing the language histories and skills of participants using behavioral and self-report measures will more accurately allow us to capture the effects of bilingualism. Lastly, we propose that diversifying research design by using more (short or long-term) longitudinal studies and by focusing more on individual differences, we can better evaluate how second language experiences affects cognition while avoiding setbacks of quasi-experimental designs. The recommendations proposed in this paper will enable us to move beyond simple group comparisons and to exploit variation to elucidate the relation among language experience, mind, and brain.

AUTHOR CONTRIBUTIONS

Each co-author contributed equally in writing the manuscript and prepared it for publication.

FUNDING

This work was supported in part by NSF postdoctoral fellowship SBE-1714925 to NA.

ACKNOWLEDGMENTS

We thank Judith Kroll for comments on an earlier version of the manuscript and the UC Riverside Bilingualism, Mind, and Brain lab for their support.

REFERENCES

- Abutalebi, J., and Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Biling. Lang. Cogn.* 19, 689–698. doi: 10.1017/S1366728916000225
- Adesope, O. O., Lavin, T., Thompson, T., and Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Rev. Educ. Res.* 80, 207–245. doi: 10.3102/0034654310368803
- Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., Shukla, A. K., et al. (2013). Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology* 81, 1938–1944. doi: 10.1212/01.wnl.0000436620.33155.a4
- Anastasi, A. (1958). *Differential Psychology: Individual and Group Differences in Behavior*. Oxford: Macmillan.
- Antoniou, M., Liang, E., Ettlinger, M., and Wong, P. C. (2015). The bilingual advantage in phonetic learning. *Biling. Lang. Cogn.* 18, 683–695. doi: 10.1017/S1366728914000777
- Atagi, N. (2018). *The Effects of Community Linguistic Diversity and Multilingualism on Children's Development of Language Awareness*. Doctoral dissertation, Ann Arbor, MI: ProQuest Dissertations Publishing.
- Au, T. K. F., Knightly, L. M., Jun, S. A., and Oh, J. S. (2002). Overhearing a language during childhood. *Psychol. Sci.* 13, 238–243. doi: 10.1111/1467-9280.00444
- Bak, T. H. (2016). Cooking pasta in La Paz. *Linguist. Approaches Biling.* 6, 699–717. doi: 10.1075/lab.16002.bak
- Baum, S., and Titone, D. (2014). Moving toward a neuroplasticity view of bilingualism, executive control, and aging. *Appl. Psycholinguist.* 35, 857–894. doi: 10.1017/S0142716414000174
- Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, and Cognition*. Cambridge: Cambridge University Press.
- Bialystok, E. (2016). The signal and the noise. *Linguist. Approaches Biling.* 6, 517–534. doi: 10.1017/CBO9780511605963
- Bialystok, E., Craik, F. I., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1075/lab.15040.bia
- Bice, K., and Kroll, J. F. (2015). Native language change during early stages of second language learning. *Neuroreport* 26, 966–971. doi: 10.1097/WNR.0000000000000453
- Carlson, S. M., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298. doi: 10.1111/j.1467-7687.2008.00675.x
- Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86. doi: 10.1111/j.1467-7687.2008.00675.x
- de Bruin, A., Treccani, B., and Della Sala, S. (2015). Cognitive advantage in bilingualism: an example of publication bias? *Psychol. Sci.* 26, 99–107. doi: 10.1177/0956797614557866
- De Houwer, A. (1990). *The Acquisition of Two Languages from Birth: A Case Study*. Cambridge: Cambridge University Press. doi: 10.1177/0956797614557866

- Delis, D. C., Kaplan, E., and Kramer, J. H. (2001). *Verbal Fluency Subtest of the Delis-Kaplan Executive Function System*. San Antonio, TX: The Psychological Corporation. doi: 10.1017/CBO9780511519789
- Fan, S. P., Liberman, Z., Keysar, B., and Kinzler, K. D. (2015). The exposure advantage: early exposure to a multilingual environment promotes effective communication. *Psychol. Sci.* 26, 1090–1097. doi: 10.1177/0956797615574699
- Fricke, M., Zirnstein, M., Navarro-Torres, C. A., and Kroll, J. F. (in press). Bilingualism reveals fundamental variation in language processing. *Biling. Lang. Cogn.* 1–8. doi: 10.1017/S1366728918000482
- Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., and Cera, C. M. (2012). Self-ratings of spoken language dominance: a Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Biling. Lang. Cogn.* 15, 594–615. doi: 10.1017/S1366728911000332
- Green, D. W., and Abutaleb, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377
- Hämäläinen, S., Sairanen, V., Leminen, A., and Lehtonen, M. (2017). Bilingualism modulates the white matter structure of language-related pathways. *Neuroimage* 152, 249–257. doi: 10.1016/j.neuroimage.2017.02.081
- Hoshino, N., and Kroll, J. F. (2008). Cognate effects in picture naming: does cross-language activation survive a change of script? *Cognition* 106, 501–511.
- Howard, L. H., Carrazza, C., and Woodward, A. L. (2014). Neighborhood linguistic diversity predicts infants' social learning. *Cognition* 133, 474–479. doi: 10.1016/j.cognition.2014.08.002
- Kaushanskaya, M., and Marian, V. (2009). The bilingual advantage in novel word learning. *Psychon. Bull. Rev.* 16, 705–710. doi: 10.3758/PBR.16.4.705
- King, J., and Just, M. A. (1991). Individual differences in syntactic processing: the role of working memory. *J. Mem. Lang.* 30, 580–602. doi: 10.1016/0749-596X(91)90027-H
- Knightly, L. M., Jun, S. A., Oh, J. S., and Au, T. K. F. (2003). Production benefits of childhood overheard. *J. Acoust. Soc. Am.* 114, 465–474. doi: 10.1121/1.1577560
- Kousaie, S., and Phillips, N. A. (2017). A behavioural and electrophysiological investigation of the effect of bilingualism on aging and cognitive control. *Neuropsychologia* 94, 23–35. doi: 10.1016/j.neuropsychologia.2016.11.013
- Kroll, J. F., and Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *J. Cogn. Psychol.* 25, 497–514. doi: 10.1080/20445911.2013.799170
- Laine, M., and Lehtonen, M. (in press). Cognitive consequences of bilingualism: where to go from here? *Lang. Cogn. Neurosci.* 1–8. doi: 10.1080/23273798.2018.1462498
- Liberman, Z., Woodward, A. L., Keysar, B., and Kinzler, K. D. (2017). Exposure to multiple languages enhances communication skills in infancy. *Dev. Sci.* 20:e12420. doi: 10.1111/desc.12420
- López, B. G., and Vaid, J. (2016). Divergence and overlap in bilingual conceptual representation: does prior language brokering experience matter? *Biling. Lang. Cogn.* 21, 1–12.
- López, B. G., and Vaid, J. (2018). Fácil or A piece of cake: does variability in bilingual language brokering experience affect idiom comprehension? *Biling. Lang. Cogn.* 21, 340–354. doi: 10.1111/14735480
- López, B. G., Vaid, J., Tosun, S., and Rao, C. (2017). Bilinguals' plausibility judgments for phrases with a literal vs. non-literal meaning: the influence of language brokering experience. *Front. Psychol.* 8:1661. doi: 10.3389/fpsyg.2017.01661
- Luk, G., and Bialystok, E. (2013). Bilingualism is not a categorical variable: interaction between language proficiency and usage. *J. Cogn. Psychol.* 25, 605–621. doi: 10.1080/20445911.2013.795574
- Marian, V., Blumenfeld, H. K., and Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Res.* 50, 940–967. doi: 10.1044/1092-4388(2007)067
- McLaughlin, J., Osterhout, L., and Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nat. Neurosci.* 7, 703–704. doi: 10.1038/nn1264
- Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x
- Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., et al. (2008). Second-language learning and changes in the brain. *J. Neurolinguistics* 21, 509–521. doi: 10.1016/j.jneuroling.2008.01.001
- Paap, K. R. (2014). The role of componential analysis, categorical hypothesising, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *J. Cogn. Psychol.* 26, 242–255. doi: 10.1080/20445911.2014.891597
- Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Paap, K. R., Johnson, H. A., and Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *J. Cogn. Psychol.* 26, 615–639. doi: 10.1080/20445911.2014.944914
- Paap, K. R., Myuz, H. A., Anders, R. T., Bockelman, M. F., Mikulinsky, R., and Sawi, O. M. (2017). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *J. Cogn. Psychol.* 29, 89–112. doi: 10.1080/20445911.2016.1248436
- Pierce, L. J., Chen, J. K., Delcenserie, A., Genesee, F., and Klein, D. (2015). Past experience shapes ongoing neural patterns for language. *Nat. Commun.* 6:10073. doi: 10.1038/ncomms10073
- Pierce, L. J., Genesee, F., Delcenserie, A., and Morgan, G. (2017). Variations in phonological working memory: linking early language experiences and language learning outcomes. *Appl. Psycholinguist.* 38, 1265–1300. doi: 10.1017/S0142176417000236
- Pierce, L. J., Klein, D., Chen, J. K., Delcenserie, A., and Genesee, F. (2014). Mapping the unconscious maintenance of a lost first language. *Proc. Natl. Acad. Sci. U.S.A.* 111, 17314–17319. doi: 10.1073/pnas.1409411111
- Poarch, G. J., and van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: evidence from second-language learners, bilinguals, and trilinguals. *J. Exp. Child Psychol.* 113, 535–551. doi: 10.1016/j.jecp.2012.06.013
- Prat, C. S., Yamasaki, B. L., Kluender, R. A., and Stocco, A. (2016). Resting-state qEEG predicts rate of second language learning in adults. *Brain Lang.* 157, 44–50. doi: 10.1016/j.bandl.2016.04.007
- Santillán, J., and Khurana, A. (2017). Developmental associations between bilingual experience and inhibitory control trajectories in Head Start children. *Dev. Sci.* 21:e12624. doi: 10.1111/desc.12624
- Schroeder, S. R., and Marian, V. (2012). A bilingual advantage for episodic memory in older adults. *J. Cogn. Psychol.* 24, 591–601. doi: 10.1080/20445911.2012.669367
- Sebastián-Gallés, N., Albareda-Castellot, B., Weikum, W. M., and Werker, J. F. (2012). A bilingual advantage in visual language discrimination in infancy. *Psychol. Sci.* 23, 994–999. doi: 10.1177/0956797612436817
- Shanahan, T., and Escamilla, K. (2009). English language learners: developing literacy in second-language learners—Report of the National Literacy Panel on Language-Minority Children and Youth. *J. Lit. Res.* 41, 432–452. doi: 10.1080/10862960903340165
- Shin, J., Sailors, M., McClung, N., Pearson, P. D., Hoffman, J. V., and Chilimanjira, M. (2015). The case of Chichewa and English in Malawi: the impact of first language reading and writing on learning English as a second language. *Biling. Res. J.* 38, 255–274. doi: 10.1080/15235882.2015.1091050
- Singh, L., Fu, C. S., Rahman, A. A., Hameed, W. B., Sanmugam, S., Agarwal, P., et al. (2015). Back to basics: a bilingual advantage in infant visual habituation. *Child Dev.* 86, 294–302. doi: 10.1111/cdev.12271
- Sparrow, W., Butvilofsky, S., Escamilla, K., Hopewell, S., and Tolento, T. (2014). Examining the longitudinal biliterate trajectory of emerging bilingual learners in a paired literacy instructional model. *Biling. Res. J.* 37, 24–42. doi: 10.1080/15235882.2014.893271
- Surrain, S., and Luk, G. (in press). Describing bilinguals: a systematic review of labels and descriptions used in the literature between 2005–2015. *Biling. Lang. Cogn.* 1–15. doi: 10.1017/S1366728917000682
- Tanner, D., and Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia* 56, 289–301. doi: 10.1016/j.neuropsychologia.2014.02.002
- Tomoschuk, B., Ferreira, V. S., and Gollan, T. H. (in press). When a seven is not a seven: self-ratings of bilingual language proficiency differ between

- and within language populations. *Biling. Lang. Cogn.* 1–21. doi: 10.1017/S1366728918000421
- Tyler, L. E. (1947). *The Psychology of Human Differences*. New York, NY: D Appleton-Century Company. doi: 10.1037/11271-000
- Valian, V. (2015). Bilingualism and cognition. *Biling. Lang. Cogn.* 18, 3–24. doi: 10.1017/S1366728914000522
- Verreyt, N., Woumans, E. V. Y., Vandelandotte, D., Szmalec, A., and Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Biling. Lang. Cogn.* 19, 181–190. doi: 10.1017/S1366728914000352
- von Bastian, C. C., Souza, A. S., and Gade, M. (2016). No evidence for bilingual cognitive advantages: a test of four hypotheses. *J. Exp. Psychol. Gen.* 145, 246–258. doi: 10.1037/xge0000120
- Weikum, W. M., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastián-Gallés, N., and Werker, J. F. (2007). Visual language discrimination in infancy. *Science* 316, 1159–1159. doi: 10.1126/science.1137686
- Xu, D., Yapanel, U., and Gray, S. (2009). *Reliability of the LENA Language Environment Analysis System in Young Children's Natural Home Environment*. Boulder: LENA Foundation.
- Zhang, H., Kang, C., Wu, Y., Ma, F., and Guo, T. (2015). Improving proactive control with training on language switching in bilinguals. *Neuroreport* 26, 354–359. doi: 10.1097/WNR.0000000000000353
- Zirnsstein, M., van Hell, J. G., and Kroll, J. F. (2018). Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition* 176, 87–106. doi: 10.1016/j.cognition.2018.03.001

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Takahesu Tabori, Mech and Atagi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Working Memory With Emotional Distraction in Monolingual and Bilingual Children

Monika Janus and Ellen Bialystok*

Department of Psychology, York University, Toronto, ON, Canada

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Lior Fern-Pollak,
University College London,
United Kingdom
Julia Ouzia,
Goldsmiths, University of London,
United Kingdom

*Correspondence:

Ellen Bialystok
ellenb@yorku.ca

Specialty section:

This article was submitted to
Cognitive Science,
a section of the journal
Frontiers in Psychology

Received: 29 June 2018

Accepted: 08 August 2018

Published: 28 August 2018

Citation:

Janus M and Bialystok E (2018)
Working Memory With Emotional
Distraction in Monolingual
and Bilingual Children.
Front. Psychol. 9:1582.
doi: 10.3389/fpsyg.2018.01582

Extensive work has demonstrated the benefits of bilingualism on executive functioning (EF) across the lifespan. Concurrently, other research has shown that EF is related to emotion regulation (ER), an ability that is integral to healthy socio-emotional development. However, no research to date has investigated whether bilingualism-related advantages in EF can also be found in emotional contexts. The current study examined the performance of 93 children who were 9-years old, about half of whom were bilingual, on the Emotional Face *N*-Back Task, an ER task used to assess the interference effect of emotional processing on working memory. Bilingual children were more accurate than monolingual children in both 1-back and 2-back conditions but were significantly slower than monolingual children on the 2-back condition. There were significant effects of emotional valence on reaction time, but these did not differ across language groups. These results confirm previous research showing better EF performance by bilinguals, but no differences in ER were found between language groups. Findings are discussed in the context of our current understanding of the ER literature with potential implications for previously unexplored differences between monolingual and bilingual children.

Keywords: bilingualism, executive control, emotional regulation, working memory, *n*-back

INTRODUCTION

Flexible and effective emotion regulation (ER) is critical for healthy psychosocial adjustment throughout development (Cole and Deater-Deckert, 2009; Eisenberg et al., 2010). The inability to properly regulate emotions, or emotional dysregulation, has been shown to underlie a range of maladaptive outcomes including aggressive behavior problems (e.g., Stieben et al., 2007; Lewis et al., 2008; Holley et al., 2017) and academic underachievement (e.g., Hawkins et al., 1999; Gumora and Arsenio, 2002; Djambazova-Popordanoska, 2016) in children. Although the past several decades have seen a steady increase in research investigating the consequences of maladaptive ER, factors that promote the development of adaptive ER remain poorly understood. However, there is evidence that ER is highly interrelated with executive functioning (EF; e.g., for a review, see Zelazo and Cunningham, 2007; Calkins and Marcovitch, 2010), and that individual differences in EF are predictive of ER abilities (e.g., Kieras et al., 2005; McDermott et al., 2009; for a review, see Schmeichel and Tang, 2015). As such, individual factors that enhance EF may be expected to promote the development of adaptive ER abilities.

Although seemingly distinct from ER, bilingualism, or proficiency in a second language (L2) is associated with advantages on a variety of EF tasks (see Adesope et al., 2010, for a meta-analysis;

for a review, see Barac et al., 2014). Some EF tasks administered in research on bilingualism have also been used in ER research to assess cognitive control as it interacts with emotional processing (e.g., Bell and Wolfe, 2004; for a review, see Cole et al., 2004). Examining the interrelation between cognitive and emotional processing (Bell and Wolfe, 2004) may be key to explaining how proficiency in an L2 may also promote ER. The current study investigates whether bilingualism supports more adaptive ER strategies in school-aged children than is found for monolingual children.

Literature on ER has attracted significant attention due to its association with a variety of important developmental outcomes (for review, see Gross, 2002; Eisenberg et al., 2010). Although some controversy continues to exist over its constituents, Calkins and Hill (2007) view ER as the range of conscious and unconscious behaviors, skills, and strategies that change one's emotional experience and expression either in an automatic or effortful way. Most definitions of ER also recognize interacting emotional and cognitive processes as integral to ER (for review, see Calkins and Marcovitch, 2010). These cognitive operations, including working memory, fall under the umbrella term of EF (Miyake et al., 2000; for a review, see Zelazo and Carlson, 2012). Working memory has been defined as a "cognitive system in which memory and attention interact to produce complex cognition" (Shipstead et al., 2015) and is a pivotal component of the EF system (Miyake et al., 2000). Working memory not only requires memory updating and retention but also relies on attentional control, which can vary between task conditions and challenge the EF system to different degrees (Miyake et al., 2000). Within ER contexts, EF interacts with emotional processing to modify appraisals, feelings, and behaviors in response to emotional experiences (for review, see Zelazo and Cunningham, 2007; Calkins and Marcovitch, 2010). Optimal ER development thus depends on the acquisition of cognitive skills such as working memory that allow the child to focus on task-relevant information with minimal interference from distracting and non-goal-oriented cues.

Research on emotion and working memory has primarily focused on adult clinical populations, but other research has investigated the emotion-cognition interaction in both clinical and non-clinical samples across development (Bradley et al., 1999; Bar-Haim et al., 2007; Ladouceur et al., 2009). One of the most popular paradigms used to assess working memory is the *n*-back task (for a review, see Owen et al., 2005; Meule, 2017). In this task, participants are asked to recall whether the location or identity of a target on the screen (e.g., the letter M) matches the location or identity of a stimulus presented *n* trials previously; memory load increases as *n* increases. Emotional faces (angry, sad, fearful, happy, neutral/calm) are often adopted as distracting emotional cues.

Studies investigating the impact of emotional valence on working memory report mixed findings. This is likely due to the extensive variability in the populations studied (typically vs. atypically developing, and different age groups), making it difficult to draw parallels between findings and particularly challenging to make predictions for typically developing children. Some authors find only significant slowing in response to

negative emotional cues relative to positive or neutral ones (*healthy adult sample*: Kensinger and Corkin, 2003; *anxious sample of 8- to 30-year-olds*: Ladouceur et al., 2009), whereas others report no reaction time (RT) differences by emotion type and instead report impaired accuracy on trials with negative compared to neutral distractors (*adult controls and ADHD participants*: Marx et al., 2011). Others have reported varying speed-accuracy trade-offs by emotion type, including higher accuracy but slower RTs for negative compared to neutral stimuli in a non-verbal working memory task with a sample of schizophrenic participants (Becerril and Barch, 2011). This finding supports previous work showing that in emotionally dysregulated populations, aversive stimuli generate a significantly larger burden on the cognitive system than positive ones, depleting resources available for working memory (Bishop et al., 2004; Hare et al., 2005). For example, emotionally dysregulated, clinically depressed patients report an inability to disengage from pervasive negative thoughts (for a review, see Gotlib and Joormann, 2010), resulting in memory challenges as well as difficulty with planning and concentration (Paelecke-Habermann et al., 2005; Rose and Ebmeier, 2006).

Similar tendencies have been reported among clinically anxious and depressed children, showing differences in performance compared to healthy controls. Ladouceur et al. (2005) administered the Emotional *N*-back Task to a sample of 75 children (8–16 years of age) categorized into one of four groups: children who met criteria for an anxiety disorder, major depressive disorder, comorbid anxiety and depression, or were identified as a normal control group. In this version of the task, the distracting emotional stimuli were neutral, negative, and positive images in the background of the to-be-remembered letters. Their results showed that children with major depressive disorder and those with comorbid anxiety and depression had significantly longer RTs on the negative condition than on the neutral condition, whereas children in the normal control group had significantly longer RTs on the positive condition than on the neutral condition. Ladouceur et al. (2009) took a developmental approach to demonstrate ER changes with age within a sample of 8- to 30-year-old participants with low versus high levels of trait anxiety. The authors used the emotional face *N*-back task with 0-back and 2-back memory load conditions and three emotional face distractor types (neutral, fearful, and happy) as well as a control condition with shapes. Their findings revealed that individuals high in trait anxiety had slower RTs on the fearful 2-back memory-load condition than on the happy and neutral trials, but that the effect was greatest in younger participants. Conversely, individuals low in trait anxiety did not reveal any emotion effects, either in RT or accuracy rates. Taken together, these findings highlight that there are differences between how children and adults process distracting emotional information and that we continue to find inconsistent results when investigating interacting cognitive and emotional processing in typically developing children.

From a separate area of research, the cognitive benefit for bilingual individuals has been identified as enhanced performance on tasks requiring non-verbal EF (for a review, see

Bialystok, 2017). Improvement in EF is believed to develop as a result of the well-documented coactivation of both languages within the bilingual brain, even when only one language is in use (e.g., Beauvillain and Grainger, 1987; Colomé, 2001; for a review, see Kroll et al., 2014). The practice of attending to one cue (one language) during interference from another “trains” the EF network (for a review, see Bialystok, 2015), becoming more effective throughout life and thereby extending the practice of verbal cognitive control to the non-verbal EF network (Green, 1998; Abutalebi et al., 2008; Luk et al., 2010). Evidence for the enhanced cognitive control in bilinguals comes from research using a variety of cognitive tasks with infants (Kovacs and Mehler, 2009), toddlers (Poulin-Dubois et al., 2011), young children (e.g., see Adesope et al., 2010, for a meta-analysis), and young adults (e.g., Costa et al., 2009). Recent findings have emerged that do not support these results with young adults (e.g., Paap and Greenberg, 2013) possibly due to differences in populations, criteria for bilingualism, or the nature of the experimental tasks used to assess cognitive ability (for a review, see Antoniou, 2019). However, the majority of the research points to bilingual benefits across a variety of cognitive control operations, especially where conflict conditions pose additional attentional demands on the EF system.

Given the central importance of working memory to EF, some bilingualism research has investigated differences in verbal and non-verbal working memory in children and adults, with mixed findings. Morales et al. (2013) conducted two studies that assessed working memory in 5-year-old (Study 1) and 5- to 7-year-old (Study 2) monolingual and bilingual children. In Study 1, the authors found that differences in performance between the groups emerged only on the most challenging condition of a Simon-type task, with bilingual children showing an EF advantage when a high level of conflict was present. In the second study, where children were required to recall the positions of frogs presented either simultaneously (easy) or sequentially (hard) within a 3×3 grid (Frog Matrices task), bilingual children had better accuracy on the more challenging sequential condition. Blom et al. (2014) investigated both visuospatial (Dot Matrix and Odd-One-Out tasks) and verbal working memory (Forward and Backward Digit Recall) performance in bilingual Turkish–Dutch children and Dutch monolingual controls from low socioeconomic backgrounds. Although no difference was found between the two language groups in 5-year-old children, by 6 years of age bilingual children showed overall benefits on the Dot Matrix task and the Backward Digit Recall task, both of which pose additional demand for EF over the other two tasks. While some have failed to reproduce this effect using simpler working memory measures (Engel de Abreu, 2011), bilingual children show advantages over their monolingual peers on conditions of heightened conflict.

In summary, several lines of evidence depict bilingual advantages in EF on tasks where successful performance depends on the ability to resolve conflict from competing cues and ignore interfering information or to maintain task rules in working memory. Ultimately, by monitoring language choice among competing linguistic systems, bilinguals must learn to more effectively regulate attention to distracting information, resulting

in an EF system that is better equipped to support processes of working memory.

What are the implications of bilingualism for ER? Importantly, emotional and cognitive processes are highly interactive and integral to ER in that effective ER in emotional contexts depends on the EF system to process relevant information without being impaired by interfering emotional cues (Gray et al., 2002). Concurrently, literature on bilingualism shows evidence of strengthened cognitive control in dual-language users, resulting in greater selective attention to relevant information and reduced interference from distracting cues (Bialystok, 2015). It is thus reasonable to hypothesize that bilingualism may promote the development of more adaptive ER by strengthening the cognitive control system and all its constituents, including working memory. Furthermore, if bilingualism contributes to the development of self-regulatory abilities in emotionally challenging contexts at earlier stages of development than is found for monolinguals, then these enhanced abilities may also have implications for children’s psychosocial outcomes. However, research assessing ER differences between monolingual and bilingual children in this manner, and more comprehensively using standardized ER tasks, is largely lacking, and a direct evaluation of ER differences between these groups has not been undertaken.

The present study aimed to investigate the effect of bilingualism on cognitive and emotional processing that is integral to ER. The Emotional Face *N*-Back Task, an emotionally based EF task of working memory with three emotion conditions (angry, happy, and neutral), was used to examine differences in ER in school-aged monolingual and bilingual children. The overarching hypothesis was that bilingual children would demonstrate an overall advantage in working memory. Given that an EF advantage was expected in the bilingual group, and that the ability to modulate attention toward or away from emotionally salient information is a marker of ER and associated with EF, we anticipated finding evidence for ER benefits for bilinguals. Although the findings in this area with healthy children are mixed, it was predicted that the ER benefits for bilinguals would be most salient on the particularly challenging angry emotion trials, with the highest EF demands. This is the first study to compare these processes in healthy young children and evaluate the influence of bilingualism on ER. The novelty of this research will contribute to our understanding of differences in emotional processing between groups of children with different language experiences, over and beyond the known advantage of cognitive control in bilingual children.

MATERIALS AND METHODS

Participants

One hundred and two children between 8- and 11-years old were recruited from four elementary schools. Based on caregiver reports of the child’s language background, an aggregate score was created to classify children as monolingual or bilingual. Nine children were removed from the study due to behavioral concerns that prevented them from completing the tasks. Complete data

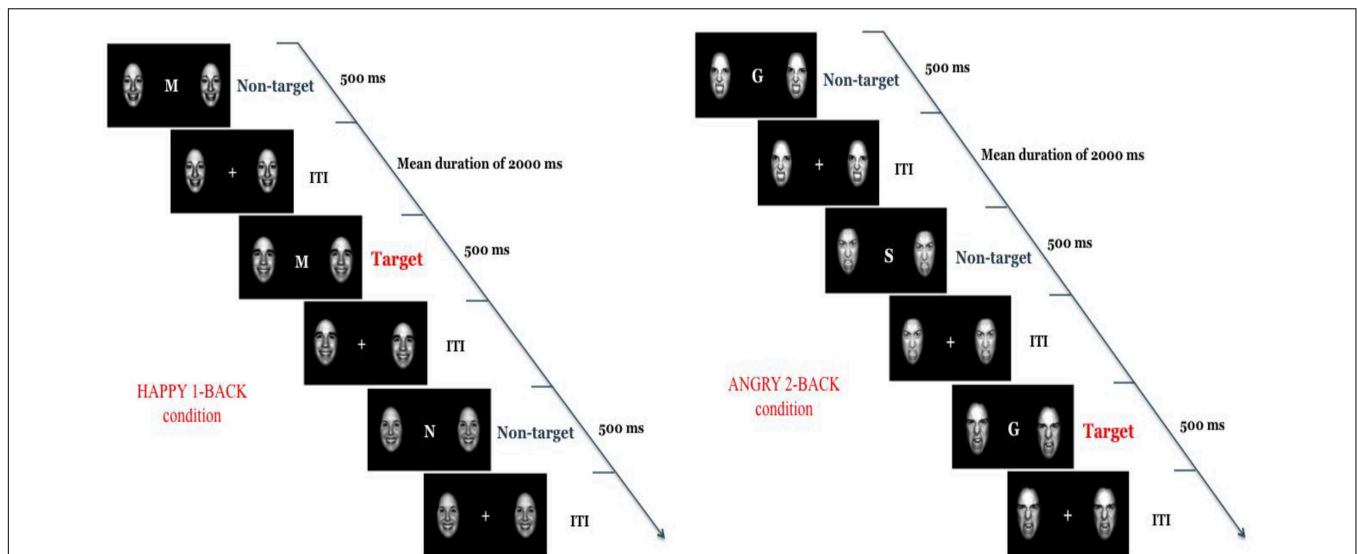


FIGURE 1 | Emotional Face *N*-Back Task (adapted from Ladouceur et al., 2009). Children see one letter at a time presented on the screen and are asked to indicate whether the letter is the same as the letter they saw one screen back (1-back) or two screens back (2-back), creating two memory conditions. A face displaying an angry, happy, or neutral expression is shown on both sides of the letter.

for analysis were available for 93 children, 48 monolinguals (M age = 9.3 years, SD = 0.6; 18 boys) and 45 bilinguals (M age = 9.4 years, SD = 0.5; 20 boys). The majority of children were born in Canada (78.5%), with 10 children born in the Philippines (10.8%) and the remainder being born in 10 different countries. Children in the bilingual group proficiently spoke a non-English language at home: Portuguese (n = 15), Philippine dialect (n = 12; Tagalog, Visayan, or Ilonggo), Italian (n = 5), Spanish (n = 6), or seven other different languages (n = 7). School instruction was in English for all children.

Procedure

Approval to test in the schools was obtained from the University Ethics Committee and from the school board's ethics committee. The principal and teachers at each school agreed to have researchers introduce the study tasks to the children within their own classrooms. A packet of questionnaires was sent home with each child so that interested parents could complete the parental informed consent, the Language and Social Background Questionnaire (LSBQ), the Strengths and Weaknesses of Attention-Deficit/Hyperactivity Disorder Symptoms and Normal Behavior Scale (SWAN), and the Emotion Regulation Checklist (ERC). Before working with a child, qualified research assistants ensured that the complete packet had been returned to the teacher. Teachers were also asked to complete the ERC for each child that returned the packet of questionnaires.

Each child who returned a completed packet to their school was withdrawn from their classroom for approximately 45 min to complete the testing session. The procedure was explained to the child, and verbal assent was obtained prior to testing. During the session, each participant completed the Peabody Picture Vocabulary Test (PPVT), a standardized test of English proficiency, the Raven Standard Progressive Matrices (Raven),

assessing spatial reasoning, and the Emotional Face *N*-Back Task (Figure 1), a task of ER. Upon completion of the *n*-back task, children were asked to subjectively rate the expression of a subset of angry, happy, and neutral faces that they had seen during the task to assess whether all children perceived the faces similarly. Throughout the session, children received stickers for completing each task. Research assistants made ongoing notes during the session to identify children whose behavior (talking, singing, refusal to continue, excessive fidgeting or movement, etc.) interfered with their ability to complete the tasks; these children were later removed from the study (n = 9). Each child was thanked for their participation and awarded a personalized certificate to recognize their effort before being walked back to their classroom.

Questionnaires and Tasks

Language and Social Background Questionnaire (LSBQ; Anderson et al., 2018)

The LSBQ is completed by parents/guardians and contains questions pertaining to the child's age, sex, handedness, time spent using video/computer games, and language fluency and use in different contexts. Parental education is indicated and used as a proxy for socioeconomic status (SES). SES was assessed as the average of mother's and father's education, using a 5-point scale with 1 indicating no high school diploma, 3 indicating some college or college diploma, and 5 indicating graduate or professional degree.

Peabody Picture Vocabulary Test (PPVT; Dunn and Dunn, 1997)

The PPVT is a standardized measure of receptive vocabulary. Children hear a word and are required to point to which one of four pictures corresponds with that word. Testing proceeds

until the child makes eight errors within a block of 12 words. The PPVT normally takes 15–20 min to complete. Scores are standardized based on the participant's age ($\mu = 100$, $SD = 15$). The PPVT has a high reliability (>0.90) across a variety of measures (i.e., internal consistency, split-half, test-retest) and a 0.91 correlation with the Wechsler's Intelligence Scale for Children's measure of verbal ability (WISC-III; Wechsler, 1991).

Raven Standard Progressive Matrices (Raven Test; Raven et al., 1996)

The Raven test is a standardized test of non-verbal spatial reasoning. Children view test figures and chose which item from a set of six options provides the best completion. The task normally takes 10–15 min to complete with children. Results are converted to standardized scores based on the participant's age ($\mu = 100$, $SD = 15$). The predictive validity of the Raven test is around 0.70, whereas test-retest reliability and internal consistency coefficients range between 0.80 and 0.93 (Raven et al., 1996).

Strengths and Weaknesses of Attention-Deficit/Hyperactivity Disorder Symptoms and Normal Behavior Scale (SWAN; Swanson et al., 2001)

The SWAN questionnaire is completed by a child's parent or guardian and teacher. The SWAN includes 18 items that are associated with the characteristic symptoms assessed for a diagnosis of ADHD as described in the DSM-5 (American Psychiatric Association, 2013). These include nine symptoms related to *inattention* (e.g., "Stays focused on tasks and activities"), six symptoms related to *hyperactivity* (e.g., "Can sit without constant fidgeting or squirming"), and three items related to *impulsivity* (e.g., "Easily waits turn, such as standing in line-ups"). Each item is positively worded and was modified slightly from the original test to improve ease of reading and decrease word difficulty for parents/guardians who may struggle with understanding English (e.g., "Sustains attention on tasks or play activities" was changed to "Stays focused on tasks and activities"). A guardian and teacher rated the child on each item using a 4-point Likert scale ranging from "Far below average" (1) to "Far above average" (4). Higher scores are indicative of better attentional abilities, lower hyperactivity, and lower impulsivity. The SWAN has excellent internal consistency and reliability (Young et al., 2009; Lakes et al., 2012).

Emotion Regulation Checklist (ERC; Shields and Cicchetti, 1997)

The ERC is a 24-item measure intended to assess the frequency of children's displays of affective behaviors. Parents/caregivers and teachers rate the frequency of the behavior using a 4-point scale. The raw scores generate two subscales: (1) ER, which assesses socially appropriate emotional responses and empathy, and (2) lability-negativity, which assess arousal more broadly, focusing on anger, dysregulation, and mood lability. High internal consistency has been shown for both the lability-negativity and ER subscales, with Cronbach's alphas of 0.96 and 0.83, respectively (Shields and Cicchetti, 1997).

Emotional Face N-Back Task (N-back; Adapted From Ladouceur et al., 2009)

The emotional variant of the *n*-back paradigm is designed to examine the interference effect of emotional information on working memory performance. The task consisted of two memory conditions (1-back and 2-back), with blocked emotional (angry, happy) and neutral conditions, for each level of difficulty. Letters were presented in the middle of the screen and two of the same emotional faces were presented simultaneously on both sides of the letter to act as the emotional distractors (see Figure 1). In the 1-back condition, participants were asked if the letter was the same as the letter on the previous trial (target, "yes") or not (non-target, "no"). In the 2-back condition, participants decided whether the current letter matched the trial that was presented *two* trials previously (target) or not (non-target). Responses were made using two mice, one assigned to each response, with the dominant hand assigned to target trials and the non-dominant hand to non-target trials. Angry, happy, and neutral faces were taken from the NimStim set available at www.macbrain.org (Tottenham et al., 2009), and modified so that only an oval-shaped face was visible, without hair or a neck. Each emotion block was made up of 15 target ("yes") and 25 non-target ("no") trials. The task took approximately 15 min to complete.

Affect Rating

After completing the *n*-back task children were presented with the angry, happy, and neutral expressions of three NimStim actors whose faces they had seen during the task. The NimStim actors were two females and one male, all demographically diverse. Children chose one adjective to describe the expression on each face without being told whether the face was meant to portray a happy, angry, or neutral expression. The purpose was to assess whether there were differences in how monolingual and bilingual children perceive emotional expressions. The top three descriptive words used to identify each emotional face were compared between the two language groups. The findings were also used to determine whether the child descriptions found in the current study replicated previous findings from the child literature depicting neutral faces as more aversive to children than happy faces.

RESULTS

The background measures for age, SES (parental education), vocabulary knowledge (PPVT), and nonverbal cognitive functioning (Raven test), are reported in Table 1. One-way ANOVAs for language group showed no differences between children in the two groups on any of these measures (all $ps > 0.14$). Mean scores on the subscales of the SWAN (attention, hyperactivity, impulsivity) and the ERC (ER, negativity/lability) are reported in Table 2 for teacher and parent/guardian reports. One-way ANOVAs for language group showed no differences between the teacher ratings for children in the two groups on any of the subscales (all $ps > 0.31$), but parent ratings revealed that monolingual and bilingual children were rated similarly on hyperactivity ($p = 0.76$), impulsivity ($p = 0.99$), ER ($p = 0.63$),

TABLE 1 | Mean score, standard deviation, and range for background measures by language group.

Background measure	Monolingual (<i>n</i> = 48)		Bilingual (<i>n</i> = 45)	
	<i>M</i> (SD)	Range	<i>M</i> (SD)	Range
Age in months	9.3 (0.6)	8.3–10.5	9.4 (0.5)	8.3–10.3
SES*	3.4 (1.0)	2–5	3.1 (1.3)	1–5
PPVT	101.6 (13.5)	81–135	97.5 (13.2)	67–125
Raven test	101.3 (14.9)	75–130	97.8 (12.9)	75–120

*SES (socioeconomic status) was measured as the average of maternal and paternal education level (3 = completed college).

TABLE 2 | Mean score and standard deviation for reports made by teachers and parents on children's behavior by language group.

Teacher and parental reports on child behavior	Monolingual (<i>n</i> = 48)		Bilingual (<i>n</i> = 45)	
	Teacher <i>M</i> (SD)	Parent <i>M</i> (SD)	Teacher <i>M</i> (SD)	Parental <i>M</i> (SD)
SWAN (out of 4)				
Attention	2.8 (0.8)	2.8 (0.7)	3.0 (0.8)	3.1 (0.6)*
Hyperactivity	3.1 (0.9)	3.0 (0.7)	3.2 (0.7)	3.1 (0.8)
Impulsivity	3.2 (0.9)	3.0 (0.7)	3.2 (0.8)	3.0 (0.8)
ERC (out of 4)				
Emotion regulation	3.2 (0.6)	3.3 (0.4)	3.2 (0.6)	3.4 (0.4)
Negativity/lability	3.4 (0.6)	3.2 (0.4)	3.5 (0.5)	3.3 (0.4)

*Significant difference in ratings between language groups, $p < 0.05$.

and negativity/lability ($p = 0.41$), but differently on attention, $F(1,90) = 4.53$, $p = 0.04$, $\eta_p^2 = 0.05$, with bilingual children (3.1) being rated as more attentive than monolingual children (2.8).

The outcomes for accuracy and RT on the Emotional Face *N*-back task are reported in **Table 3** and **Figure 2**, respectively. Accuracy on correct target trials was analyzed using a three-way ANOVA for *n*-back condition (1-back, 2-back), emotion (angry, happy, neutral), and language group (monolingual, bilingual). The analysis revealed a main effect of condition, $F(1,91) = 239.97$, $p < 0.001$, $\eta_p^2 = 0.73$, with children scoring higher on the 1-back (75.78%) than on the more challenging 2-back (52.70%) condition, and a main effect of language group, $F(1,91) = 9.71$, $p < 0.01$, $\eta_p^2 = 0.08$, with bilingual children (67.83%) outperforming monolingual children (60.6%). There was no main effect of emotion, $p = 0.26$, and no interactions, all $F_s < 1.37$, $p_s > 0.26$.

Accuracy on *nontarget* trials was also investigated to determine whether the higher accuracy scores on target trials for the bilingual group reflected a response bias to say “yes” (i.e., identify more trials as *target* trials). Nontarget trials were analyzed using a three-way ANOVA for condition, emotion, and language group. The results revealed only a main effect of *n*-back condition, $F(1,91) = 27.81$, $p < 0.001$, $\eta_p^2 = 0.22$, with children correctly identifying more nontarget trials (“no” response) on the 1-back (87.48%) than on the 2-back (79.04%), as expected.

TABLE 3 | Mean score and standard deviation for accuracy on the Emotional Face *N*-back Task by language group.

Working memory by emotion condition	Monolingual (<i>n</i> = 48)	Bilingual (<i>n</i> = 45)
	<i>M</i> (SD)	<i>M</i> (SD)
Accuracy on target (“yes”) trials (% correct)		
1-back		
Angry	71.4 (15.3)	78.67 (12.5)*
Happy	73.9 (13.1)	81.6 (12.3)*
Neutral	71.5 (17.9)	77.6 (14.4)*
2-back		
Angry	49.4 (17.6)	54.5 (21.5)*
Happy	49.0 (17.5)	57.0 (17.6)*
Neutral	48.8 (17.8)	57.5 (14.9)*
Accuracy on nontarget (“no”) trials (% correct)		
1-back		
Angry	86.8 (16.1)	90.2 (12.3)
Happy	86.5 (17.0)	90.0 (12.1)
Neutral	85.5 (15.0)	86.1 (14.7)
2-back		
Angry	80.5 (17.3)	78.0 (18.8)
Happy	78.8 (16.5)	80.0 (13.8)
Neutral	78.3 (18.4)	78.6 (15.4)

*Significant difference in ratings between language groups, $p < .05$.

There was no main effect of language group ($p = 0.64$) or emotion ($p = 0.16$), or any interactions, all $F_s < 0.60$, $p_s > 0.35$.

Reaction time data for the Emotional Face *N*-back task were analyzed the same way as accuracy data, using a three-way ANOVA for *n*-back condition, emotion, and language group on target trials (see **Figure 2**). There was a main effect for condition, $F(1,91) = 21.86$, $p < 0.001$, $\eta_p^2 = 0.18$, with children performing slower on the more challenging 2-back (958 ms) than the 1-back (868 ms). A main effect of emotion, $F(1,186) = 13.03$, $p < 0.001$, $\eta_p^2 = 0.13$, revealed that RTs were significantly slower on the neutral trials (949 ms) than on angry (902 ms), $p < 0.001$, or happy (886), $p < 0.001$, trials, with no difference between the latter two, $p = 0.57$. Furthermore, a two-way interaction of *n*-back condition and emotion, $F(2,182) = 16.71$, $p < 0.001$, $\eta_p^2 = 0.08$, revealed that the effect of emotion was present on the 1-back condition, $F(2,182) = 39.09$, $p < 0.001$, $\eta_p^2 = 0.10$, but not on the 2-back condition, $p = 0.89$. Finally, a main effect of language group, $F(1,91) = 5.46$, $p = 0.02$, $\eta_p^2 = 0.07$, revealed that bilingual children (956 ms) were significantly slower than their monolingual peers (870 ms), but a two-way interaction of condition and language group, $F(1,91) = 9.03$, $p < 0.01$, $\eta_p^2 = 0.08$, restricted this difference to the 2-back condition $F(1,91) = 9.10$, $p < 0.01$, $\eta_p^2 = 0.10$, with no difference between groups in the 1-back condition, $p = 0.23$.

A correlation was computed between accuracy and RT for each condition to determine whether there were speed-accuracy trade-offs. There were no significant correlations in the 1-back condition, $r(93) = 0.19$, $p = 0.16$, but the relation was significant in the 2-back condition, $r(93) = 0.34$, $p = 0.001$.

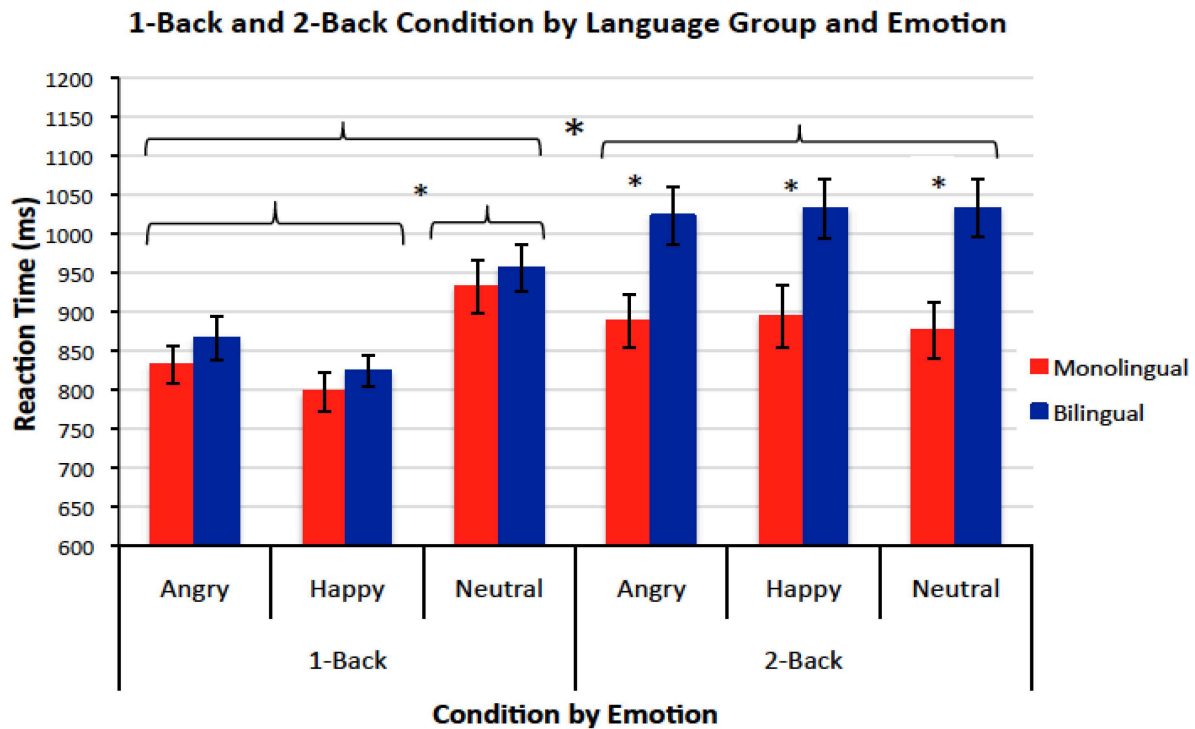


FIGURE 2 | Mean reaction times (and standard errors) on the Emotional Face *N*-back Task by condition (1-back vs. 2-back), emotion (angry, happy, neutral), and language group (monolingual vs. bilingual).

Given that bilingual children performed significantly slower than monolingual children on the 2-back, a correlation was run separately by language group to determine whether the speed-accuracy trade-off in the 2-back was driven by the bilingual group. The correlation revealed that the speed-accuracy trade-off was significant for the bilingual children, $r(45) = 0.34$, $p = 0.02$, but only marginal for monolingual children, $r(48) = 0.24$, $p = 0.09$. However, the Fisher r -to- z transformation revealed that the difference between the two correlations was not significant, $p = 0.61$.

Reaction time on *nontarget* trials was also investigated using a three-way ANOVA for *n*-back condition, emotion, and language group. There was a main effect of emotion, $F(2,182) = 19.29$, $p < 0.001$, $\eta_p^2 = 0.18$, with children performing more slowly on neutral trials (969 ms) than on angry (909 ms) or happy (901 ms) trials, all $ps < 0.001$, with no difference in speed of performance between angry and happy emotions, $p = 1.00$. A main effect of language group was also found, $F(1,91) = 5.30$, $p = 0.02$, $\eta_p^2 = 0.07$, in which bilingual children (971 ms) were slower than their monolingual peers (880 ms). An interaction of *n*-back condition by emotion, $F(2,182) = 21.06$, $p < 0.001$, $\eta_p^2 = 0.20$, revealed that differences in speed of responding between emotion blocks emerged only on the 1-back version of the task, $F(2,182) = 43.13$, $p < 0.001$, $\eta_p^2 = 0.42$, where children performed slower on angry (900 ms) than happy (856 ms) trials, slower on neutral (995 ms) than angry trials, and slower on happy than neutral trials, all $ps < 0.01$; no differences between angry (918 ms), happy (946 ms), and neutral (944 ms) trials were found on the 2-back

version of the task, all $ps > 0.14$. No other main or interaction effects were found, all $F_s < 2.5$, $ps > 0.14$.

Children's affect ratings of nine preselected facial expressions (three per emotional type) used in the task were recorded and evaluated. The three most frequently occurring words used by children in each language group were tabulated by facial expression and emotion (Table 4). These top three words were then inspected to determine whether there were notable differences in the valence of the words within each emotion block and across language groups. The word inspection revealed that children used all positively valenced words (e.g., happy, excited, joyful) to describe the three standardized happy faces, and they rated all standardized angry faces using negatively valenced words (e.g., angry, mad, scary, furious), as expected. The three standardized neutral faces rated by the children generated the greatest amount of variety in valence. Children described neutral faces using words with a positive (i.e., happy), negative (i.e., shocked, sad, scared, serious), and neutral (i.e., normal, bored, no emotion/expression) valence. No outstanding differences were detected between language groups in the words children selected to describe any of the actors' faces in either emotion block.

DISCUSSION

The present study investigated the interrelation between cognitive and emotional processing in typically developing

TABLE 4 | Three most commonly used words by children to describe the emotional expressions of three actors with standardized angry, happy, and neutral affects viewed during the Emotional Face *N*-back Task by language group.

Standardized <i>n</i> -back faces by emotion	Actor 1 (Asian, female)		Actor 2 (African American, female)		Actor 3 (Caucasian, male)	
	Monolingual (<i>n</i> = 48)	Bilingual (<i>n</i> = 45)	Monolingual (<i>n</i> = 48)	Bilingual (<i>n</i> = 45)	Monolingual (<i>n</i> = 48)	Bilingual (<i>n</i> = 45)
Angry	Angry (34) Mad (14) –	Angry (36) Mad (9) Scary (1)	Angry (31) Mad (13) Furious (2)	Angry (30) Mad (11) Frustrated (1) Furious (1)	Angry (22) Mad (17) Frustrated (2)	Angry (18) Mad (22) Furious (2)
Happy	Happy (48) –	Happy (46) –	Happy (35) Excited (6) Energized (1) Joyful (1)	Happy (36) Excited (10) –	Happy (46) Glad (1) Joyful (1)	Happy (40) Excited (2) Silly (2) Cheerful (2)
Neutral	Surprised (19) Shocked (5) Sad (4)	Surprised (16) Shocked (6) Scared (5)	Sad (11) Bored (9) Normal (7)	Sad (19) Bored (6) Normal (5)	Normal (14) Bored (5) No emotion/ expression (4)	Normal (14) Happy (7) Serious (4)

monolingual and bilingual children. Children in the two language groups were similar on age, SES, English proficiency, and nonverbal cognitive functioning. Parents provided information on children's language background, and both parents and teachers reported on children's emotional and behavioral functioning. All children were tested using the Emotional Face *N*-Back Task, an ER task assessing working memory within an emotional context.

Ratings of attention obtained from the SWAN showed that all children were rated similarly by teachers, but bilingual children were rated by parents as significantly more attentive than monolingual children. Although it is common to find discrepancies between informants when gathering ratings on children's behavior (Gresham et al., 2010; Wray et al., 2013), little is known about characteristics that predict discrepancies in ratings. Nonetheless, parental ratings indicating greater attention in bilingual children bring to light an important consideration, namely the frame of reference of the informant. Teachers have experience with children from different communities and cultures and so have a wide frame of reference for rating children's performance relative to their same-aged peers. Parents, in contrast, may be limited to observing the children living in their own community or even household. Furthermore, parents of monolingual or bilingual children may have different culture-specific expectations that influence their parenting practices and expectations for normative development. Thus, while it was beyond the scope of this study, future bilingualism researchers may consider gathering information on familial expectations and background as these may be related to differences between language groups in parental ratings of children's behavior.

The Emotional Face *N*-back Task was used to investigate differences in ER between monolingual and bilingual children by manipulating working memory load (1-back or 2-back) and emotional distraction (angry, happy, and neutral faces). As expected, all children were more accurate and faster on the easier condition than on the more challenging condition of the task. Also in line with our predictions, bilingual children demonstrated better cognitive performance on both working

memory conditions. Specifically, depending on the emotion block and condition, accuracy rates of bilingual children ranged from 6% to 9% higher than those of their monolingual peers. This observed working memory advantage for bilingual children supports research highlighting the cognitive benefits of bilingualism on a variety of EF tasks (Adesope et al., 2010).

The findings also revealed significantly slower RTs on neutral emotion trials than on angry and happy trials for children in both language groups, but only on the 1-back condition. Thus, the easy working memory condition showed differences between the emotional stimuli but the difficult working memory condition did not, presumably because the effort required for the working memory response overwhelmed the more subtle difference between emotion conditions. As in previous research with emotional *n*-back tasks (e.g., Ladouceur et al., 2005; Cromheeke and Mueller, 2016; Villemonteix et al., 2017), accuracy of responses was not impacted by the emotion condition.

The longer RTs on neutral trials may reflect the challenges children experienced in accurately labeling neutral faces. When asked to generate affect ratings for the angry, happy, and neutral emotional expressions for the subset of stimuli used in the task, all children gave accurate positively valenced ratings to happy faces and negatively valenced ratings to angry faces, but neutral faces elicited variable responses, ranging from positively to negatively valenced descriptions. This difficulty in interpreting neutral faces has been found on previous emotion recognition tasks (Kujawa et al., 2014; see Herba et al., 2006; Thomas et al., 2007, for research on age-related changes in emotion recognition). Consequently, the slow responses to neutral emotion trials may reflect children's difficulty in interpreting ambiguous neutral face stimuli.

The absence of an emotion effect on the 2-back condition may have been influenced by the difficulty that this task posed for children this age in both language groups. Support for this interpretation comes from the neuroscientific literature; studies using neuroimaging during ER tasks have found that increasing requirements for EF can override emotional effects

(Hart et al., 2010; Mueller et al., 2017). For example, Erk et al. (2006) found no effects of emotional cueing on working memory accuracy on their ER task; however, the fMRI results revealed a valence-specific regulation effect on brain regions whereby participants had significantly reduced activity in brain areas responsible for emotional processing during high cognitive effort conditions than during low cognitive effort conditions, and significantly greater recruitment of regions implicated in working memory as the complexity of the task increased. This research supports our behavioral findings in that the effects of emotional context tend to be reduced under conditions of high cognitive effort as participants attempt to meet the demands of increasing task complexity (e.g., on the 2-back version of the task).

In this context, it might therefore be expected that bilingual children would respond differently than monolingual children in the 1-back version of the task, but this did not happen. Overall, the study failed to capture the anticipated differences in ER between the language groups. Instead, the main finding was that bilingual children were significantly slower than their monolingual peers on the 2-back condition of the Emotional Face N-back Task.

In healthy adult populations, researchers find that positive emotional stimuli are generally processed more quickly and automatically than emotionally neutral stimuli (see Pool et al., 2016, for a meta-analysis), as was found in the present study. Research with typically developing children is sparse and has generated mixed results; however, it is generally accepted that children's ability to modulate attention toward or away from emotionally salient information is a marker of ER that distinguishes healthy from at-risk or atypically developing children (e.g., Shackman et al., 2007; Nuske et al., 2017), and that responses are less consistent than those observed in adults. For example, Mueller et al. (2012) tested healthy versus anxious 12-year-old youth using an antisaccade task with emotional faces and found that healthy youth were more accurate during angry trials and happy trials relative to neutral trials, but revealed no emotion effects in RT. However, typically developing children have also been shown to demonstrate longer RTs on positive emotional conditions than on neutral emotional conditions (Ladouceur et al., 2005). Conversely, studies with anxious children, youth, and adults on ER tasks consistently find a threat bias, also described as biased allocation of attentional resources toward threatening stimuli, which is reflected in longer RTs on trials with aversive stimuli (Ladouceur et al., 2005, 2009; for a review, see Williams et al., 1997). Similar findings have been observed in depressed individuals, whose responses are characterized by impaired disengagement from negative stimuli and deficits in cognitive control while processing negative information (for a review, see Gotlib and Joormann, 2010). Taken together, prolonged engagement with emotionally threatening information is believed to be mediated by deficits in cognitive control. However, our finding that bilingual children exhibited longer RTs, particularly on the 2-back, cannot be explained by this theoretical proposal, because: (1) the slowing on (or lack of quick disengagement from) emotional stimuli for bilinguals was generalized to the whole 2-back condition and was nonspecific to either emotion condition, and (2) bilingual

children demonstrated overall enhanced working memory relative to monolingual children, with no cognitive deficits noted across any conditions on the task. As such, it is reasonable to assume that there is a continuum between typical capture of attention by emotional cues and dysregulated or maladaptive emotional processing.

A possible explanation for the longer RTs for bilinguals and lack of ER differences between the language groups may lie in differences in monitoring and cognitive flexibility (or shifting) abilities between the groups (e.g., Bialystok and Viswanathan, 2009; Prior and MacWhinney, 2010). Shifting or cognitive flexibility is a component of EF, and the ability to think flexibly that includes switching strategies or responses as task demands change (Miyake et al., 2000). Vitiello et al. (2011) have linked enhanced cognitive flexibility to school success in children. It is notable that children in both language groups accommodated the difficulty of the 2-back condition by slowing down, but only the bilinguals maintained high accuracy in the difficult condition. Hur et al. (2017) observed that particularly on a more challenging task as the 2-back, "participants' efforts are generally focused more on performing the task accurately than responding as fast as they can" (p. 4). Many studies show an increase in RT and decrease in accuracy with increasing task difficulty on *n*-back tasks (for a review, see Meule, 2017). Therefore, the slowing for bilingual children cannot be explained by impaired cognitive processing within emotional contexts, but may reflect normal development in healthy bilingual children who are better able than monolingual children to adjust their behavior to task demands.

In summary, this study demonstrated advantages in working memory for bilingual children compared to monolingual children, consistent with previous research showing EF benefits in bilingual individuals, but no evidence for better ER in bilinguals. ER has not been previously investigated with monolingual and bilingual individuals using a working memory task outside of the linguistic context. Although behavioral responses to negative emotional stimuli have commonly been studied within the context of dysregulation and maladjustment, and responses to positive emotions have been studied within the context of healthy socioemotional development, viewing child behavior through this narrow lens may be an oversimplification of functioning and undermine the importance of individual differences that modulate interacting cognitive and emotional processing. Continued research using ER tasks such as the Emotional Face N-back Task has the potential of advancing our understanding of the developmental mechanisms underlying ER in children, and more specifically in elucidating any differences in emotional processing between children with different language experiences.

ETHICS STATEMENT

This research was approved by the Human Participants Review Committee of York University. Parents signed informed consent prior to the experiment and children provided verbal assent before each task.

AUTHOR CONTRIBUTIONS

MJ conducted the study under the supervision of EB in partial fulfillment of requirements for the degree of Doctor of Philosophy.

REFERENCES

- Abutalebi, J., Annoni, J. M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., et al. (2008). Language control and lexical competition in bilinguals: an event-related fMRI study. *Cereb. Cortex* 18, 1496–1505. doi: 10.1093/cercor/bhm182
- Adesope, O., Lavin, T., Thompson, T., and Ungerleider, C. (2010). A systematic review and Meta-analysis of the cognitive correlates of bilingualism. *Rev. Educ. Res.* 80, 207–245. doi: 10.3102/0034654310368803
- American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn. Washington, DC: Author. doi: 10.1176/appi.books.9780890425596
- Anderson, J. A. E., Mak, L., Chahi, A. K., and Bialystok, E. (2018). The language and social background questionnaire: assessing degree of bilingualism in a diverse population. *Behav. Res. Methods* 1–14, 250–263. doi: 10.3758/s13428-017-0867-9 doi: 10.3758/s13428-017-0867-9
- Antoniou, M. (2019). The advantages of bilingualism debate. *Annu. Rev. Linguist.* 5, 1–21. doi: 10.1146/annurev-linguistics011718-011820
- Barac, R., Bialystok, E., Castro, D. C., and Sanchez, M. (2014). The cognitive development of young dual language learners: a critical review. *Early Childhood Res. Q.* 29, 699–714. doi: 10.1016/j.ecresq.2014.02.003
- Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., and van Ijzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychol. Bull.* 133, 1–24. doi: 10.1037/0033-2909.133.1.1
- Beauvillain, C., and Grainger, J. (1987). Accessing interlexical homographs: some limitations of a language-selective access. *J. Mem. Lang.* 26, 658–672. doi: 10.1016/0749-596X(87)90108-2
- Becerril, K., and Barch, D. (2011). Influence of emotional processing on working memory in schizophrenia. *Schizophrenia Bull.* 37, 1027–1038. doi: 10.1093/schbul/sbq009
- Bell, M. A., and Wolfe, C. D. (2004). Emotion and cognition: an intricately bound developmental process. *Child Dev.* 75, 366–370. doi: 10.1111/j.1467-8624.2004.00679.x
- Bialystok, E. (2015). Bilingualism and the development of executive function: the role of attention. *Child Dev. Perspect.* 9, 117–121. doi: 10.1111/cdep.12116
- Bialystok, E. (2017). The bilingual adaptation: how minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099
- Bialystok, E., and Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition* 112, 494–500. doi: 10.1016/j.cognition.2009.06.014
- Bishop, S., Duncan, J., Brett, M., and Lawrence, A. D. (2004). Prefrontal cortical function and anxiety: controlling attention to threat-related stimuli. *Nat. Neurosci.* 7, 184–188. doi: 10.1038/nn1173
- Blom, E., Kuntay, A. C., Messer, M., Verhagen, J., and Leseman, P. (2014). The benefits of being bilingual: working memory in bilingual turkish-dutch children. *J. Exp. Child Psychol.* 128, 105–119. doi: 10.1016/j.jecp.2014.06.007
- Bradley, B. P., Mogg, K., White, J., Groom, C., and de Bono, J. (1999). Attentional bias for emotional faces in generalized anxiety disorder. *Br. J. Clin. Psychol.* 38, 267–278. doi: 10.1348/014466599162845
- Calkins, S. D., and Hill, A. (2007). “Caregiver influences on emerging emotion regulation: Biological and environmental transactions in early development,” in *Handbook of Emotion Regulation*, ed. J. Gross (New York, NY: Guilford Press), 229–248.
- Calkins, S. D., and Marcovitch, S. (2010). “Emotion regulation and executive functioning in early development: Integrated mechanisms of control supporting adaptive functioning,” in *Child Development: At the Intersection of Emotion and Cognition*, eds S. D. Calkins and M. A. Bell (Washington, DC: APA Press), 37–58.
- Cole, P. M., and Deater-Deckard, K. (2009). Emotion regulation, risk, and psychopathology. *J. Child Psychol. Psychiatry* 50, 1327–1330. doi: 10.1111/j.1469-7610.2009.02180.x
- Cole, P. M., Martin, S., and Dennis, T. (2004). Emotion regulation as a scientific construct: methodological challenges and directions for child development research. *Child Dev.* 75, 317–333. doi: 10.1111/j.1467-8624.2004.00673.x
- Colomé, A. (2001). Lexical activation in bilinguals’ speech production: language-specific or language-independent? *J. Mem. Lang.* 45, 721–736. doi: 10.1006/jmla.2001.2793
- Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don’t. *Cognition* 113, 135–149. doi: 10.1016/j.cognition.2009.08.001
- Cromheeke, S., and Mueller, S. C. (2016). The power of a smile: stronger working memory effects for happy faces in adolescents compared to adults. *Cogn. Emot.* 30, 288–301. doi: 10.1080/02699931.2014.997196
- Djambazova-Popordanoska, S. (2016). Implications of emotion regulation on young children’s emotional wellbeing and educational achievement. *Educ. Rev.* 68, 497–515. doi: 10.1080/00131911.2016.1144559
- Dunn, L. M., and Dunn, L. M. (1997). *Peabody Picture Vocabulary Test-III*. Circle Pines, MN: American Guidance Service.
- Eisenberg, N., Spinrad, T. L., and Eggum, N. D. (2010). Emotion-related self-regulation and its relation to children’s maladjustment. *Annu. Rev. Clin. Psychol.* 6, 495–525. doi: 10.1146/annurev.clinpsy.121208.131208
- Engel de Abreu, P. M. (2011). Working memory in multilingual children: is there a bilingual effect? *Memory* 19, 529–537. doi: 10.1080/09658211.2011.590504
- Erk, S., Abler, B., and Walter, H. (2006). Cognitive modulation of emotion anticipation. *Eur. J. Neurosci.* 24, 1227–1236. doi: 10.1111/j.1460-9568.2006.04976.x
- Gotlib, I. H., and Joormann, J. (2010). Cognition and depression: current status and future directions. *Annu. Rev. Clin. Psychol.* 6, 285–312. doi: 10.1146/annurev.clinpsy.121208.131305
- Gray, J. R., Braver, T. S., and Raichle, M. E. (2002). Integration of emotion and cognition in the lateral prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 99, 4115–4120. doi: 10.1073/pnas.062381899
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism* 1, 67–81. doi: 10.1017/S1366728998000133
- Gresham, F. M., Elliott, S. N., Cook, C. R., Vance, M. J., and Kettler, R. (2010). Cross-informant agreement for social skill and problem behavior ratings: an investigation of the Social Skills Improvement System – Rating Scales. *Psychol. Assess.* 22, 157–166. doi: 10.1037/a0018124
- Gross, J. J. (2002). Emotion regulation: affective, cognitive, and social consequences. *Psychophysiology* 39, 281–291. doi: 10.1017/S0048577201393198
- Gumora, G., and Arsenio, W. F. (2002). Emotionality, emotion regulation, and school performance in middle school children. *J. Sch. Psychol.* 40, 395–413. doi: 10.1016/S0022-4405(02)00108-5
- Hare, T. A., Tottenham, N., Davidson, M. C., Glover, G. H., and Casey, B. J. (2005). Contributions of amygdala and striatal activity in emotion regulation. *Biol. Psychiatry* 57, 624–632. doi: 10.1016/j.biopsych.2004.12.038
- Hart, S. J., Green, S. R., Casp, M., and Belger, A. (2010). Emotional priming effects during Stroop task performance. *Neuroimage* 49, 2662–2670. doi: 10.1016/j.neuroimage.2009.10.076
- Hawkins, J., Catalano, R., Kosterman, R., Abbott, R., and Hill, K. (1999). Preventing adolescent health-risk behaviors by strengthening protection during childhood. *Arch. Pediatr. Adolesc. Med.* 153, 226–234. doi: 10.1001/archpedi.153.3.226
- Herba, C. M., Landau, S., Russell, T., Ecker, C., and Phillips, M. L. (2006). The development of emotion processing in children: effects of age, emotion, and

FUNDING

The study was funded by Grant R01HD052523 from the U.S. National Institutes of Health and Grant A2559 from the Natural Sciences and Engineering Research Council of Canada to EB.

- intensity. *J. Child Psychol. Psychiatry* 47, 1098–1106. doi: 10.1111/j.1469-7610.2006.01652.x
- Holley, S. R., Ewing, S. T., Stiver, J. T., and Bloch, L. (2017). The relationship between emotion regulation, executive functioning, and aggressive behaviors. *J. Interpers. Violence* 32, 1692–1707. doi: 10.1177/0886260515592619
- Hur, J., Jordan, A. D., Dolcos, F., and Berenbaum, H. (2017). Emotional influences on perception and working memory. *Cogn. Emot.* 31, 1294–1302. doi: 10.1080/02699931.2016.1213703
- Kensinger, E. A., and Corkin, S. (2003). Effect of negative emotional content on working memory and long-term memory. *Emotion* 3, 378–393. doi: 10.1037/1528-3542.3.4.378
- Kieras, J. E., Tobin, R. M., Graziano, W. G., and Rothbart, M. K. (2005). You can't always get what you want: effortful control and children's responses to undesirable gifts. *Psychol. Sci.* 16, 391–396. doi: 10.1111/j.0956-7976.2005.01546.x
- Kovacs, A. M., and Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6556–6560. doi: 10.1073/pnas.0811323106
- Kroll, J. F., Bobb, S. C., and Hoshino, N. (2014). Two languages in mind: bilingualism as a tool to investigate language, cognition, and the brain. *Curr. Dir. Psychol. Sci.* 23, 159–163. doi: 10.1177/096372141414528511
- Kujawa, A., Dougherty, L., Durbin, C. E., Laptook, R., Torpey, D., and Klein, D. N. (2014). Emotion recognition in preschool children: associations with maternal depression and early parenting. *Dev. Psychopathol.* 26, 159–170. doi: 10.1017/S0954579413000928
- Ladouceur, C. D., Dahl, R. E., Williamson, D. E., Birmaher, B., Ryan, N. D., and Casey, B. J. (2005). Altered emotional processing in pediatric anxiety, depression, and comorbid anxiety-depression. *J. Abnorm. Child Psychol.* 33, 165–177. doi: 10.1007/s10802-005-1825-z
- Ladouceur, C. D., Silk, J. S., Dahl, R. E., Ostapenko, L., Kronhaus, D. M., and Phillips, M. L. (2009). Fearful faces influence attentional control processes in anxious youth and adults. *Emotion* 9, 855–864. doi: 10.1037/a0017747
- Lakes, K. D., Swanson, J. M., and Riggs, M. (2012). The reliability and validity of the English and Spanish Strengths and Weaknesses of ADHD and Normal Behavior Rating Scales in a preschool sample: continuum measures of hyperactivity and inattention. *J. Atten. Disord.* 16, 510–516. doi: 10.1177/1087054711413550
- Lewis, M. D., Granic, I., Lamm, C., Zelazo, P. D., Stieben, J., Todd, R. M., et al. (2008). Changes in the neural bases of emotion regulation associated with clinical improvement in children with behavior problems. *Dev. Psychopathol.* 20, 913–939. doi: 10.1017/S0954579408000448
- Luk, G., Anderson, J. A., Craik, F. I., Grady, C., and Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: response inhibition versus interference suppression. *Brain Cogn.* 74, 347–357. doi: 10.1016/j.bandc.2010.09.004
- Marx, I., Domes, G., Havenstein, C., Berger, C., Sczulle, L., and Herperts, S. C. (2011). Enhanced emotional interference on working memory performance in adults with ADHD. *J. Biol. Psychiatry* 12, 70–75. doi: 10.3109/15622975.2011.599213
- McDermott, J. M., Perez-Edgar, K., Henderson, H. A., Chronis-Tuscano, A., Pine, D. S., and Fox, N. A. (2009). A history of childhood behavioral inhibition and enhanced response monitoring in adolescence are linked to clinical anxiety. *Biol. Psychiatry* 65, 445–448. doi: 10.1016/j.biopsych.2008.10.043
- Meule, A. (2017). Reporting and interpreting working memory performance in n-back tasks. *Front. Psychol.* 8:352. doi: 10.3389/fpsyg.2017.00352
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wagar, T. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734
- Morales, J., Calvo, A., and Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *J. Exp. Child Psychol.* 114, 187–202. doi: 10.1016/j.jecp.2012.09.002
- Mueller, S. C., Cromheeke, S., Siugzdaite, R., and Boehler, N. C. (2017). Evidence for the triadic model of adolescent brain development: cognitive load and task-relevance of emotion differentially affect adolescents and adults. *Dev. Cogn. Neurosci.* 26, 91–100. doi: 10.1016/j.dcn.2017.06.004
- Mueller, S. C., Hardin, M. G., Mogg, K., Benson, V., Bradley, B. P., Reinholdt-Dunne, M. L., et al. (2012). The influence of emotional stimuli on attention orienting and inhibitory control in pediatric anxiety. *J. Child Psychol. Psychiatry* 53, 856–863. doi: 10.1111/j.1469-7610.2012.02541.x
- Nuske, H. J., Hedley, D., Woollacott, A., Thomson, P., Macari, S., and Dissanayake, C. (2017). Developmental delays in emotion regulation strategies in preschoolers with autism. *Autism Res.* 10, 1808–1822. doi: 10.1002/aur.1827
- Owen, A. M., McMillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging. *Hum. Brain Mapp.* 25, 46–59. doi: 10.1002/hbm.20131
- Paap, K. R., and Greenberg, Z. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Paelecke-Habermann, Y., Pohl, J., and Lepow, B. (2005). Attention and executive functions in remitted major depressed patients. *J. Affect. Disord.* 89, 125–135. doi: 10.1016/j.jad.2005.09.006
- Pool, E., Brosch, T., Delplanque, S., and Sander, D. (2016). Attentional bias for positive emotional stimuli: a meta-analytic investigation. *Psychol. Bull.* 142, 79–106. doi: 10.1037/bul0000026
- Poulin-Dubois, D., Blaye, A., Coutya, J., and Bialystok, E. (2011). The effects of bilingualism on toddlers' executive functioning. *J. Exp. Child Psychol.* 108, 567–579. doi: 10.1016/j.jecp.2010.10.009
- Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism* 13, 253–262. doi: 10.1017/S1366728909990526
- Raven, J. C., Court, J. H., and Raven, J. (1996). *Coloured Progressive Matrices*. London: H. K. Lewis.
- Rose, E. J., and Ebmeier, K. P. (2006). Pattern of impaired working memory during major depression. *J. Affect. Disord.* 90, 149–161. doi: 10.1016/j.jad.2005.11.003
- Schmeichel, B. J., and Tang, D. C. (2015). Individual differences in executive functioning and their relationship to emotional processes and responses. *Curr. Dir. Psychol. Sci.* 24, 93–98. doi: 10.1177/0963721414555178
- Shackman, J. E., Shackman, A. J., and Pollak, S. D. (2007). Physical abuse amplifies attention to threat and increases anxiety in children. *Emotion* 7, 838–852. doi: 10.1037/1528-3542.7.4.838
- Shields, A., and Cicchetti, D. (1997). Emotion regulation among school-age children: the development and validation of a new criterion Q-sort scale. *Dev. Psychol.* 33, 906–916. doi: 10.1037/0012-1649.33.6.906
- Shipstead, Z., Harrison, T. L., and Engle, R. W. (2015). Working memory capacity and the scope and control of attention. *Atten. Percept. Psychophys.* 77, 1863–1880. doi: 10.3758/s13414-015-0899-0
- Stieben, J., Lewis, M. D., Granic, I., Zelazo, P. D., and Segalowitz, S. (2007). Neurophysiological mechanisms of emotion regulation for subtypes of externalizing children. *Dev. Psychopathol.* 19, 455–480. doi: 10.1017/S0954579407070228
- Swanson, J., Schuck, S., Mann, M., Carlson, C., Hartman, C., Sergeant, J., et al. (2001). *Over-identification of extreme behavior in the evaluation and diagnosis of ADHD/HKD*. Available at: http://www.adhd.net/SWAN_Paper.pdf
- Thomas, L. A., De Bellis, M. D., Graham, R., and LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Dev. Sci.* 10, 547–558. doi: 10.1111/j.1467-7687.2007.00614.x
- Tottenham, N., Tanaka, J., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res.* 168, 242–249. doi: 10.1016/j.psychres.2008.05.006
- Villemonteix, T., Marx, I., Septier, M., Berger, C., Hacker, T., Bahadori, S., et al. (2017). Attentional control of emotional interference in children with ADHD and typically developing children: an emotional N-back study. *Psychiatry Res.* 254, 1–7. doi: 10.1016/j.psychres.2017.04.027
- Vitiello, V. E., Greenfield, D. B., Munis, P., and George, J. (2011). Cognitive flexibility, approaches to learning, and academic school readiness in Head Start Preschool Children. *Early Educ. Dev.* 3, 388–410. doi: 10.1080/10409289.2011.538366

- Wechsler, D. (1991). *The Wechsler intelligence scale for children*, 3rd Edn. San Antonio, TX: The Psychological Corporation.
- Williams, J. M., Watts, F. N., MacLeod, C., and Mathews, A. (1997). *Cognitive psychology and emotional disorders*, 2nd Edn. Chichester: John Wiley & Sons.
- Wray, E., Shashi, V., Schoch, K., Curtiss, K., and Hooper, S. R. (2013). Discrepancies in parent and teacher ratings of social-behavioral functioning of children with chromosome 22q11.2 deletion syndrome: implications for assessment. *Am. J. Intell. Dev. Disab.* 118, 339–352. doi: 10.1352/1944-7558-118.5.339
- Young, D. J., Levy, F., Neilson, C. M., and Hay, D. A. (2009). Attention Deficit Hyperactivity Disorder: a Rasch analysis of the SWAN Rating Scale. *Child Psychiatry Hum. Dev.* 40, 543–559. doi: 10.1007/s10578-009-0143-z
- Zelazo, P. D., and Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: development and plasticity. *Child Dev. Perspect.* 6, 354–360. doi: 10.1111/j.1750-8606.2012.00246.x
- Zelazo, P. D., and Cunningham, W. A. (2007). “Executive function: Mechanisms underlying emotion regulation,” in *Handbook of Emotion Regulation*, ed. J. J. Gross (New York, NY: Guilford), 135–158.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- The reviewer LF-P and handling Editor declared their shared affiliation at the time of the review.
- Copyright © 2018 Janus and Bialystok. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Do Bilinguals Have an Advantage in Theory of Mind? A Meta-Analysis

Scott R. Schroeder*

Department of Speech-Language-Hearing Sciences, Hofstra University, Hempstead, NY, United States

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Ellen Bialystok,
York University, Canada
Greg Poarch,
Universität Münster, Germany

*Correspondence:

Scott R. Schroeder
scott.r.schroeder@hofstra.edu

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 28 June 2018

Accepted: 27 July 2018

Published: 24 August 2018

Citation:

Schroeder SR (2018) Do Bilinguals
Have an Advantage in Theory of
Mind? A Meta-Analysis.
Front. Commun. 3:36.
doi: 10.3389/fcomm.2018.00036

Bilingualism might help children develop Theory of Mind, but the evidence is mixed. To address the disagreement in the literature, a meta-analysis was conducted on studies that compared bilingual and monolingual children on false belief and other Theory of Mind tests. The meta-analysis of 16 studies and 1,283 children revealed a small bilingual advantage (Cohen's $d = 0.22$, $p = 0.050$). A secondary analysis was conducted on studies ($k = 8$) that statistically adjusted the Theory of Mind scores to correct for a bilingual disadvantage in language proficiency. This secondary analysis indicated a medium-size bilingual advantage (Cohen's $d = 0.58$, $p < 0.001$). There was no evidence for publication bias in either analysis. Taken together, the results provide support for a beneficial effect of acquiring two languages on mental state reasoning. Explanations for this bilingual advantage, which include bilingual-monolingual differences in executive functioning, metalinguistic awareness, and socio-pragmatic abilities, are discussed.

Keywords: bilingualism, Theory of Mind, false belief, executive functioning, cognitive development

INTRODUCTION

Bilingualism research in the modern era has been dominated by a potential bilingual advantage in executive functioning. This advantage, which is supported by many studies (Bialystok, 1999; Bialystok et al., 2004; Costa et al., 2008), has been communicated to the general public through significant media coverage (Bhattacharjee, 2012; Reville, 2014). Yet, several recent studies have failed to replicate this finding (Morton and Harper, 2007; Paap and Greenberg, 2013; Antón et al., 2014), leading many researchers to doubt its validity (de Bruin et al., 2015; Paap et al., 2016), and creating a division in the bilingualism research community between believers and skeptics (Bak, 2016; Bialystok, 2016; Titone et al., 2017). This ambiguous state of the literature is not limited to executive functioning. It extends to other aspects of mental functioning, such as Theory of Mind, a socio-cognitive ability that is thought to be closely linked to executive functioning (Devine and Hughes, 2014).

In the research addressing whether bilingual children have an advantage over their monolingual peers in the development of Theory of Mind, the answer has been mixed (Goetz, 2003; Kovács, 2009; Kyuchukov and De Villiers, 2009; Fan et al., 2015; Gordon, 2016; Dahlgren et al., 2017), often even within a single study (Bialystok and Senman, 2004; Chan, 2004; Nguyen and Astington, 2014; Diaz and Farrar, 2018). To help disambiguate the ambiguous literature, the current study

statistically combined data from many previous studies through a meta-analysis. To provide the background for the meta-analysis, the rest of the Introduction describes the concept of a Theory of Mind and common tests of this ability, followed by reasons for why bilingual children might perform better than their monolingual peers on these tests.

Theory of Mind refers to the ability to attribute mental states to other people and to predict and explain other people's behavior on the basis of those attributed mental states. This ability is often assessed through a false belief test, such as the unexpected-transfer test (Wimmer and Perner, 1983; Baron-Cohen et al., 1985) and the unexpected-contents test (Hogrefe et al., 1986; Perner et al., 1987).

In a popular version of the unexpected-transfer test, known as the Sally-Anne test (Baron-Cohen et al., 1985), participants see a character, who is named Sally, put a marble into a basket. Sally then leaves the scene, and while away, a second character, who is named Anne, removes the marble from the basket and puts it into a box. Next, Sally returns to the scene to retrieve the marble. The key question for the participant is: "Where will Sally look for the marble?" The correct answer is that Sally will look for the marble in the basket, which is where she put it (and not in the box, where it currently is). Answering correctly requires assigning the correct mental state to Sally (namely, the false belief that the marble is in the basket) and predicting her behavior on the basis of that assigned mental state (namely, that she will look in the basket because she falsely believes that the marble is in the basket).

In another commonly-used assessment of Theory of Mind called an unexpected-contents test (Hogrefe et al., 1986; Perner et al., 1987), participants are shown, for example, a tube of Smarties candies and are asked what the tube contains. Participants invariably answer "Smarties" but when the tube is opened, pencils unexpectedly appear (rather than the anticipated Smarties candies). Participants are then asked what someone else, such as a classmate, would predict is contained in the Smarties tube. The correct answer, which is Smarties candies (rather than pencils), requires assigning the correct mental state to someone else (namely, the false belief that the tube contains Smarties candies) and predicting a person's behavior on the basis of that assigned mental state (namely, that the person will say that they think Smarties are contained in the tube).

These false belief tests and other Theory of Mind tests are failed by many children before they turn four years old (Wellman et al., 2001), but there is significant variability among children. For example, many children on the autism spectrum fail to pass these tests even when they are several years older than four (Happé, 1995). Even among typically developing children, there is detectable variability, such as differences across cultures. For example, meta-analyses have revealed faster Theory of Mind development for children from mainland China, Canada, and the United States relative to children from Hong Kong (Liu et al., 2008), and for children from Australia and Canada relative to children from Austria and Japan (Wellman et al., 2001), differences that are thought to be related to certain environmental factors, such as the child's linguistic environment. These cultural differences in the rate of Theory of Mind

development suggest that Theory of Mind is malleable and could potentially be facilitated by a dual-language (i.e., bilingual) environment.

Consistent with this line of thinking, previous studies have provided evidence that bilingualism accelerates Theory of Mind development (Goetz, 2003; Farhadian et al., 2010; Han and Lee, 2013; Diaz and Farrar, 2017). For example, Kovács (2009) found that more than twice as many 2 and 3 year-old Romanian-Hungarian bilingual children passed an unexpected-transfer test than intelligence-matched 2 and 3 year-old Romanian monolingual children.

There are three main accounts for why bilingual children might pass Theory of Mind tests earlier than monolinguals: the "executive functioning" account (Goetz, 2003; Bialystok and Senman, 2004; Kovács, 2009; Greenberg et al., 2013), the "metalinguistic awareness" account (Goetz, 2003; Diaz and Farrar, 2017), and the "socio-pragmatic" account (Goetz, 2003; Fan et al., 2015).

The first account, "executive functioning," is based on evidence that bilingualism improves executive functioning (Carlson and Meltzoff, 2008; Bialystok and Viswanathan, 2009) and that level of executive functioning is a significant predictor of Theory of Mind performance (Devine and Hughes, 2014). The supposed enhanced attentional control abilities of bilinguals could be used to down-regulate their own mental state (i.e., their own beliefs and knowledge) while up-regulating someone else's mental state. The second account, "metalinguistic awareness," is based on evidence that bilingualism enhances metalinguistic awareness (Ben-Zeev, 1977; Bialystok, 1988) and that metalinguistic awareness is linked to Theory of Mind development (Doherty and Perner, 1998; Doherty, 2000). Bilinguals' metalinguistic understanding that there are two labels for the same concept (i.e., one label in each language) might facilitate the understanding that two people can have a different mental state in relation to the same event (and thus that someone else's mental state can differ from their own). The third account, "socio-pragmatic," is that bilinguals come to understand that some people speak only one of their languages (either language A or language B) and some people speak both of their languages (languages A and B). This understanding that two people can have different (or similar) language knowledge may transfer to the more general understanding that two people can have a different (or similar) mental state.

All three of these accounts predicts a bilingual advantage in Theory of Mind development. This prediction has received support both from studies that have used traditional false belief tests, such as the unexpected-location and unexpected-contents tests (Goetz, 2003; Kovács, 2009; Farhadian et al., 2010), as well as studies that have used non-traditional Theory of Mind tests, such as tests that assess the ability to take someone else's visual-spatial perspective when it differs from one's own visual-spatial perspective (Greenberg et al., 2013; Fan et al., 2015). In contrast, several other studies have failed to find a bilingual advantage, both on traditional false belief tests (Kyuchukov and De Villiers, 2009; Pearson, 2013; Nguyen and Astington, 2014; Gordon, 2016; Dahlgren et al., 2017) and non-traditional Theory of Mind tests (Gordon, 2016; Dahlgren et al., 2017).

The inconsistent results across individual studies make it difficult to draw a conclusion about the effects of bilingualism on Theory of Mind development. To help draw a conclusion, the current study statistically combined data from many studies through a meta-analysis. Specifically, a main analysis was conducted, which involved aggregating raw Theory of Mind scores across studies that have compared bilingual and monolingual children on Theory of Mind tests. A secondary analysis was then conducted on the subset of these studies that reported Theory of Mind scores that were statistically adjusted to account for a bilingual disadvantage in language proficiency. It has been argued that bilinguals' lower receptive language proficiency hurts their performance on language-based Theory of Mind tests, thereby concealing a bilingual advantage that would have otherwise emerged (Chan, 2004; Nguyen and Astington, 2014; Diaz and Farrar, 2017, 2018). Thus, the current study presents a main meta-analysis on raw Theory of Mind scores and a secondary meta-analysis on language-adjusted Theory of Mind scores.

METHOD

Literature Search

To identify eligible studies, a three-step process was planned. First, a search through the databases PsycARTICLES, PsycINFO, and MEDLINE was to be conducted using the search terms “bilingual,” “Theory of Mind,” and “false belief.” Second, after identifying eligible articles through the database search, the reference lists of these eligible studies were to be scanned for additional studies that might not have been detected in the database search (i.e., cited studies were to be searched). Third, after eligible articles were identified through both the database search and the reference list search, the studies that cited these eligible studies were to be checked for eligibility (i.e., cited-by studies were to be searched). (Then, in a re-iterative process, the reference lists of the studies identified in the second step and the reference lists and citations of the studies identified in the third step were to be checked.) After completing the search plan in March-May of 2018, a total of 2,032 studies had been considered (though a small subset were duplicates), of which 16 satisfied the inclusion criteria.

Inclusion Criteria

To be eligible for inclusion, a study had to satisfy the following requirements: the study (1) tested bilinguals and monolinguals, (2) tested children rather than adults, (3) tested spoken language users rather than sign language users, and (4) tested participants on a valid Theory of Mind test¹. Included studies also provided

sufficient data to compute an effect size and were reported in a journal ($k = 13$) or a dissertation ($k = 3$).

The 16 studies that were included in the main analysis (i.e., the analysis of raw Theory of Mind scores) are shown in **Table 1**. (Note that the order of the studies in the table was arranged to duplicate the order in **Figure 1**.) Collectively, the 16 studies tested 1,283 participants (655 monolinguals, 628 bilinguals). A subset of these studies ($k = 8$) was included in a secondary analysis (i.e., the analysis of language proficiency adjusted Theory of Mind scores). This secondary analysis used studies that reported Theory of Mind data that were statistically adjusted to account for the confounding variable of bilinguals' reduced language proficiency. These 8 studies, which included 569 participants (311 monolinguals, 258 bilinguals), are marked with an asterisk in **Table 1**.

Most of the studies in the meta-analyses used a version of the unexpected-location or unexpected-transfer false belief test, but some studies used non-traditional Theory of Mind tests (see **Table 1**; Greenberg et al., 2013; Han and Lee, 2013; Fan et al., 2015). Additionally, most of the studies tested English-speaking monolinguals and bilinguals, but some tested non-English speakers (see **Table 1**; Kovács, 2008, 2009; Kyuchukov and De Villiers, 2009; Farhadian et al., 2010; Dahlgren et al., 2017). Furthermore, most of the studies tested children between the ages of 3 and 5, but there were some exceptions (Greenberg et al., 2013; Fan et al., 2015; Dahlgren et al., 2017).

Statistical Analyses

For both the main and secondary analyses, Cohen's d , also known as the Standardized Difference in Means, was used as the effect size measure (Cohen, 1992). The main analysis used raw means and variances to compute Cohen's d , whereas the secondary analysis used statistically adjusted means and variances (typically from an Analysis of Covariance) to compute Cohen's d . When a study used multiple Theory of Mind tests, the effect sizes were pooled together to create a single grand effect size for each study. The Cohen's d effect sizes were entered into a random-effects model. The computing of the effect sizes and the running of the random-effects model were performed in the software Comprehensive Meta-Analysis (Borenstein et al., 2005).

In addition to analyses of effect sizes, potential publication bias was also examined. To this end, a funnel plot with effect sizes and standard errors was visually inspected for symmetry. Following visual inspection, Egger's regression intercept test (Egger et al., 1997) was conducted. The software Comprehensive Meta-Analysis (Borenstein et al., 2005) was used to complete the tests of potential publication bias.

RESULTS

Main Analysis: Raw Scores

The meta-analysis of raw Theory of Mind scores from the 16 studies indicated a small bilingual advantage, Cohen's $d = 0.22$,

included in the meta-analysis. It is important to note, however, that the results of both of these studies were consistent with the results of the meta-analysis (i.e., a bilingual advantage).

¹Two studies (Berguno and Bowler, 2004; Yow and Markman, 2015) did not use a measure of Theory of Mind that was deemed valid for the current purpose and were thus not included in the meta-analysis. Berguno and Bowler (2004) did not use any Theory of Mind tests that assessed the attribution of mental states to others (only to oneself). Because this study did not assess the key component of Theory of Mind, it was not included in the meta-analysis. Yow and Markman (2015) used a word learning test that included a Theory of Mind component. Because there is evidence for a bilingual advantage in word learning (Kaushanskaya et al., 2014), better performance on the word learning test might not reflect a bilingual advantage in Theory of Mind *per se*. Due to this confound, this study was not

TABLE 1 | Studies included in main analysis.

	Monolingual N	Bilingual N	Monolingual languages	Bilingual languages	Theory of Mind Tests
Diaz and Farrar, 2018*	33	32	English	English & Spanish	Unexpected-Transfer, Unexpected-Contents, Other
Kyuchukov and De Villiers, 2009 ^a	60	60	Bulgarian	Bulgarian & Romani	Unexpected-Transfer, Unexpected-Contents
Pearson, 2013	40	28	English	English & Spanish	Unexpected-Transfer
Bialystok and Senman, 2004* ^b	52	43	English	English & Other	Other
Gordon, 2016* ^c	26	26	English	English & Spanish	Unexpected-Transfer, Unexpected-Contents, Other
Dahlgren et al., 2017	14	14	Swedish or Slavic	Swedish & Slavic	Unexpected-Transfer, Other
Nguyen and Astington, 2014*	48	24	English or French	English & French	Unexpected-Transfer, Unexpected-Contents
Diaz and Farrar, 2017* ^d	35	38	English	English & Spanish	Unexpected-Transfer, Unexpected-Contents, Other
Kovács, 2009 ^e	28	28	Italian	Romanian & Slovenian	Unexpected-Transfer, Unexpected-Contents
Chan, 2004*	29	31	English	English & Chinese	Unexpected-Transfer, Unexpected-Contents, Other
Goetz, 2003*	64	40	English or Chinese	English & Chinese	Unexpected-Transfer, Unexpected-Contents, Other
Han and Lee, 2013	60	73	Korean	English & Korean	Other
Greenberg et al., 2013	45	37	English	English & Other	Other
Kovács, 2009	32	32	Romanian	Romanian & Hungarian	Unexpected-Transfer, Other
Farhadian et al., 2010	65	98	Persian	Persian & Kurdish	Unexpected-Transfer, Unexpected-Contents
Fan et al., 2015*	24	24	English	English & Other	Other

*The study was included in the secondary analysis

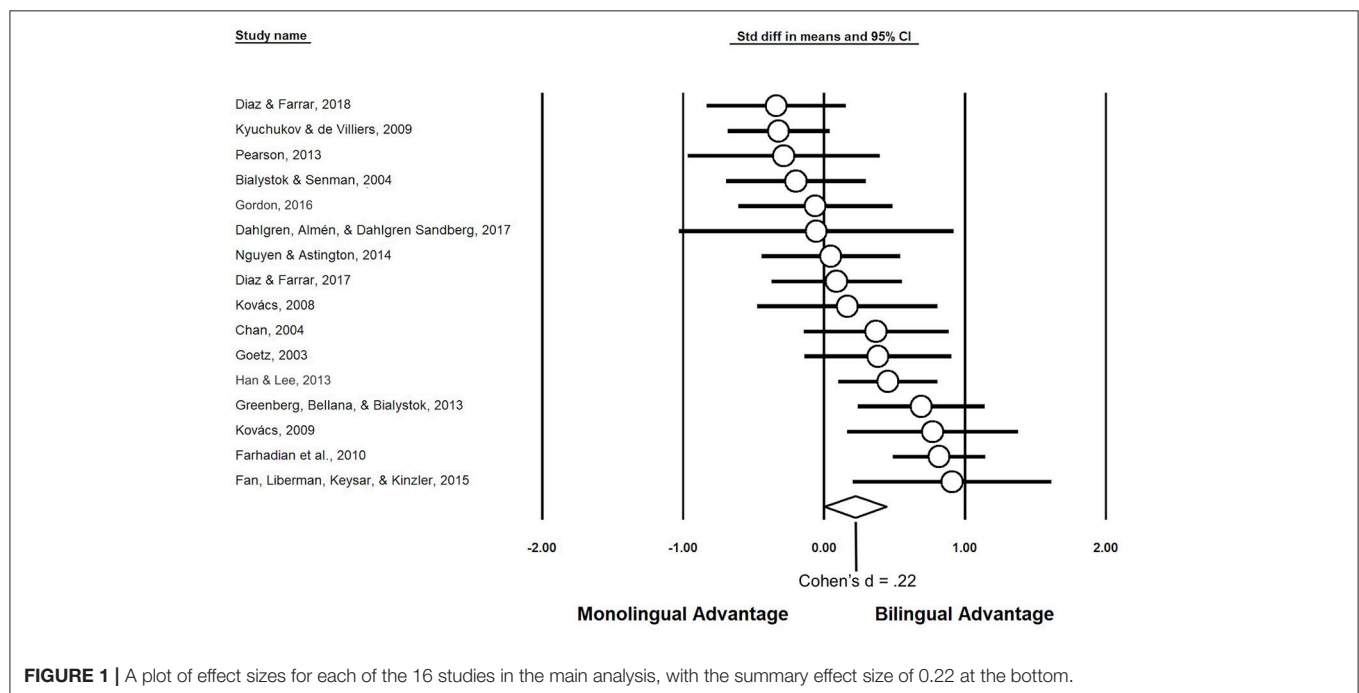
^aBilinguals were tested in both their L1 and L2. The L1 test scores were used to compute the effect size so that Theory of Mind scores for both monolinguals and bilinguals were based on their native language performance.

^bIn the secondary analysis, only the appearance results of the appearance-reality test were used to compute an effect size because the reality questions did not require a Theory of Mind.

^cGordon's dissertation Millett, 2010 was used to extract additional data that were not included in the journal version.

^dA subset of participants was tested a second time but for a more accurate statistical calculation only time 1 data were included in the effect size calculation.

^eThis dissertation contained an experiment that was published in Kovács (2009). So as to not give this experiment double weight, it was not included when computing the effect size for Kovács (2008).

**FIGURE 1** | A plot of effect sizes for each of the 16 studies in the main analysis, with the summary effect size of 0.22 at the bottom.

$p = 0.050$, $z = 1.96$, $SE = 0.11$, 95% Confidence Interval = 0.00–0.44. A plot of the effect sizes for each of the 16 studies and the summary effect size (Cohen's $d = 0.22$) is displayed in **Figure 1**.

To assess the possibility of publication bias, a funnel plot was generated. See **Figure 2** for the plot. There is no apparent asymmetry in the plot, suggestive of no publication bias.

Confirming the lack of publication bias, the Eggers regression intercept test was not significant, $t_{(14)} = 0.65, p = 0.53$.

Secondary Analysis: Language Proficiency Adjusted Scores

The secondary meta-analysis was conducted on the 8 studies that reported Theory of Mind scores that were adjusted for bilingual-monolingual differences in language proficiency. This analysis indicated a medium-size bilingual advantage, Cohen's $d = 0.58, p < 0.001, z = 6.70, SE = 0.09, 95\%$ Confidence Interval = 0.41–0.75. See **Figure 3** for a plot of the effect sizes.

A funnel plot, which is displayed in **Figure 4**, was created to check for potential publication bias. There is no obvious asymmetry in the plot, implying no publication bias. Eggers regression intercept test also indicated no publication bias, as the test was not significant, $t_{(6)} = 0.17, p = 0.87$.

DISCUSSION

The main meta-analysis, which compared bilingual and monolingual children's raw Theory of Mind scores, revealed a small bilingual advantage. The size of this bilingual-monolingual

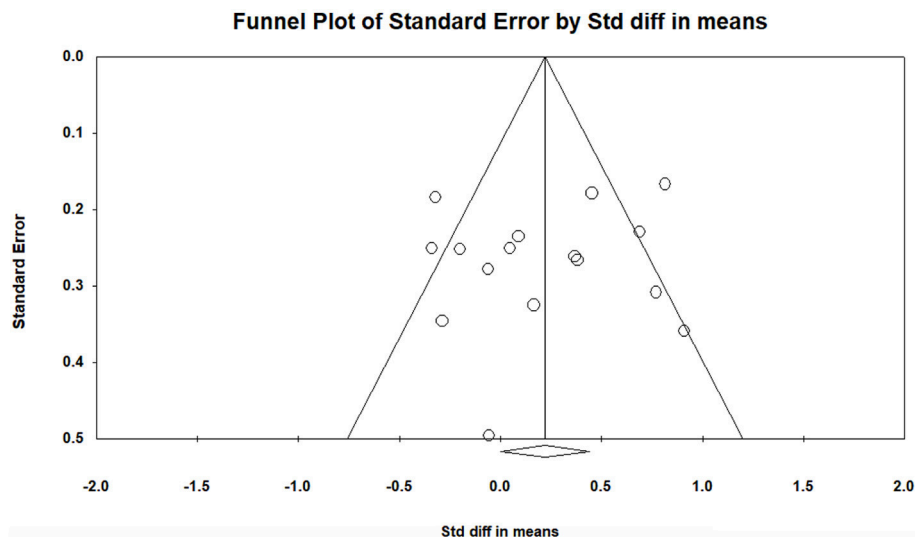


FIGURE 2 | A funnel plot to assess potential publication bias in the main analysis.

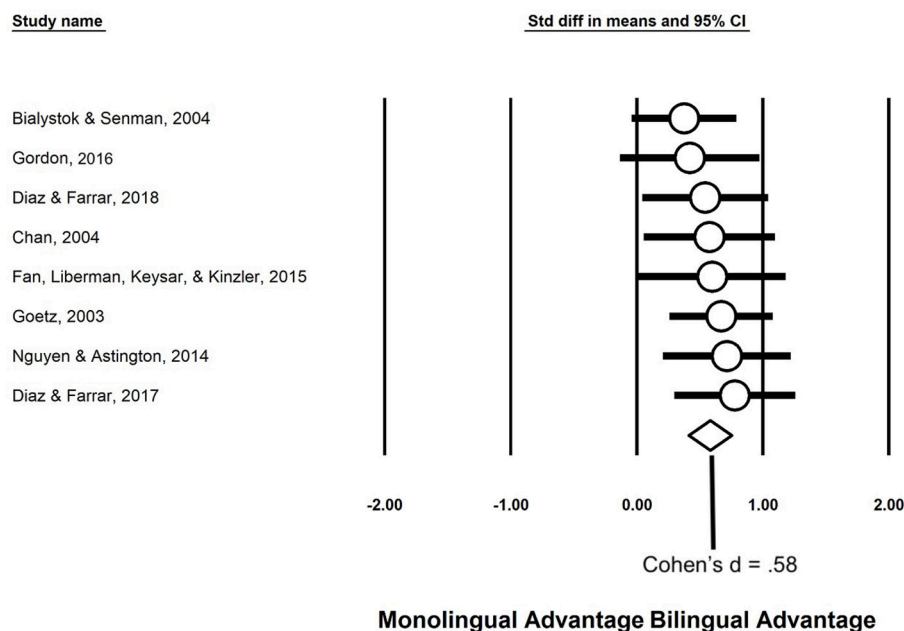


FIGURE 3 | A plot of effect sizes for each of the 8 studies in the secondary analysis, with the summary effect size of .58 at the bottom.

difference (i.e., a Cohen's d in the “small” range) is similar to the effect of early education interventions on cognitive, school, and social outcomes (Camilli et al., 2010). The secondary meta-analysis, which used transformed Theory of Mind scores that were adjusted for language proficiency, revealed a medium-size bilingual advantage. This secondary analysis, however, should be interpreted with caution, given that these studies may have violated assumptions of the Analysis of Covariance (Miller and Chapman, 2001; Paap et al., 2015). Even with skepticism for the secondary analysis, the main analysis provides evidence that acquiring two languages helps Theory of Mind development.

This meta-analytical finding of a bilingual advantage would have less validity if evidence for publication bias had been found. Indeed, in the high-profile meta-analysis by de Bruin et al. (2015), a bilingual advantage in executive functioning was revealed, but so was a publication bias. Using the same method as de Bruin et al. (i.e., the Eggers test), there was no evidence for publication bias in either the main analysis or the secondary analysis.

While the current study indicates a bilingual advantage in Theory of Mind, it does not address the reasons why. In the Introduction, three accounts for why bilinguals might have an advantage in mental state reasoning were laid out—i.e., the “executive functioning” account, the “metalinguistic awareness” account, and the “socio-pragmatic” account. Though future research is needed to determine the relative contributions of these accounts and others, some of the studies included in this meta-analysis provide germane preliminary evidence. Regarding the “executive functioning” account, evidence for this account comes from the Kovács (2008) finding that a bilingual advantage emerges when the Theory of Mind test has high inhibitory demands but not when it has low inhibitory demands. However, evidence against this account comes from several other studies that have found that measures of executive

functioning (such as the dimensional change card sorting test) do not statistically mediate the bilingual advantage in Theory of Mind (Nguyen and Astington, 2014; Fan et al., 2015; Diaz and Farrar, 2017, 2018). Regarding the “metalinguistic awareness” account, Diaz and Farrar (2017) and Chan (2004) found that measures of metalinguistic awareness (such as symbol substitution, synonym judgment, and homonym selection) statistically mediate the bilingual advantage. Regarding the “socio-pragmatic” account, while there is no statistical mediation evidence, Fan et al. (2015) found a Theory of Mind advantage in children who were not bilingual but were exposed to a second language. The performance by these children suggests that Theory of Mind may be augmented by learning that one's linguistic knowledge can be different from that of other people.

Regardless of the source, this bilingual advantage is likely to have meaningful real-world consequences. On the one hand, an enhanced Theory of Mind may help in the development of prosocial behavior. For example, a recent meta-analysis of 20 studies revealed that children who scored higher on Theory of Mind tests were more popular among their peers (Slaughter et al., 2015). On the other hand, negative effects of an enhanced Theory of Mind are possible. For instance, a recent study found that Theory of Mind training led honest children to begin lying (Ding et al., 2015).

In sum, the current study took a meta-analytical approach to the question of whether learning two languages has a positive impact on mental state reasoning. The results indicated a small- or medium-size positive effect (depending on the analysis), an effect that may carry real-world implications for bilingual children's social competence. Though plausible accounts of this bilingual advantage have been put forward, future research is needed to determine more precisely why a dual-language environment is helpful for Theory of Mind development.

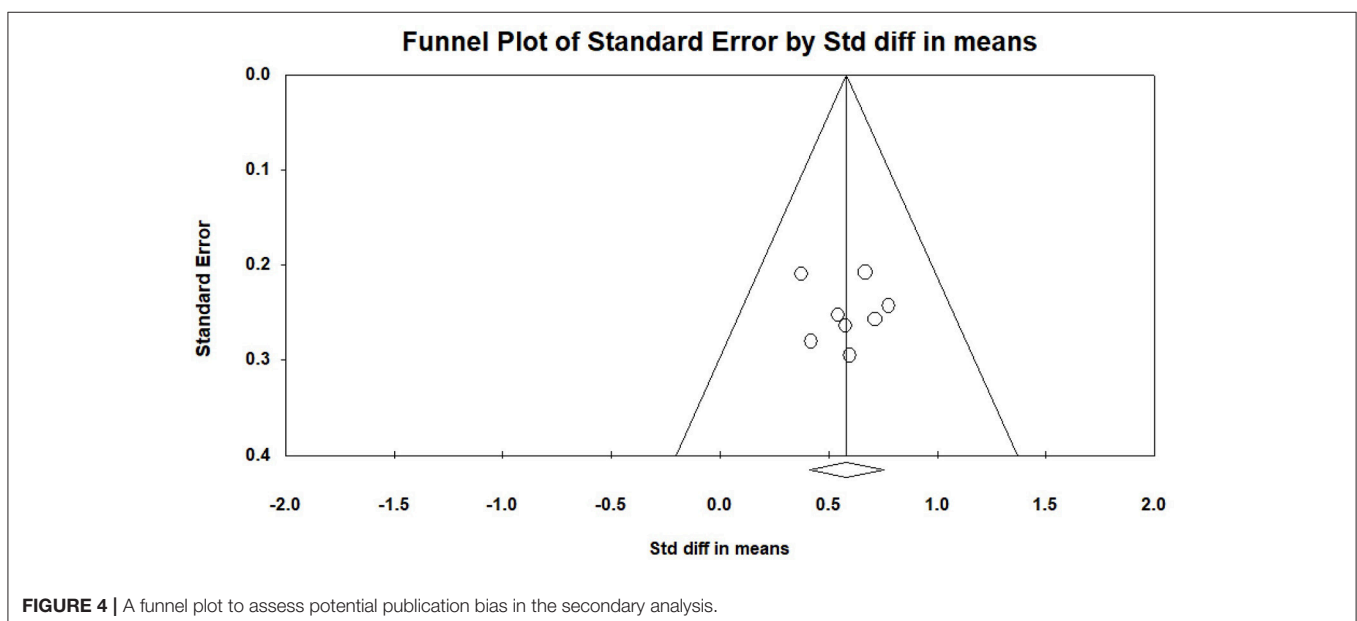


FIGURE 4 | A funnel plot to assess potential publication bias in the secondary analysis.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

REFERENCES

- Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., et al. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Front. Psychol.* 5:398. doi: 10.3389/fpsyg.2014.00398
- Bak, T. H. (2016). Cooking pasta in La Paz. *Linguisti. Approach. Bilingualism*. 6, 699–717. doi: 10.1075/lab.16002.bak
- Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). Does the autistic child have a “theory of mind?” *Cognition* 21, 37–46.
- Ben-Zeev, S. (1977). The influence of bilingualism on cognitive strategy and cognitive development. *Child Dev.* 48, 1009–1018. doi: 10.2307/1128353
- Berguno, G., and Bowler, D. M. (2004). Communicative interactions, knowledge of a second language, and theory of mind in young children. *J. Genet. Psychol.* 165, 293–309. doi: 10.3200/GNTP.165.3.293-309
- Bhattacharjee, Y. (2012). The benefits of bilingualism. Available online at: <https://www.nytimes.com/2012/03/18/opinion/sunday/the-benefits-of-bilingualism.html>
- Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Dev. Psychol.* 24, 560–567. doi: 10.1037/0012-1649.24.4.560
- Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Dev.* 70, 636–644. doi: 10.1111/1467-8624.00046
- Bialystok, E. (2016). The signal and the noise. *Linguisti. Approaches Bilingualism* 6, 517–534. doi: 10.1075/lab.15040.bia
- Bialystok, E., Craik, F. I., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19:290. doi: 10.1037/0882-7974.19.2.290
- Bialystok, E., and Senman, L. (2004). Executive processes in appearance-reality tasks: the role of inhibition of attention and symbolic representation. *Child Dev.* 75, 562–579. doi: 10.1111/j.1467-8624.2004.00693.x
- Bialystok, E., and Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition* 112, 494–500. doi: 10.1016/j.cognition.2009.06.014
- Borenstein, M., Hedges, L., Higgins, J., and Rothstein, H. (2005). *Comprehensive Meta-Analysis Version 2*. Englewood, NJ: Biostat.
- Camilli, G., Vargas, S., Ryan, S., and Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teach. Coll. Rec.* 112, 579–620.
- Carlson, S. M., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298. doi: 10.1111/j.1467-7687.2008.00675.x
- Chan, K. T. (2004). *Chinese-English Bilinguals' Theory-of-Mind Development*, Unpublished Doctoral Dissertation, University of Toronto, Toronto, CA.
- Cohen, J. (1992). A power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037/0033-2909.112.1.155
- Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86. doi: 10.1016/j.cognition.2006.12.013
- Dahlgren, S., Almén, H., and Dahlgren Sandberg, A. (2017). Theory of mind and executive functions in young bilingual children. *J. Genet. Psychol.* 178, 303–307. doi: 10.1080/00221325.2017.1361376
- de Bruin, A., Treccani, B., and Della Sala, S. (2015). Cognitive advantage in bilingualism: an example of publication bias? *Psychol. Sci.* 26, 99–107. doi: 10.1177/0956797614557866
- Devine, R. T., and Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: a meta-analysis. *Child Dev.* 85, 1777–1794. doi: 10.1111/cdev.12237
- Diaz, V., and Farrar, M. J. (2017). The missing explanation of the false-belief advantage in bilingual children: a longitudinal study. *Dev. Sci.* 21:e12594. doi: 10.1111/desc.12594
- Diaz, V., and Farrar, M. J. (2018). Do bilingual and monolingual preschoolers acquire false belief understanding similarly? The role of executive functioning and language. *First Lang.* 38, 382–398. doi: 10.1177/0142723717752741
- Ding, X. P., Wellman, H. M., Wang, Y., Fu, G., and Lee, K. (2015). Theory-of-mind training causes honest young children to lie. *Psychol. Sci.* 26, 1812–1821. doi: 10.1177/0956797615604628
- Doherty, M., and Perner, J. (1998). Metalinguistic awareness and theory of mind: just two words for the same thing? *Cogn. Dev.* 13, 279–305.
- Doherty, M. J. (2000). Children's understanding of homonymy: metalinguistic awareness and false belief. *J. Child Lang.* 27, 367–392. doi: 10.1017/S0305000900004153
- Egger, M., Davey Smith, G., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Br. Med. J.* 315, 629–634.
- Fan, S. P., Liberman, Z., Keysar, B., and Kinzler, K. D. (2015). The exposure advantage: early exposure to a multilingual environment promotes effective communication. *Psychol. Sci.* 26, 1090–1097. doi: 10.1177/0956797615574699
- Farhadian, M., Abdullah, R., Mansor, M., Redzuan, M., Gazanizadand, N., and Kumar, V. (2010). Theory of mind in bilingual and monolingual preschool children. *J. Psychol.* 1, 39–46. doi: 10.1080/09764224.2010.11885444
- Goetz, P. J. (2003). The effects of bilingualism on theory of mind development. *Bilingualism Lang. Cogn.* 6, 1–15. doi: 10.1017/S1366728903001007
- Gordon, K. R. (2016). High proficiency across two languages is related to better mental state reasoning for bilingual children. *J. Child Lang.* 43, 407–424. doi: 10.1017/S0305000915000276
- Greenberg, A., Bellana, B., and Bialystok, E. (2013). Perspective-taking ability in bilingual children: extending advantages in executive control to spatial reasoning. *Cogn. Dev.* 28, 41–50. doi: 10.1016/j.cogdev.2012.10.002
- Han, S., and Lee, K. (2013). Cognitive and affective perspective-taking ability of young bilinguals in South Korea. *Child Stud. Div. Contexts* 3, 69–80. doi: 10.5723/cscd.2013.3.1.069
- Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Dev.* 66, 843–855. doi: 10.2307/1131954
- Hogrefe, G., Wimmer, H., and Perner, J. (1986). Ignorance versus false belief: a developmental lag in attribution of epistemic states. *Child Dev.* 57, 567–582. doi: 10.2307/1130337
- Kaushanskaya, M., Gross, M., and Buac, M. (2014). Effects of classroom bilingualism on task-shifting, verbal memory, and word learning in children. *Dev. Sci.* 17, 564–583. doi: 10.1111/desc.12142
- Kovács, Á. M. (2008). *Learning Two Languages Simultaneously: Mechanisms Recruited for Dealing With a Mixed Linguistic Input*. Unpublished doctoral dissertation, SISSA, Trieste.
- Kovács, Á. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Dev. Sci.* 12, 48–54. doi: 10.1111/j.1467-7687.2008.00742.x
- Kyuchukov, H., and De Villiers, J. (2009). Theory of mind and evidentiality in romanian-bulgarian bilingual children. *Psychol. Lang. Commun.* 13, 21–34. doi: 10.2478/v10057-009-0007-4
- Liu, D., Wellman, H. M., Tardif, T., and Sabbagh, M. A. (2008). Theory of mind development in Chinese children: a meta-analysis of false-belief understanding across cultures and languages. *Dev. Psychol.* 44:523. doi: 10.1037/0012-1649.44.2.523
- Miller, G. A., and Chapman, J. P. (2001). Misunderstanding analysis of covariance. *J. Abnorm. Psychol.* 110, 40–48. doi: 10.1037/0021-843X.110.1.40
- Millett, K. R. G. (2010). *The Cognitive Effects of Bilingualism: Does Knowing Two Languages Impact Children's Ability to Reason about Mental States?*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, MN.

ACKNOWLEDGMENTS

Thank you to Viorica Marian and the Northwestern *Bilingualism and Psycholinguistics Research Group* for helpful discussions of this topic.

- Morton, J. B., and Harper, S. N. (2007). What did Simon say? revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726. doi: 10.1111/j.1467-7687.2007.00623.x
- Nguyen, T.-K., and Astington, J. W. (2014). Reassessing the bilingual advantage in theory of mind and its cognitive underpinnings. *Bilingualism Lang. Cogn.* 17, 396–409. doi: 10.1017/S1366728913000394
- Paap, K. R., and Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Paap, K. R., Johnson, H. A., and Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69, 265–278. doi: 10.1016/j.cortex.2015.04.014
- Paap, K. R., Johnson, H. A., and Sawi, O. (2016). Should the search for bilingual advantages in executive functioning continue? *Cortex* 74, 305–314. doi: 10.1016/j.cortex.2015.09.010
- Pearson, D. K. (2013). *Effect of Language Background on Metalinguistic Awareness and Theory of Mind*. Unpublished doctoral dissertation, University of Stirling, Scotland.
- Perner, J., Leekam, S. R., and Wimmer, H. (1987). Three-year-olds' difficulty with false belief: the case for a conceptual deficit. *Br. J. Dev. Psychol.* 5, 125–137.
- Reville, W. (2014). *The Bilingual Brain Is More Nimble And Efficient*. Available online at: <https://www.irishtimes.com/news/science/the-bilingual-brain-is-more-nimble-and-efficient-1.1764064>
- Slaughter, V., Imuta, K., Peterson, C. C., and Henry, J. D. (2015). Meta-analysis of theory of mind and peer popularity in the preschool and early school years. *Child Dev.* 86, 1159–1174. doi: 10.1111/cdev.12372
- Titone, D., Gullifer, J., Subramaniapillai, S., Rajah, N., and Baum, S. (2017). “History-inspired reflections on the bilingual advantages hypothesis,” in *Growing Old with Two Languages: Effects of Bilingualism on Cognitive Aging*, eds E. Bialystok and M. D. Sullivan (Amsterdam: John Benjamins Publishing), 265–295.
- Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72, 655–684. doi: 10.1111/1467-8624.00304
- Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128.
- Yow, W. Q., and Markman, E. M. (2015). A bilingual advantage in how children integrate multiple cues to understand a speaker's referential intent. *Bilingualism Lang. Cogn.* 18, 391–399. doi: 10.1017/S1366728914000133

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Schroeder. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Bilingual Advantages in Inhibition or Selective Attention: More Challenges

Kenneth R. Paap*, Regina Anders-Jefferson, Lauren Mason, Katerinne Alvarado and Brandon Zimiga

Department of Psychology, San Francisco State University, San Francisco, CA, United States

A large sample ($N = 141$) of college students participated in both a conjunctive visual search task and an ambiguous figures task that have been used as tests of selective attention. Tests for effects of bilingualism on attentional control were conducted by both partitioning the participants into bilinguals and monolinguals and by treating bilingualism as a continuous variable, but there were no effects of bilingualism in any of the tests. Bayes factor analyses confirmed that the evidence substantially favored the null hypothesis. These new findings mesh with failures to replicate language-group differences in congruency-sequence effects, inhibition-of-return, and working memory capacity. The evidence that bilinguals are better than monolinguals at attentional control is equivocal at best.

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Claudia C. von Bastian,
University of Sheffield,
United Kingdom
Manuel Perea,
Universitat de València, Spain

*Correspondence:

Kenneth R. Paap
kenp@sfsu.edu

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 21 June 2018

Accepted: 19 July 2018

Published: 15 August 2018

Citation:

Paap KR, Anders-Jefferson R,
Mason L, Alvarado K and Zimiga B
(2018) Bilingual Advantages
in Inhibition or Selective Attention:
More Challenges.
Front. Psychol. 9:1409.
doi: 10.3389/fpsyg.2018.01409

Keywords: bilingualism, inhibitory control, selective attention, visual search, ambiguous figures

INTRODUCTION

Fluent bilinguals have acquired two lexicons and two grammars and must be able to select the intended words and rules as they switch back and forth between their two languages. This is usually viewed as non-trivial because both languages are coactivated during production and comprehension (see Paap, 2019 for a review). For example, the intention to say “gato” may coactivate “cat” in a Spanish-English bilingual. A common assumption (e.g., Blumenfeld and Marian, 2014) is that the competition from “cat” is usually resolved early by inhibiting the CAT representation within the lexicon. Furthermore, having nipped CAT in the bud the articulatory features for producing “cat” may not always emerge as a competitor that requires response inhibition. If the inhibitory control exercised at either the lexical or articulatory-response levels involves a general (domain-free) inhibitory-control mechanism and if this recruitment of general inhibitory control is functionally greater than the levels sustained by monolinguals in speaking a single language and in pursuing the myriad of goals required by everyday life, then bilingual advantages in inhibitory control would result. As we have repeatedly speculated (Paap and Greenberg, 2013; Paap et al., 2015; Paap, 2018, 2019) this logical chain can be broken at any link and consequently it should not be a surprise that the evidence for a bilingual advantage in inhibitory control is weak, at best.

The main purpose of this article is to consider Bialystok et al. (2009) revised hypothesis that bilingual advantages occur in attentional systems rather than in general inhibitory control.

The roots for this shift can readily be traced to the 2009 review by Bialystok et al. (2004) that opens the debate with a section headed *Inhibition or selection?* The authors point out that most (but not all) of the evidence taken to support the assumption that bilingual language control recruits a healthy dose of inhibitory control merely supports the less specific assumption that

there is ubiquitous competition between the two languages that must be resolved by some conflict resolution mechanism: inhibition, selection, or some combination of mechanisms. Bialystok (2017) is convinced that “*Joint activation requires that there is a mechanism for language selection to assure that use of the target language proceeds fluently.*” Furthermore, “*the assumptions are that this mechanism is part of a domain-general process and that the constant engagement of this process for language selection fortifies it for other purposes, including non-verbal ones.* . . .” p. 234. A key outcome of this analysis is that in the absence of additional qualifying assumptions regarding the nature of specific tasks, both the original hypothesis based on inhibition and the revised hypothesis based on selection make the same prediction (viz., a bilingual advantage) for non-verbal interference tasks such as the Simon, spatial Stroop, or flanker.

We will return to this issue in the discussion, but suffice to say that because the revised hypothesis, like the original, assumes that joint activation and competition involves some form of general-purpose control that is strengthened through practice (bilingual experience), the vast literature comparing bilinguals to monolinguals on tests requiring conflict resolution remain relevant to the strength and scope of bilingual advantages in cognitive control. That literature will be reviewed next followed by a review of studies using tasks that more distinctively focus on attentional control.

LANGUAGE GROUP DIFFERENCES IN INTERFERENCE CONTROL

Average Effect Sizes

An early meta-analysis appeared to provide compelling evidence ($g = 0.40$) for bilingual advantages in cognition (Adescope et al., 2010). However, the analysis was very broad in scope and with hindsight very likely influenced by the file-drawer problem and publication bias. Direct evidence for these biases was provided by de Bruin et al. (2014). Both Hilchey et al. (2015) and Sanchez-Azanza et al. (2017), who used a bibliometric approach, speculated that the 2013 article by Paap and Greenberg may have been a turning point whereby challenges to the bilingual advantage hypothesis were more common than not, in part, because of a decrease in bias in the published literature. No doubt the steady drum beat of null results in large-scale studies (with highly proficient and balanced bilinguals and ages ranging from six to older adults) published by the Basque Center on Cognition, Brain, and Language (BCBL) also contributed to this shift (Antón et al., 2014, 2016; Duñabeitia et al., 2014).

More recent meta-analyses converge on the conclusion that significant bilingual advantages in inhibitory control are relatively rare (15% of all comparisons in Paap, 2018), that the average effect sizes are very small, and that there remains some amount of publication bias, which when taken into account, completely eliminates the effect. In Paap (2018) the mean advantage across 146 comparisons using interference scores derived from non-verbal interference tasks was +4.4 ms. If the 146 effect sizes are treated as a single sample the Bayes

Factor (using the JZS prior and Rouder’s calculator) favoring the alternative is only 2.9.

A meta-analysis by Lehtonen et al. (2018) examined bilingual advantages across six domains of executive functioning (with very similar outcomes), but their analysis of inhibitory control is central to this discussion. Their meta-analysis used a wider definition of inhibitory control tasks and identified a more heterogeneous set of 212 effect sizes compared to Paap (2018). The Lehtonen et al. (2018) analysis was restricted to comparisons that were independent, yielded standardized effect sizes, and based on participants 18 years and older. In contrast, the Paap meta-analysis included participants 6 years and older. The mean effect size for inhibition in Lehtonen et al. (2018) was Hedge’s $g = +0.11$ [$+0.05$, $+0.18$], but when corrected for bias the mean was no longer significant, $g = -0.02$ [-0.12 , $+0.08$]. The differences between the two meta-analyses are complementary and the fact that they converge on the same outcome leads to the conclusion that the evidence for a bilingual advantage in inhibitory control is extremely weak.

Advantages in the Elderly?

The topic editors have called for a greater focus on possible developmental effects. In that regard Bialystok (2017) often characterizes the research on inhibitory control as showing more consistent bilingual advantages in older adults compared to younger adults. When the two “extraordinary” outliers¹ of older adults from the Bialystok et al. (2004) study are excluded there are 19 other comparisons using non-verbal interference tasks in the Paap (2018) database of non-verbal interference tasks. Only two show significant bilingual advantages and the mean advantage of +9.7 ms has a 95% CI that straddles zero (CI: -0.4 , $+19.9$). Seniors do not show consistent bilingual advantages.

Advantages in School Children?

Similarly, there is a lore that bilingual advantages in inhibitory control occur consistently in children. In order to test this view the Paap (2018) database was searched for studies using children in the range of 6–15 years old. Only 3 of 30 comparisons produced significant bilingual advantages and the mean bilingual advantage was +2.2 ms (95% CI: -7.9 , $+12.2$). School children do not show consistent bilingual advantages in these non-verbal interference tasks.

Task Differences?

Another challenge to testing for bilingual advantages is that the interference scores derived from different non-verbal interference tasks often show weak and non-significant inter-task correlations (Paap and Sawi, 2014) and low test-retest reliability (Paap and Sawi, 2016). Even more disconcerting the arrows version of the flanker task does not correlate with original letter version (Salthouse, 2010). A more promising outcome was recently reported by Paap et al. (unpublished) in a study comparing four closely matched versions of the Simon, horizontal spatial-Stroop, vertical spatial-Stroop, and flanker

¹See Paap (2018) for the criterion used to identify outliers and a discussion of how strikingly anomalous these Simon effects were for older monolinguals.

tasks in that the interference scores from the first three showed moderate inter-task correlations (r 's ≈ 0.4). The flanker task did not significantly correlate with the Simon or spatial Stroop tasks suggesting that the nature of conflict resolution may depend on whether the conflict arises from two dimensions of the same stimulus or between adjacent but separate stimuli. The latter characterizes the flanker task because participants must select the relevant central arrow among the irrelevant flankers using visuospatial attention. Many theorists have suggested that conflict in the flanker task is resolved by spatially attending to the target stimulus (e.g., Magen and Cohen's, 2007, dimension-action model). If spatial attention is construed as a filter or the upregulation of task relevant information then it clearly contrasts with inhibition. This interpretation of the flanker task is timely with respect to the bilingual advantage controversy given Bialystok's (2017) reframing of the hypothesis from inhibition to attentional control. However, it must be noted that there were no bilingual advantages in inhibitory control in any of the four tasks reported by Paap et al. (unpublished).

LANGUAGE GROUP DIFFERENCES IN OTHER TASKS REQUIRING ATTENTIONAL CONTROL

Bialystok's revised hypothesis assumes that when lexical entries in the two lexicons are co-activated that it is the disengagement of attention from the non-target language, not inhibition, that is the mechanism responsible for facilitating selection of the target language and the mechanism that creates bilingual advantages in domain general cognitive control. The evidence for this revised hypothesis has been drawn from the five tasks discussed in this section. The first two, conjunctive visual search and the ambiguous figures task are quite new to the bilingual advantage debate and will be the focus of the new studies reported below.

Conjunctive Visual Search

In a test of the bilingual advantage in selective-attention hypothesis Friesen et al. (2014) reported that bilingual adults outperformed their monolingual counterparts in a conjunctive visual search task. Participants were instructed to decide as quickly and accurately as possible whether a specified target was present among an array of distractors. Displays including a target were designated as target-present trials whereas displays consisting only of distractors were designated as target-absent trials. Task difficulty was manipulated by search type: (target present vs. target absent), discriminability (low vs. high), and distractor set size (5, 15, 25). Latency was the primary dependent measure with faster RTs indicating better performance.

As expected significant bilingual advantages in search time occurred only in the conjunctive search condition that had low discriminability stimuli. For unexplained reasons, results only for the target-present trials were reported and analyzed. The significant group differences led Friesen et al. (2014) to conclude that bilingualism improves top-down selective attention in young adults. More specifically the extensive

practice bilinguals receive at disengaging attention from the non-target language produces far transfer in the form of an enhanced ability to disengage attention from the distractors and more quickly find the target in the conjunctive visual search condition.

Given the difficulties in replicating studies showing bilingual advantages in EF it is perhaps no surprise that Ratiu et al. (2017) failed to find any bilingual advantages in conjunctive search across a series of three experiments that use eye movements to separate search time from decision time during conjunctive visual search. The study is data rich and if bilingualism confers advantages in selective attention a consistent difference should have been observed across the three experiments. The only reliable group difference was observed in Experiment 3 and that was a bilingual disadvantage in decision times. Ratiu et al. (2017) conclude that their results show no bilingual advantages in attentional guidance, response initiation or overall search performance.

Although the Ratiu et al. (2017) results are very consistent across all three of their experiments and test the same research question as Friesen et al. (2014) (viz., Are bilinguals better than monolinguals in conjunctive visual search under difficult conditions?), their materials and procedures were quite different. Thus, one purpose of the present study was to conduct a close replication of the critical conditions of the Friesen et al. (2014) search task. Another failure to replicate would deepen the skepticism that bilinguals show consistent advantages in selective attention, at least as reflected in performance during conjunctive visual search.

The Ambiguous Figures Task

It has been proposed that the ambiguous figures task also provides a measure of selective attention (Chun-Fat-Yim et al., 2017). Young adult participants were presented with seven sequences of 11 drawings one at a time. The first drawing in each set was an unambiguous object that gradually changed into a different unambiguous object. Based on prior testing the sixth card was the most ambiguous. The instructions were to predict the alternative object using the fewest number of drawings. The series continued until a correct response was made or the participant reached the last figure. The dependent measure was the mean number of drawings it took to identify the alternative object. Lower scores presumably reflect better selective attention. Given that the bilingual group identified the alternative object earlier on in the series compared to the monolingual group, Chun-Fat-Yim et al. (2017) suggested that bilinguals were better able to disengage attention from the salient features consistent with the first interpretation to those consistent with the second and evolving interpretation. Although Chun-Fat-Yim et al. (2017) allow that the ability to disengage the focus of attention in order to selectively attend to new information "*involves EF*," they emphasize that it is not equivalent to EF and "*is not defined by its components such as inhibition*" p. 371. A second purpose of the present study is to conduct a close replication of the Chun-Fat-Yim et al. (2017) study using the same ambiguous figures task.

Congruency Sequence Effects

Grundy et al. (2017) further pursued the hypothesis that bilinguals are better than monolinguals at disengaging attention by comparing the magnitude of congruency sequence effects (CSEs). CSEs are robust context effects observed in many choice RT tasks that include both congruent and incongruent trials. Alternative names include the Gratton Effect (Gratton et al., 1992), sequential congruency effects, and conflict adaptation effects. The term CSE will be used descriptively to describe a specific outcome, namely, that the congruency effect is significantly smaller following incongruent trials than following congruent trials. In their first two experiments using a flanker task Grundy et al. (2017) observed no language-group differences in the magnitude of the simple flanker effect, but bilinguals did have significantly smaller CSEs compared to monolinguals. The smaller CSE was interpreted as reflecting “...more rapid disengagement of attention and greater ability to refocus on the current trial” p. 45. The findings are asserted to “...provide insight into why some studies show bilingual advantages on executive control tasks and some do not” p. 52.

Paap (2018) discusses several reasons why the Grundy et al. (2017) results and interpretation should be discounted. First, the results have consistently failed to replicate. See Table 3 of Paap (2018) for descriptive and inferential statistics associated with 10 failures to replicate across three different laboratories. Second, the Grundy et al. (2017) account does not mesh with Botvinick’s influential Conflict Adaptation Model which assumes that CSEs are the consequence of activating control plans for trial n based on the amount of conflict detected on trial $n-1$ rather than a potentially disruptive carryover effect from trial $n-1$ to n . Third, the assumption that smaller CSEs are good and are caused by a more rapid disengagement of attention and better ability to refocus on the present trial produces a contradiction. Grundy et al. (2017) reported null results (no group differences in the magnitude of the CSEs) in their Experiment 3 which they suggest is due to the relatively long response stimulus intervals “during which all participants would have had sufficient time to disengage attention” p. 51. But, this cannot be the case because the CSEs were equally robust for both groups. If CSEs are the product of carryover effects and if all participants had sufficient time to disengage attention, then all participants should have CSEs near zero. A related, but subtly different point is that CSE magnitudes are unrelated to overall task performance² and, consequently, do not provide insights into the necessary and sufficient conditions for predicting bilingual advantages that matter in everyday life.

Switch Costs

Grundy et al. (2017) point out that in cued switching tasks, when the task shifts from one dimension (e.g., sort on color) to another (e.g., sort on shape), participants must rapidly disengage from information that was relevant and refocus on information that was previously irrelevant. They cite

²When non-verbal interference tasks include an equal number of congruent and incongruent trials the CSE (the smaller congruency effect following an incongruent trial) is caused by a symmetrical speed-up on incongruent trials and slow-down on congruent trials (see Figure 1 of Paap et al., 2016). The net result is that overall task performance remains unchanged and independent of the magnitude of the CSE.

Prior and MacWhinney (2010) and Prior and Gollan (2013) as showing that switch costs are smaller for bilinguals compared to monolinguals. These two early studies showing bilingual advantages are very difficult to replicate. For example, Paap et al. (2017) reported null effects in a large sample study using three-different switching tasks. More generally, Lehtonen et al.’s (2018) meta-analysis based on 77 comparisons across various types of switching tasks showed an average effect size of $g = 0.15$ [$+0.06$, $+0.24$] that disappeared when corrected for publication bias, $g = 0.02$ [-0.09 , $+0.14$]. The results are no different when the analysis is restricted to the clearly non-verbal color-shape task with manual responses: Based on our current database of 16 articles³ and 25 such tests the mean bilingual advantage is 4.9 ms and the 95% confidence interval straddles zero, $t(24) = 0.87$, $p = 0.39$, $CI[-7, +16]$. The apparently robust bilingual advantage in switch costs reported in the seminal article by Prior and MacWhinney (2010) has turned out to be anomalous. If Grundy et al. (2017) are correct to characterize switch costs as a valid measure of the disengagement of attention, then the meta-analyses offer no support for the hypothesis of bilingual advantages in attentional control.

Inhibition-of-Return (IOR)

In Posner’s cue-target paradigm (described in Klein, 2000) the interval between the rapid onset of a peripheral cue and a later target is varied. When the target appears in the cued location the typical finding is a brief period of facilitation followed by a longer period of inhibition known as inhibition of return (IOR). According to Klein (2000), the appearance of IOR is dependent on how quickly attention is endogenously disengaged from the cued location. Even though using “inhibition” as a marker of attentional control is somewhat ironic in the present context, it appears that the relative timing of IOR provides a fairly direct test of the hypothesis that bilinguals have learned how to rapidly disengage attention. Grundy et al. (2017) cite Mishra et al.’s (2012) report that high-proficiency bilinguals display IOR effects at earlier SOAs than low-proficiency bilinguals as support for this hypothesis. However, the Mishra et al. (2012) study did not include monolinguals. In a study that actually did compare bilinguals ($n = 24$) to monolinguals ($n = 28$) there were no group differences in the time course of IOR (Hernández et al., 2010). Furthermore, in a replication and extension of their earlier work Saint-Aubin et al. (2018) tested a large sample of English–French bilinguals and reported no effects of L2 proficiency on the IOR. Saint-Aubin et al. (2018) concluded that there is no reliable evidence that mastering a second language leads to faster or more potent disengagement of endogenous attention.

Working Memory as Executive Attention

Bialystok (2017) asserts that working memory (WM) capacity, conceptualized not as storage space, but as the extent to which resources are available to control attention “...is compatible with the evidence found across the life span for bilingualism-dependent plasticity” p. 249. A recent meta-analysis by von Bastian et al. (2017) evaluated this conceptualization of EF for

³The articles are listed in Table 2 of Paap (2018).

bilingual advantages. A set of 88 studies with 108 independent comparisons were included. The average effect size was $g = +0.11$ [$+0.03, +0.19$]. Considering the Bayes Factor associated with each effect size there was a high degree of heterogeneity, mostly null effects, and little evidence for the alternative hypothesis. Neither age (children, younger adults, older adults) nor task mode (verbal versus non-verbal) moderated the variability in effect sizes. Lehtonen et al. (2018) also examined the WM domain and their meta-analysis of 243 effect sizes yielded a mean effect size of $g = +0.07$ [$0.00, +0.13$] that shifted to a disadvantage when corrected for bias, $g = -0.07$ [$-0.17, +0.03$]. The Lehtonen et al. (2018) meta-analysis reinforces the conclusion of von Bastian et al. (2017) that the findings “challenge executive-attention accounts of bilingual advantages.”

MATERIALS AND METHODS

Procedures

All participants completed the following activities in this order: (1) the conjunctive visual search task, (2) the Raven's test, (3) the language background questionnaire, (4) demographic questions, (5) the ambiguous figures task and (6) the multilingual naming task (MINT) of productive vocabulary (Gollan et al., 2012).

Participants

The 141 participants were San Francisco State University (SFSU) undergraduate students who participated for credit or extra credit in a psychology course. The protocol was approved by the SFSU Institutional Review Board. All subjects gave written informed consent in accordance with the Declaration of Helsinki. Participants were 18–29 years old. The language background questionnaire is the same as that used by Paap et al. (unpublished) and appears in the appendix to that article. The means for bilinguals and monolinguals on several background and language-use variables are shown in Table 1.

The groups do not significantly differ on the Raven's measure of general fluid intelligence, but the non-significant difference favors the monolinguals. The results of the t -tests reported later do not change if Raven's scores are taken as a covariate. SES is

a composite measure of mother's education, father's education, and family income. When we sample from the SFSU student population the monolinguals typically have a significantly higher degree of SES compared to the bilinguals. However, in this population the measures of SES are never significantly correlated with any of the measures of executive functioning and, indeed, those correlations are often near zero (Paap and Greenberg, 2013; Paap et al., 2014, 2017; Paap et al., unpublished). As reported in the results SES did not significantly correlate with the measures of selective attention. Variables that are uncorrelated with the dependent variable cannot be the cause of a null result that would otherwise show a bilingual advantage.

As shown in Table 1 the bilinguals actively use two languages. On average their second most proficient language is self-rated as a 5.2 and a rating value of 5 was labeled “Almost as good as a typical native speaker on both everyday topics and specialized topics I know about.” They use their other language about one-third of the time. Their mean frequency of switching is 3.5 on a five-point scale where 3 is “a couple of times a day” and 4 is “several times a day.”

Visual Search Task

The visual search task was modeled on that used by Friesen et al. (2014). Participants were instructed to search for a blue-triangle target and to press the “1” key if it was present and to press the “0” key if it was not. The visual arrays remained on the screen until a response was made. The next visual array appeared immediately after a response was made. The target randomly appeared in one of the 26 locations on the screen. Given that Friesen et al. (2014) reported a bilingual advantage only in the low discriminability conjunctive search condition, the feature-search and high discriminability conditions were omitted. Thus, search type (target present vs. target absent) and distractor set size (5, 15, 25) were manipulated. In conjunctive search, two features (e.g., shape and color) need to be identified in the target stimulus (e.g., blue triangle) in order to distinguish it from the distractor stimuli. The distractors were purple triangles and blue diamonds. The targets and distractors have low discriminability because purple is similar to blue and diamonds are similar to triangles. There were 24 target-present trials and 18 target-absent trials. There

TABLE 1 | Differences between bilinguals and monolinguals on demographics and language use.

Measure	Bilinguals		Monolinguals		Diff	SE	t	p
	n	Mean	n	Mean				
Age	79	22.3	44	21.4	0.9	0.48	+1.92	0.058
Ravens	77	8.2	44	9.1	−0.9	0.48	−1.89	0.061
SES	79	4.1	44	4.9	−0.8	0.23	−3.50	0.001
Most proficient language	79	6.4	44	6.5	−0.1	0.11	−0.71	0.479
English MINT	69	61.7	39	63.9	−2.2	0.70	−3.31	0.002
Second most proficient language	79	5.2	44	0.8	+4.4	0.24	18.09	<0.001
% of time use most proficient	79	65.4	44	98.6	−33.2	3.31	−10.0	<0.001
Language switches per day	79	3.5	44	0.7	+2.9	0.22	12.98	<0.001

Diff, Group Difference; SE, standard error.

were six displays in each combination of number of distractors and positive versus negative trials. The 42 displays were presented in a different random order for each participant.

Ambiguous Figures

In the ambiguous figures task participants were presented with seven sequences of 11 black-and-white line drawings. For each sequence the drawings were presented one at a time. The participant sat about 45 cm away from a Dell computer screen and each of the line drawings projected a visual angle of about 6.0°. In each set of figures, the first was an unambiguous object that morphed in discrete steps into a different unambiguous figure. Participants were shown the first figure and were prompted with the label that most observers readily see, for example, “most people see this drawing as a seal.” As each successive figure from the series was presented they were asked “Does it still look like a seal?” If the participant indicated that it no longer looked like the start object, they were asked to guess what it might be morphing into. The first dependent variable for this ambiguous figures task (AF1) was the trial number of the drawing that no longer looked like the start object. The sequence continued until the participant correctly identified the new object. The second dependent variable (AF2) was the trial number of the drawing that was correctly identified as the new object.

To illustrate the difference between the variables consider the following scenario. A participant is shown the first figure of the seal/horse set and told that most people see a seal. As second, third, and fourth figures are shown the participant continues to report seeing a seal, but when shown the fifth figure from the sequence she says it no longer looks like a seal. Her AF1 score for this set is therefore “5”. If her response to the follow-up question is that it now looks like a horse, then her AF2 score would also be “5.” However, if she does not guess the identity of the new object until she is shown the seventh figure, then her AF2 score would be “7.” If the participant was unable to correctly identify the new object after seeing the 11th and last figure in the sequence the AF2 score was assigned a value of 11.

Given the assumptions of Chun-Fat-Yim et al. (2017) higher scores signal poorer ability to disengage attention. All participants saw the sets in the following order: Seal/Horse, Old Man/Lady, Apple/Face, Rat/man, Lady/Sax Swan/Squirrel, Body/Face. Chun-Fat-Yim et al. (2017) did not prompt their

participants with the label of the start object and used only the second dependent measure. During pilot testing we discovered that some participants did not correctly recognize the start object or saw it as a visually similar but different object. Although we were reluctant to deviate from Chun-Fat-Yim et al.’s (2017) procedure, the upside is that the two dependent variables may reflect two stages of selective attention: a disengagement of the salient features that promote the interpretation of the start object (AF1) versus an engagement of the salient features associated with the other object (AF2).

RESULTS

Visual Search

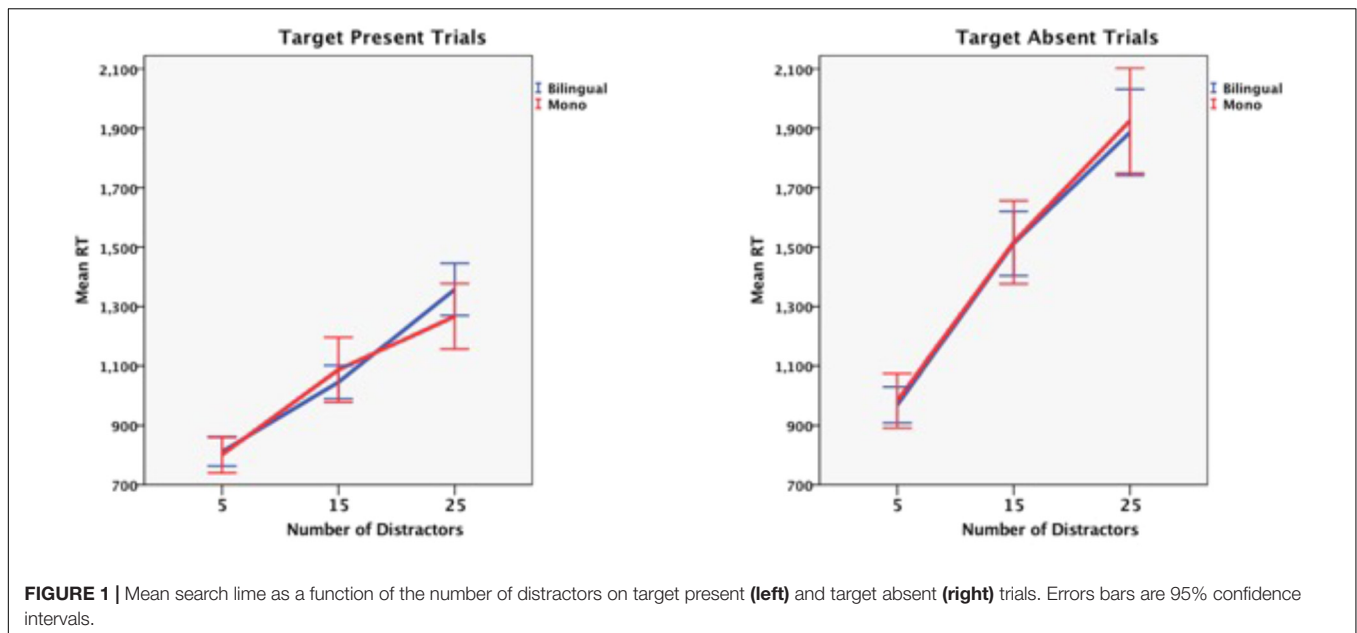
Search times less than 200 ms or more than 2.5 standard deviations above the participant’s mean for each condition were removed as were incorrect responses. This was identical to the procedures used by Friesen et al. (2014) Trials consisting of a target with no distractors provide a measure of the speed of basic perceptual-motor processes. There was no difference between the groups on these trials, $t(115) = -1.14, p = 0.257$.

Three-way mixed ANOVAs were performed separately on the RT and proportion correct (PC) data with Language Group (bilingual vs. monolingual) as a between-subjects factor and Trial Type (target present vs. target absent) and Number of Distractors (5, 15, 25) as repeated measures. The means and SEs in each condition are shown in **Table 2**. As expected, the RT analysis showed a significant main effect of Trial Type, $F(1,115) = 140, p < 0.001$ whereby it took longer to respond when no target was present and participants always had to search the entire display. Likewise the significant main effect of Number of Distractors, $F(1,115) = 382, p < 0.001$ confirmed that search times increase as the number of distractors increase. However, there was no significant main effect of Group, $F(1,115) = 0.03, p = 0.854$, nor was Group involved in any significant interactions. **Figure 1** shows the mean search time (for each group) as a function of the number of distractors for target-present and target-absent trials. Visual inspection confirms that there are no trends favoring bilingual advantages in search time.

Given that Friesen et al. (2014) obtained bilingual advantages only in the low discriminability condition it is important to show that the low discriminability displays used in the present study produced comparable levels of difficulty. The means (estimated

TABLE 2 | Means and standard deviations for reaction time and proportion correct for monolinguals and bilinguals in each condition defined by trial type and number of distractors.

	Positive trials Number of distractors				Negative trials Number of distractors		
	0	5	15	25	5	15	25
RT (SD)							
Bilingual	617 (8)	812 (24)	1046 (28)	1358 (44)	969 (30)	1512 (54)	1886 (73)
Monolingual	655 (30)	824 (37)	1097 (54)	1273 (54)	961 (50)	1482 (76)	1881 (96)
PC (SD)							
Bilingual	0.98 (0.06)	0.97 (0.09)	0.89 (0.15)	0.87 (0.14)	0.97 (0.07)	0.97 (0.08)	0.95 (0.09)
Monolingual	0.97 (0.09)	0.97 (0.07)	0.89 (0.12)	0.85 (0.16)	0.94 (0.17)	0.93 (0.17)	0.92 (0.18)



from Figure 3 of Friesen et al. (2014) and averaged across both language groups) for positive trials with 5, 15, and 25 distractors were about 980, 1300, and 1460 ms, respectively. Based on these values the slope of the best fitting straight line was 24 ms and this is very close to the 25 ms slope obtained for the positive trials in the present study. It seems that the conjunctive search conditions in the two studies are equally difficult.

Slope is arguably a purer measure of the ability to disengage and re-engage attention than overall search time. Consequently the slope for each participant across set sizes of 5, 15, and 25 distractors were computed for positive and negative trials separately (Pfister et al., 2013). Independent *t*-tests on these individual slopes compared language groups and showed no difference for either positive trials, $t(119) = -0.55$, $p = 0.581$, or negative trials, $t(116) = 0.27$, $p = 0.781$.

The mean proportion correct (PC) and SDs in each condition are shown in the bottom part of **Table 2**. The three-way mixed ANOVA showed no significant main effect of Language Group, $F(1,115) = 2.274$, $p = 0.134$; nor was Group involved in any significant interactions with Number of Distractors or Trial Type.

Continuous Measures of Bilingualism

Rather than relying exclusively on categorizing participants as bilinguals, monolinguals, and undetermined; the entire sample can be used to examine the relationships between aspects of bilingualism (proficiency of the less dominant language, percentage of most used language, frequency of daily switching) and the measures of selective attention (RT, slope, and PC). These bivariate correlations are shown in **Table 3** and no aspect of bilingualism significantly predicts performance in the search task.

Ambiguous Figures

The first dependent variable, AF1, was the mean number of drawings examined before it no longer looks like the start object.

The means for monolinguals ($n = 43$, $M = 4.0$) and bilinguals ($n = 79$, $M = 4.1$) did not significantly differ, $t(120) = -0.247$, $p = 0.806$. The second dependent variable, AF2, was the mean number of figures examined before correctly identifying the second object. Again, the means for monolinguals ($M = 6.3$) and bilinguals ($M = 6.5$) did not differ, $t(120) = -0.735$, $p = 0.464$, on this dependent variable either.

Despite the change in procedure that led to the addition of the first dependent variable, the overall mean of the ambiguous figure that yielded a correct identification of the new object was 6.4 in both studies. Our results offer no evidence of a bilingual advantage in the disengagement of attention. The continuous measures of bilingualism reported for the visual search task were also correlated with both dependent variables in the ambiguous figures task. All six correlations had magnitudes less than 0.09 and, consequently were not significant despite an N of 128.

Bayes Factor Analyses

Bayes factor analyses calculate the ratio of probability of the null hypothesis given the data to the probability of the alternative given the data. The means and *t* values for the tests reported above were entered into Rouder's Bayes Factor (BF) calculator (Rouder et al., 2009)⁴ using the default prior of $r = 0.707$. All of the Bayes factor analyses are greater than 3 which according to Jeffrey's (1961) guidelines provide substantial evidence for the null hypothesis: overall search RT (4.5), overall search PC (4.7), slope on target trials (4.3), slope on no-target trials (4.7), AF1 (4.5), and AF2 (4.8).

Correlations Between Measures

Table 4 shows the within and between task correlations for the visual search and ambiguous figures tasks. The target only

⁴pcl.missouri.edu

TABLE 3 | Correlations between aspects of bilingualism and performance measures in the visual search task for all 127 participants.

Predictor	RT		PC		Slope	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Proficiency of second most proficient language	+0.004	0.965	−0.041	0.646	−0.017	0.851
Percentage use of most used language	−0.062	0.486	+0.083	0.356	+0.148	0.096
Frequency of daily switches	−0.048	0.591	−0.037	0.681	−0.033	0.761

RT, reaction time; PC, proportion correct; *r*, Pearson correlation; *p*, probability.

condition is the mean for the displays consisting of a single target with no distractors. Individual differences in the target only condition are likely to reflect differences in basic perceptual-motor processing. The correlations of the target only condition with the two slope measures are near zero and this is consistent with the assumption that search rate is independent of basic speed of processing. The correlation between the target-present and target-absent slopes is significant, but small, suggesting non-trivial differences in how the two types of displays are searched. This is consistent with the version of the guided search model developed by Chun and Wolfe (1996) that posits the setting of an activation threshold that terminates a non-exhaustive search more often for target absent trials than those where the target is present.

Turning to the two dependent variables measured in the ambiguous figures task it is not surprising that they are highly correlated as no longer seeing a figure as the start object should facilitate being able to organize the features into a new object. Of primary interest is whether there are cross-task correlations that would support the possibility that both tasks are tapping into a shared attentional control mechanism. But, as evident in **Table 4** neither slope measure significantly correlates with either AF measure. There is a significant correlation between the target only RT and the first AF measure, but there is no obvious reason why general processing speed should be related to a judgment made under no time pressure.

DISCUSSION

The main empirical goal of this study was to conduct a close, but not exact, replication of two studies interpreted to support bilingual advantages in attentional control, particularly the ability to disengage attention. The conjunctive visual search task that

yielded a bilingual advantage in Friesen et al. (2014) showed null results in the present experiment despite the fact that the studies produced nearly identical slopes of search time as a function of number of distractors. Furthermore, by examining slopes, target-absent trials, Bayes factors, and continuous measures of bilingualism the present study provided more tests of the hypothesis. Thus, the present study, together with the null results reported by Ratiu et al. (2017) seriously dampen the likelihood that bilingual advantages will consistently occur in search tasks.

The close replication of the Chun-Fat-Yim et al. (2017) study yielded overall means for identifying the new object that were identical in the two studies, but the present study showed no differences between bilinguals and monolinguals. The present study added a dependent variable (AF1, the drawing that no longer looks like the start object) that potentially separates attentional disengagement from re-engagement, but still no group differences were observed. One possible reason for the group differences reported by Chun-Fat-Yim et al. (2017) is that their bilinguals had higher maternal education, marginally higher fluid intelligence ($p = 0.051$), and a higher proportion of immigrants.

The Revised Hypothesis Revisited

Bialystok's revised hypothesis is plausible and quite appealing, but before it can be rigorously tested it needs further specification. The looseness of the construct is reflected in the absence of a *pater familias* as in different articles and across different contexts the revised hypothesis is described in terms of executive attention, selective attention, or the disengagement of attention. Here we will introduce the term *attentional control* for a hypothetical construct that is presumed to be critical for bilingual language control. What is its essence? Are there any defining features or are there only characteristic features? If an important aspect of attentional control is the ability to focus on task relevant information and ignore irrelevant distracting information, then different types of selection are possible. In a flanker task a designated target object can be selected at the expense of the irrelevant object by spatially attending to the target. At least in theory, selection could also be the conflict resolution mechanism in a Simon task, but not via spatial attention because the task relevant information (e.g., color) and irrelevant information (e.g., location) are two attributes of the same stimulus. Neither of these types of selection seems to have much in common with selecting the lexical entry “gato” (or the entire Spanish lexicon) and leaving “cat” (or the entire English lexicon) behind when asked

TABLE 4 | Pearson correlations between specified measures from visual search and ambiguous figures task.

	Present slope	Absent slope	AF1	AF2
Target only	−0.05	+0.01	+0.20*	+0.13
Target present slope		+0.21*	−0.08	+0.07
Target absent slope			+0.04	+0.15
AF1				+0.67**
AF2				1

AF1, ambiguous figure 1 (no longer looks like the start object); AF2, ambiguous figure 2 (correct identification of the new object).

to name a picture of a domesticated feline in Spanish. The point here is that shifting the conflict-resolution mechanism from inhibition to attentional control doesn't solve the problem of identifying the specific mechanism(s) used during bilingual language control and the degree to which the mechanisms are shared with non-verbal tasks.

In the absence of a more detailed proposal regarding the attentional control involved in bilingual language control, it is difficult to predict when bilingual advantages should occur and when they would be unlikely to occur. This allows the non-productive practice of attributing bilingual advantages to attentional control when differences occur and ignoring the null results. Are we foisting the results of the non-verbal interference tasks on Bialystok's revised hypothesis? Beyond the logical argument drawn above consider that Bialystok (2017) includes the antisaccade, stop-signal, color-shape switching, and Simon as tasks that fall "broadly into a category of attention tasks" (p. 241). Furthermore, Bialystok suggests that the attentional system enhanced by bilingualism is similar in many respects to Posner's "executive attention." Yet executive attention is operationally defined in the seminal article by Fan et al. (2002) as the flanker interference effect (incongruent RT – congruent RT) in the attentional network task (ANT). Furthermore, Fan et al. (2002) state that executive control is defined as resolving conflict among responses.

To reiterate, if resolving the conflict between a bilingual's two languages is the presumed cause of bilingual advantages and if this conflict-resolution mechanism recruits a general control ability, then bilingual advantages should occur in a wide array of non-verbal interference tasks. The only way to avoid this prediction is to make an additional *post hoc* assumption that only a subset of interference tasks use the general-purpose attentional control mechanism as a conflict resolution mechanism. Therefore, what is needed is a principled way to sort interference tasks into those where the conflict resolution mechanism is clearly attentional selection (and according to the revised hypothesis should show bilingual advantages) and those where conflict resolution relies on inhibition or some other task-specific mechanism (and consequently, according to the revised hypothesis should not show bilingual advantages). One step toward clarifying a construct of attentional control might use latent-variable analyses to determine if measures assumed to reflect

attentional control all load on a common factor even if subsets are separable. If no such latent structure exists, then the hypothetical attentional-control construct may simply be chimerical.

CONCLUSION

The review of the relevant prior literature showed that significant bilingual advantages in executive functioning (and especially the inhibitory control component) were relatively rare and that the average effect size was very small and plausibly due to file drawer and publication biases. Despite the exciting early reports of bilingual advantages, advantages in inhibitory control for bilinguals age six and older and for bilinguals who are older adults are more myth than reality. The proposal that bilingual advantages are rooted in attentional control rather than executive functioning is worthy of investigation, but the challenges are mounting rapidly as this revised hypothesis is tested in conjunctive visual search, the ambiguous figures task, CSEs, and IOR. Furthermore, to the extent that tasks such as the flanker or color-shape switching also recruit attentional control, these too should consistently produce bilingual advantages, not null results and effect sizes that straddle zero.

AUTHOR CONTRIBUTIONS

The selective attention studies formed part of a masters thesis completed by RA-J. All authors contributed to the design and conduct of the experiments. RA-J, LM, KA, and BZ coded the MINT and ambiguous figures recordings and contributed to the APS poster that reported the results of the experiments. RA-J and KP performed the data-analyses.

ACKNOWLEDGMENTS

We thank the following members of the Language, Attention, and Cognitive Engineering (LACE) laboratory for their contributions to this project: Ester Avadavat, Katrina Lao, Divya Subramanian, Lesley Primero, Jennifer Lai, and Karla Barajas.

REFERENCES

- Adescope, O. O., Lavin, T., Thompson, T., and Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Rev. Educ. Res.* 80, 207–245. doi: 10.3102/0034654310368803
- Antón, E., Duñabeitia, J. A., Estévez, A., Hernández, J. A., Castillo, A., Fuentes, L. J., et al. (2014). Is there a bilingual advantage in the ANT task? Evidence from children. *Front. Psychol. Lang. Sci.* 5:398. doi: 10.3389/fpsyg.2014.00398
- Antón, E., García, Y. F., Carreiras, M., and Duñabeitia, J. A. (2016). Does bilingualism shape inhibitory control in the elderly? *J. Mem. Lang.* 90, 147–160. doi: 10.1016/j.jml.2016.04.007
- Bialystok, E. (2017). The bilingual adaptation: how minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1016/j.tics.2012.03.001
- Bialystok, E., Craik, F. I., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303. doi: 10.1037/0882-7974.19.2.290
- Bialystok, E., Craik, F. I. M., Green, D. W., and Gollan, T. H. (2009). Bilingual minds. *Psychol. Sci. Public Interest* 10, 89–129. doi: 10.1177/1529100610387084
- Blumenfeld, H. K., and Marian, V. (2014). Cognitive control in bilinguals: advantages in stimulus-stimulus inhibition. *Bilingual. Lang. Cogn.* 17, 610–629. doi: 10.1017/S1366728913000564
- Chun, M. M., and Wolfe, J. M. (1996). Just say no: how are visual searches terminated when there is no target present? *Cogn. Psychol.* 30, 39–78. doi: 10.1006/cogp.1996.0002
- Chun-Fat-Yim, A., Sorge, G. B., and Bialystok, E. (2017). The relationship between bilingualism and selective attention in young adults: evidence from

- an ambiguous figures task. *Quart. J. Exp. Psychol.* 70, 366–372. doi: 10.1080/17470218.2017.1221435
- de Bruin, A., Treccani, B., and Della Sala, S. (2014). Cognitive advantage in bilingualism: an example of publication bias? *Psychol. Sci.* 26, 99–107. doi: 10.1177/0956797614557866
- Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J., et al. (2014). The inhibitory advantage in bilingual children revisited: myth or reality? *Exp. Psychol.* 61, 234–251. doi: 10.1027/1618-3169/a000243
- Fan, J., McCandliss, B. D., Sommer, T., Raz, A., and Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *J. Cogn. Neurosci.* 14, 340–347. doi: 10.1162/089892902317361886
- Friesen, D. C., Latman, V., Calvo, A., and Bialystok, E. (2014). Attention during visual search: the benefit of bilingualism. *Int. J. Bilingual.* 19, 693–702. doi: 10.1177/1367006914534331
- Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., and Cera, C. M. (2012). Self-ratings of spoken language dominance: a Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingual. Lang. Cogn.* 15, 594–615. doi: 10.1017/S1366728911000332
- Gratton, G., Coles, M. G., and Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *J. Exp. Psychol. Gen.* 121, 480–506. doi: 10.1037/0096-3445.121.4.480
- Grundy, J. G., Chung-Fat-Yim, A., Friesen, D. C., Mak, L., and Bialystok, E. (2017). Sequential congruency effects reveal differences in disengagement of attention for monolingual and bilingual young adults. *Cognition* 163, 42–55. doi: 10.1016/j.cognition.2017.02.010
- Hernández, M., Costa, A., Fuentes, L. J., Vivas, A., and Sebastian-Galles, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Bilingual. Lang. Cogn.* 13, 315–325. doi: 10.1017/S13667289099909010
- Hilchey, M. D., Saint-Aubin, J., and Klein, R. M. (2015). “Does bilingual exercise enhance cognitive fitness in non-linguistic executive processing tasks,” in *Cambridge Handbook of Bilingual Processing*, ed. J. W. Schwieter (Cambridge: Cambridge University Press), 586–613.
- Jeffreys, H. (1961). *Theory of Probability*, 3rd Edn. Oxford: Oxford University Press, Clarendon Press.
- Klein, R. M. (2000). Inhibition of return. *Trends Cogn. Sci.* 4, 138–147. doi: 10.1016/S1364-6613(00)01452-2
- Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., and Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychol. Bull.* 144, 394–425. doi: 10.1037/bul0000142
- Magen, H., and Cohen, A. (2007). Modularity beyond perception: evidence from single task interference paradigms. *Cogn. Psychol.* 55, 1–36. doi: 10.1016/j.cogpsych.2006.09.003
- Mishra, R. K., Hilchey, M. D., Singh, N., and Klein, R. M. (2012). On the time course of exogenous cuing effects in bilinguals: higher proficiency in a second language is associated with more rapid endogenous disengagement. *Quart. J. Exp. Psychol.* 65, 1502–1510. doi: 10.1080/17470218.2012.657656
- Paap, K. R. (2019). “Bilingualism in cognitive science: the characteristics and consequences of bilingual language control,” in *The Cambridge Handbook of Bilingualism*, eds A. De Houwer and L. Ortega (Cambridge: Cambridge University Press), 435–465.
- Paap, K. R. (2018). “The bilingual advantage debate: quantity and quality of the evidence,” in *The Handbook of the Neuroscience of Multilingualism*, ed. J. W. Schwieter (West Sussex: Wiley-Blackwell).
- Paap, K. R., and Greenberg, Z. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cogn. Psychol.* 66, 232–258. doi: 10.1016/j.cogpsych.2012.12.002
- Paap, K. R., and Sawi, O. (2014). Bilingual advantages in executive functioning: problems in convergent validity, divergent validity, and the identification of the theoretical constructs. *Front. Psychol.* 5:962. doi: 10.3389/fpsyg.2014.00962
- Paap, K. R., and Sawi, O. (2016). The role of test-retest reliability in measuring individual and group differences in executive functioning. *J. Neurosci. Methods* 274, 81–93. doi: 10.1016/j.jneumeth.2016.10.002
- Paap, K. R., Johnson, H. A., and Sawi, O. (2014). Are bilingual advantages dependent upon specific tasks or specific bilingual experiences? *J. Cogn. Psychol.* 26, 615–639. doi: 10.1080/20445911.2014.944914
- Paap, K. R., Johnson, H. A., and Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69, 265–278. doi: 10.1017/j.cortex.2015.04.014
- Paap, K. R., Johnson, H. A., and Sawi, O. (2016). Should the search for bilingual advantages in executive functioning continue? *Cortex* 74, 305–314. doi: 10.1016/j.cortex.2015.09.010
- Paap, K. R., Myuz, H. A., Anders, R. T., Bockelman, M. F., Mikulinsky, R., and Sawi, O. M. (2017). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *J. Cogn. Psychol.* 29, 89–112. doi: 10.1080/20445911.2016.1248436
- Pfister, R., Schwarz, K., Carson, R., and Janczyk, M. (2013). Easy methods for extracting individual regression slopes: comparing SPSS, R, and Excel. *Tutorials Quant. Methods Psychol.* 9, 72–78. doi: 10.20982/tqmp09.2.p072
- Prior, A., and Gollan, T. H. (2013). The elusive link between language control and executive control: a case of limited transfer. *J. Cogn. Psychol.* 25, 622–645. doi: 10.1080/20445911.2013.821993
- Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingual. Lang. Cogn.* 13, 253–262. doi: 10.1017/S1366728909990526
- Ratiu, I., Hout, M. C., Walenchok, S. C., Azuma, T., and Goldinger, S. D. (2017). Comparing visual search and eye movements in bilinguals and monolinguals. *Attent. Percept. Psychophys.* 79, 1695–1725. doi: 10.3758/s13414-017-1328-3
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* 16, 225–237. doi: 10.3758/PBR.16.2.225
- Saint-Aubin, J., Hilchey, M. D., Mishra, R., Singh, N., Savoie, D., Guitard, D., et al. (2018). Does the relation between the control of attention and second language proficiency generalize from India to Canada? *Can. J. Exp. Psychol.*
- Salthouse, T. A. (2010). Is flanker-based inhibition related to age? Identifying specific influences of individual differences on neurocognitive variables. *Brain Cogn.* 73, 51–61. doi: 10.1016/j.bandc.2010.02.003
- Sanchez-Azanza, V. A., López-Penadés, R., Buil-Legaz, L., Aguillar-Mediavilla, E., and Adrover-Roig, D. (2017). Is bilingualism losing its advantage? A bibliometric approach. *PLoS One* 12:e0176151. doi: 10.1371/journal.pone.0176151
- von Bastian, C., de Simoni, C., Kane, M., Carruth, N., and Miyake, A. (2017). Does being bilingual entail advantages in working memory? A meta-analysis. *Paper Presented at the Meeting of the Psychonomic Society*, Madison, WI.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Paap, Anders-Jefferson, Mason, Alvarado and Zimiga. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Recognizing the Effects of Language Mode on the Cognitive Advantages of Bilingualism

Ziying Yu^{1,2*} and John W. Schwieter^{3,4*}

¹ Department of English Language and Literature, Fudan University, Shanghai, China, ² Department of Linguistics, University of California, Santa Barbara, Santa Barbara, CA, United States, ³ Language Acquisition, Multilingualism, and Cognition Laboratory, Wilfrid Laurier University, Waterloo, ON, Canada, ⁴ Bilingualism, Translation, and Cognition Laboratory, University of California, Santa Barbara, Santa Barbara, CA, United States

OPEN ACCESS

Edited by:

Roberto Filippi,
University College London,
United Kingdom

Reviewed by:

Antonella Sorace,
University of Edinburgh,
United Kingdom
Sara Incera,
Eastern Kentucky University,
United States

*Correspondence:

Ziying Yu
ziying_yu@mail.ucsb.edu
John W. Schwieter
jschwieter@wlu.ca

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Psychology

Received: 18 January 2018

Accepted: 05 March 2018

Published: 20 March 2018

Citation:

Yu Z and Schwieter JW (2018)
Recognizing the Effects of Language
Mode on the Cognitive Advantages
of Bilingualism. *Front. Psychol.* 9:366.
doi: 10.3389/fpsyg.2018.00366

For bilinguals, it is argued that a cognitive advantage can be linked to the constant management and need for conflict resolution that occurs when the two languages are co-activated (Bialystok, 2015). Language mode (Grosjean, 1998, 2001) is a significant variable that defines and shapes the language experiences of bilinguals and consequently, the cognitive advantages of bilingualism. Previous work, however, has not sufficiently tested the effects of language mode on the bilingual experience. In this brief conceptual analysis, we discuss the significance of language mode in bilingual work on speech perception, production, and reading. We offer possible explanations for conflicting findings and ways in which future work should control for its modulating effects.

Keywords: language mode, language activation, cognitive benefits of bilingualism, language control, multilingualism

INTRODUCTION

The claim that the knowledge and use of multiple languages gives rise to cognitive benefits is a hotly debated area of research in psycholinguistics and bilingualism (see Barac et al., 2014; Blom et al., 2017 for recent reviews). At any given point in time, and based on numerous psychosocial, situational, and linguistic factors, a bilingual must decide which language to use and how much of the other irrelevant language must be controlled or suppressed (Green, 1998). It is this constant management and monitoring of more than one language system that may be most responsible for the reported advantages in general executive functions (Bialystok et al., 2012; but see Hilchey and Klein, 2011; Paap et al., 2015 for alternate views).

Although a bilingual's two languages are constantly in a state of co-activation, Green and Abutalebi's (2013) Adaptive Control Hypothesis argues that the relative degree¹ of such activation for each language is dynamically adaptive. This hypothesis builds on the fact that bilinguals vary regarding language use in several contexts (Green, 2011; Prior and Gollan, 2011) and links this variation with underlying cognitive and neural control mechanisms. Green and Abutalebi (2013) argue that the control mechanisms adapt in response to bilingual experiences and to the recurrent demands placed on them in interactional situations. It might be the case that bilinguals outperform their monolingual counterparts in some cognitive tasks because this advantage may

¹By 'degree of activation,' we refer to the *magnitude* of language activation. While even today, it "remains to be determined what it means to say that languages can be activated to different degrees" (Dijkstra, 2005, p. 199), a recent study by Incera and McLennan (2018) found that differences in the timing of interference, but not in the magnitude of interference, led to differential effects within and between languages.

be a representation of their superior ability to be adaptive to situational (e.g., experimental) needs. If this is the case, various experiences with language mode may confound the implications for a bilingual advantage. In order to study the nature of these effects more accurately and directly, researchers must take into account speakers' language experiences thoroughly (Luk and Bialystok, 2013; Schwieter and Ferreira, 2016). These experiences are modulated by the limited capacity and goal-directed selectivity of the human executive functions.

Grosjean (1998, 2001) proposed and developed a notion of *language mode* which refers to the state of activation of the bilingual's languages and language processing mechanisms at a given point in time. In other words, language mode concerns the degree of activation of the two languages in a bilingual's mind. According to Grosjean, due to the influence of the environment, bilinguals continuously and naturally find themselves on a situational continuum of language activation, ranging from a monolingual to a bilingual mode. Three hypothetical positions can be visualized in this framework. When the bilingual is said to be in or close to being in a *monolingual mode*, a base language (i.e., the primary language being processed or produced at the time, not necessarily L1 for a bilingual) is the most active in terms of environmental activation (since the base language can also be L2 for an unbalanced bilingual), while the other non-target language is much less activated (but never totally deactivated). When the bilingual moves on the continuum and stops at an *intermediate mode*, the non-target language is more active than in the monolingual mode, whereas the base language remains the most activated language. When the bilingual is in a *bilingual mode*, in which the two languages are utilized from time to time in the form of code-switching or borrowings, the non-target language is highly active (but not as active as the base language). At all three positions, the base language remains fully active, as it is the main language that governs language perception and production.

Language activation is modulated by several variables including participants' characteristics. While language mode focuses on environmental characteristics, participants' characteristics including language proficiency and language dominance may change the activation level of languages. Dunn and Fox Tree (2014, p. 611) argued that "although there was no interaction between language mode and bilingual dominance, language mode can be made clearer when bilingual proficiency is controlled." Consequently, we discuss participants' characteristics in this paper in cases where researchers are careful to minimize their confounding influence.

Given that language mode plays an important role in language activation, it likely should be considered a modulating factor in the bilingual advantage debate. However, these possible effects have been unintentionally ignored, oftentimes by employing experimental designs that place and maintain participants in an intermediate mode. This misrepresents the true bilingual experience which consists of diverse interactions with and placements on the language mode spectrum and consequently uncovers findings that may be ambiguous or conflicting. Below, we discuss the important role of language mode in research on bilingual language activation, including speech perception,

speech production, and reading. We offer ways in which studies investigating the cognitive advantages of bilingualism can consider the role of language mode.

SPEECH PERCEPTION

In the area of bilingual speech perception, bilingual participants from some studies (e.g., Spivey and Marian, 1999; Colomé, 2001) may have been closer to an intermediate mode on the language mode continuum. Although the researchers examined one base language throughout the experiment, a number of confounding factors may have activated the participants' other languages, consequently moving them away from the monolingual endpoint. Although it is difficult and perhaps impossible to place a bilingual in a complete monolingual mode throughout a task, a few experiments have attempted to control for language mode. In a lexical decision task, Soares and Grosjean (1984) compared Portuguese-English bilinguals' reaction times (RTs) to words and non-words in English monolingual mode, Portuguese monolingual mode, and code-switching bilingual mode. Their results demonstrated that bilinguals were slower to access code-switched words in the bilingual mode than they were for words in the monolingual modes.

Similar findings were reported by Dunn and Fox Tree (2014) whose study made strides toward controlling for language mode effects. The study examined both English monolinguals and English-Spanish bilinguals who were divided into two groups before the experiment. Bilingual participants were randomly yet equally assigned to either the bilingual mode group or the monolingual mode group (consisting of monolinguals), and their bilingual language proficiency was first roughly assessed on an online survey during the online registration process (i.e., target questions about language ability were hidden in a variety of questions), and then was further assessed by a language dominance scale assessment and an individual interview after the completion of the experiment, minimizing the influence of confounding variables such as language dominance and language proficiency. Therefore, all participants had little reason to expect that their multilingual ability would be relevant for the study. The researchers also scheduled data collection sessions at times that minimized participants' chance of encountering bilingual speakers or bilingual situations in the laboratory.

Dunn and Fox Tree's (2014) study used a matched-pair design in the three experimental parts: in Part 1, all participants including English-speaking monolinguals and Spanish-English bilinguals were approached by a non-Latino experimenter and were asked to perform an English lexical decision with all instructions in English. In Part 2, all participants viewed a silent video clip about the Pink Panther and were asked to retell the story to the experimenter. In the monolingual mode group, since the need for Spanish was not mentioned, both the bilinguals and monolinguals should assume that their retellings be done in English, an assumption on which they acted corrected. However, in the bilingual mode group, bilinguals were approached by a Spanish-speaking experimenter. They were told in the instructions that Spanish retellings would enrich the

database and were therefore asked to retell the story in Spanish. In Part 3, all participants performed another English lexical decision task. The results from the first lexical decision task (Part 1) showed that RTs did not differ between the two language mode groups. However, results from the second lexical decision task (Part 3) suggested that language mode significantly affected RTs such that bilinguals processed non-words slower in bilingual mode than in monolingual mode. Further analyses demonstrated similar RTs for English monolinguals and bilinguals in the monolingual mode. This finding not only supported the language mode hypothesis, but also appeared to diverge from Soares and Grosjean's (1984) argument that bilinguals, regardless of which language mode they find themselves in, access words slower than monolinguals. It is likely that these differential results can be explained by the influence of several confounding factors in an experimental context in which participants were not fully in a monolingual mode.

In addition to the lexical decision task, evidence supporting the notion of language modes comes from picture-word interference tasks. Marian and Spivey (2003) found that in a monolingual mode, interference from the second language (L2) on the first language (L1) was not found, forming a contrast with the cross-linguistic effects while in a bilingual mode. However, interference from L1 to L2 was significant. Using the same paradigm, Canseco-Gonzalez et al. (2010) found that early Spanish-English bilinguals displayed inter-lingual competition that was significantly larger when they were tested in a bilingual mode (9.5% more fixations on the cohort than on the unrelated object) than in a monolingual mode (5%). These two studies suggest that not only language mode (environmental characteristics) but also participants' characteristics such as language dominance and proficiency may modulate language activation. Importantly, the location on the language mode continuum will have a direct effect on language activation in terms of speech perception with inter-lingual competition becoming larger when the participants stay closer to the bilingual mode.

SPEECH PRODUCTION

Language mode also exerts considerable influence on bilinguals' language production. Jared and Kroll (2001) simulated a monolingual mode in the first part of a picture-naming experiment and found little activation of the non-target language when participants expected to see only one language and were given stimuli and instructions in the same language. This consistently appeared to be the case except for when the non-target language was the participant's dominant language. Nevertheless, when stimuli were given in both languages, as in Hermans et al. (2011), cross-language phonological co-activation appeared sensitive to the cognate status of the stimuli. These results supported the modulating effects of language mode: the higher the ratio of cognates to non-cognates, the higher the activation level of the non-target language.

In another study by Boukadi et al. (2015), Tunisian Arabic-French bilinguals named pictures in their L2 while ignoring

auditory distractors. Bilingual participants were all native speakers of Tunisian Arabic who started learning French from primary school. Their L2 proficiency was assessed by means of self-ratings and a lexical decision task. In Experiment 1, the non-target language (Tunisian Arabic) was entirely absent in the experimental setting: all instructions were exclusively given in French and the students were not informed that the research was related to bilingualism until the end of the study. The participants were asked not to use their native language under any circumstances and to only communicate with the experimenter in French. The target stimuli were line-drawings of common objects and the auditory distractors were presented in French. Four French words were selected for each picture to serve as distractors based on the following conditions: phono-translation (the distractor was phonologically related to the picture name in L1); semantic (the distractor and target picture were semantically related); phonological (the distractor was phonologically related to the picture name in L1); and unrelated (the distractor had no relation to the picture name). No significant differences between the unrelated, phono-translation, or semantic condition were observed, which indicated that lexical selection proceeded in a language-specific way when the experimental setting was maintained in a monolingual mode. More importantly, the phono-translation effect remained insignificant even when L2 proficiency was taken into account. In Experiment 2, both languages appeared in the task in order to create a bilingual experimental setting, and bilinguals, who were selected from the same pool as Experiment 1, knew that the research had to do with a topic on bilingualism. They were allowed to speak in their L1 and were asked to name pictures in their L2 while ignoring an auditory distractor in their L1. Although the explicit instructions in Experiment 1 may have activated the irrelevant language, breaking a purely monolingual environment, Experiment 1 still created an environmental situation in which participants were closer to the monolingual end on the continuum compared with Experiment 2. In terms of the stimuli, Experiment 2 used the same pictures as in Experiment 1, but the auditory distractors were in Tunisian Arabic (the semantic distractors were the equivalent Tunisian Arabic translation of the French semantic distractors in Experiment 1). The results showed that RTs were significantly longer in both the phono-translation (965 ms) and semantic condition (934 ms) compared to the unrelated condition (918 ms). Taken together, the results of the two experiments suggest that language selection during bilingual speech production is a dynamic process modulated by language mode; the closer to the bilingual end of the continuum, the more activated the non-target language becomes. These findings also support the notion that the language mode of the experiment has modulating effects on the activation of bilinguals' languages.

READING

Language mode also affects bilingual lexical access during word reading, as shown in a study carried out by Dijkstra et al. (2000). Dutch-English bilinguals with an average English learning time of 11.4 years participated in an English lexical decision

task including English-Dutch homographs and cognates, as well as exclusively English control words. The total stimulus set was composed of homographs (no semantic similarity across languages), controls, English fillers, Dutch fillers, and non-words (orthographically permissible in English and not homophonic to Dutch words). The experiment included two parts, each consisting of 28 blocks of 8 stimuli including one homograph and one control item. In part 1, the remaining six item slots of each block were randomly filled with only English fillers or non-words, whereas in part 2, Dutch words were also included. The participants received the same instructions and communication with the experimenter (in English) for an English lexical decision task, but they were explicitly told that word forms that exist in both English and Dutch (homograph) required a “yes” response, while words only belonging to Dutch required a “no” response. After the experiment, all participants filled in a questionnaire to assess their L2 (English) proficiency. In this regard, part 1 of this experiment could be regarded as being close to a monolingual mode. The results showed that the RTs for homographs in part 2 were considerably slower (613 ms) than in part 1 (575 ms). This suggested that lexical selection took more time in the bilingual mode than in the monolingual mode and that participant moved closer to non-selective language activation. In addition, it should also be taken into account that the transition from part 1 to part 2 was rather abrupt, as RTs to interlingual homographs (from 581 to 663 ms) were considerably slower immediately after the transition. Consequently, encountering non-target language items during the experiment changes the language mode and exerts immediate and severe effects on bilingual lexical access during reading.

In Experiment 3 of a study by De Groot et al. (2000), the researchers mixed real words in the non-target language and non-words in the target language, forming a comparison with their Experiment 2 in which all the non-words were neither real words in the irrelevant language nor were they a mixture of the two languages. As a result, the participants responded to homographs faster in Experiment 2 (557 ms) than in Experiment 3 (619 ms). This provides support that the participants performed the task differently depending on the language mode simulated in the experiments: bilinguals processed words faster when the setting was more language-specific. In another study, Lemhöfer and Radach (2009) conducted a pure-German, a pure-English, and a mixed lexical decision task on the same set of non-words. Results showed that RTs varied according to the context of the task: in the monolingual context, participants made more mistakes and took longer to reject non-words that were more similar to the target language; in the bilingual context, RTs were significantly slower than RTs in the monolingual task with non-words that resembled the participants’ less-dominant language being harder to reject.

While some experiments have manipulated language mode by changing the composition of the stimuli, other studies have adjusted experimental settings of the task. Elston-Güttler et al. (2005) modified language mode by showing films with narration in different languages. They found that in an all-L2 sentence task with L2 pre-task priming (a film in the L2), RTs were significantly faster and decision thresholds were raised high enough to

eliminate observable L1 influence on the L2. However, cross-linguistic interference was observed in the other experiment group who had L1 pre-task priming (a film in the L1). More recently, Khachatryan et al. (2016) manipulated the length of stimulus presentation in an L1 semantic priming task. Most of the subjects who saw stimuli presented for a shorter duration were aware of the presence of L2 manipulation, whereas none of the subjects in the other group were aware of this, placing the former group closer to bilingual mode and the latter group closer to monolingual mode. A significant facilitative effect of related word pairs in L2 was found when stimuli presentations were shorter but not when they were longer, indicating that the awareness of covert manipulation of L2 can influence the language mode and consequently what is measured in the laboratory. In short, these experiments suggest that both the selection of stimuli and the experimental contexts have the potential to modulate language activation in reading among bilinguals, and that the level of activation of the non-target language increases as the stimuli involve more words in the irrelevant language or as the setting moves closer to the bilingual context.

CONFLICTING FINDINGS

Although as shown above, several studies have reported on the role of language mode and its influence on language activation, there are contradictory findings. In our opinion, there are at least two possible explanations for these conflicting results. First, language activation may have been artificially induced by the experimental paradigms. Some experiments claim to have provided a “monolingual mode,” which in fact is an intermediate mode in disguise. Since language mode is quite sensitive to a wide range of factors, it takes lengthy efforts to create a purely monolingual environment, and therefore movement along the language mode continuum can be rather easy. For instance, according to previous studies (e.g., Hermans et al., 2011; Khachatryan et al., 2016), the subject’s awareness of the purpose of the study or a small proportion of cognate filler items suffice to activate the non-target language; hence making it arbitrary to assert the non-selectivity of language activation in all modes (see also Costa et al., 2000; Van Hell and Dijkstra, 2002; Duyck et al., 2007). Furthermore, the presence of speakers of the non-target language (e.g., bilingual experimenters or interlocutors with whom participants may come into contact), the language of all instructions, the discussion with or reports from other participants, and even a certain location may all artificially activate the non-target language to some extent, consequently moving bilinguals away from a purely monolingual mode.

Furthermore, research specifically testing participants’ and languages’ characteristics including language dominance, proficiency, and typology can explain some well-controlled yet conflicting experiments. Studies have found that language mode activation may vary when testing a dominant language vs. less-dominant language or when comparing balanced bilinguals to less-proficient bilinguals (e.g., Marian and Spivey, 2003; Lemhöfer and Dijkstra, 2004; Elston-Güttler et al., 2005; Lemhöfer and Radach, 2009; Dunn and Fox Tree, 2014).

Moreover, variability in bilingual proficiency remains one of the main elements modulating non-target language activation and of the network responsible for language control (Green, 2011). According to Abutalebi and Green (2007), cross-language competition is greater among less-proficient bilinguals compared to highly proficient bilinguals which explains why in a pure monolingual mode (Colomé and Miozzo, 2010), the non-target language is invariably activated. In addition, multilinguals whose languages widely differ at lexical, grammatical, or phonological levels showed smaller interference effects as other multilinguals (van Heuven et al., 2011; Boukadi et al., 2015).

Green and Wei (2014) offer a similar account to speech planning and the cognitive processes involved in speech production, particularly in cases of code-switching. From a competitive account, Green and Wei (2014, p. 509) argue the importance of understanding “the interactional contexts of the bilingual speaker.” Bilinguals utilize processes that are most appropriate to certain situations and when they find themselves code-switching, these switches are “coordinated cooperatively and operate in a coupled or in an open-control mode. The former permits alternations and insertions whereas the latter is required for dense code-switching” (p. 499). For our purposes here, Green and Wei’s (2014) work implies that certain situations of multiple language use such as code-switching entail unique demands on control mechanisms and we could hypothesize the same for the unique demands needed as determined by many factors, including language mode.

To have a fully monolingual mode, it seems best to recruit both monolingual and bilingual participants so that the purpose of studying a topic related to bilingualism would not be revealed to the participants. Besides, during the experiment, the purpose of the study should always remain unknown (although it may be inadvertently disclosed after the critical experiments when asking about things like language proficiency or background). Alternatively, researchers can design several experiments to shift the participants’ focus away from the study’s purpose or they can invent a fictitious purpose as to prevent any activation of the irrelevant language. Ideally, participants should be recruited who have not academic knowledge of language selectivity, bilingualism, or language activation. During the experiment, all the experimental settings should be controlled carefully. For instance, the environment of the study (such as the language of the keyboard or computer system, posters on the wall, or any visible written words) should be strictly controlled. The experimenter should be highly proficient in the target language, preferably an L1 speaker and all experimental instructions should be given in that language as well. In addition, all materials (both visual and audio) for the study should be in the target language. The stimuli involved in lexical decision tasks should avoid any homographs or cognates and written words can be replaced with simple drawings in picture-naming tasks. In this

regard, it is easier for researchers who work on two typologically different languages to simulate a more monolingual experimental setting, but the language competition between two different languages may be much weaker than that between two similar languages.

Taking language dominance and proficiency into consideration, it might be ideal to have a matched-pair design in order to make reliable comparisons with the bilingual mode. Consequently, the ideal location would be a place where two languages are equally used and the community attitude toward bilingualism should be positive. Ideally participants in monolingual and bilingual mode groups should be matched on their language proficiency in both languages, especially in L2. This can easily be done post-experiment by conducting a series of standardized tests on listening, writing, speaking, or reading abilities.

CONCLUDING REMARKS

In line with Festman and Schwieter (2015) who argue that bilingual language control and activation should be studied using methods that include both mixed- and single-language experimental blocks, we would like to underscore here the importance of language mode as a confounding variable in studies looking at bilingual language activation and consequently, its implication for the cognitive benefits of bilingualism. Language mode is an important variable that modulates language activation. Simulating different points of the language mode continuum will elicit different results in studies of bilingual speech perception, production, and reading. It appears as though the more monolingual the language mode is, the more likely bilinguals will perform selective language processing. Consequently, language mode modulates language activation and alters the bilingual experience accordingly. However, language activation is also modulated by the interplay of several variables including task and participant characteristics making it challenging to create a pure monolingual mode in which selective language processing may occur. Language mode should be invariably considered as a potential and possible influence on multilingual experience. Given the importance and timeliness of this issue, future studies should specifically test the role that language mode plays in the bilingual experience and the modulating effects it may have on the cognitive benefits associated with bilingualism.

AUTHOR CONTRIBUTIONS

ZY and JS have contributed equally to the development of this paper.

REFERENCES

- Abutalebi, J., and Green, D. (2007). Bilingual language production: the neurocognition of language representation and control. *J. Neurolinguistics* 20, 242–275. doi: 10.1016/j.jneuroling.2006.10.003
- Barac, R., Bialystok, E., Castro, D., and Sanchez, M. (2014). The cognitive development of young dual language learners: a critical review. *Early Child. Res. Q.* 29, 699–714. doi: 10.1016/j.jecresq.2014.02.003
- Bialystok, E. (2015). Bilingualism and the development of executive function: the role of attention. *Child Dev. Perspect.* 9, 117–121. doi: 10.1111/cdep.12116

- Bialystok, E., Craik, F., and Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends Cogn. Sci.* 16, 240–250. doi: 10.1016/j.tics.2012.03.001
- Blom, E., Boerma, T., Bosma, E., Cornips, L., and Everaert, E. (2017). Cognitive advantages of bilingual children in different sociolinguistic contexts. *Front. Psychol.* 8:552. doi: 10.3389/fpsyg.2017.00552
- Boukadi, M., Davies, R., and Wilson, M. (2015). Bilingual lexical selection as a dynamic process: evidence from Arabic-French bilinguals. *Can. J. Exp. Psychol.* 69, 297–313. doi: 10.1037/cep0000063
- Canseco-Gonzalez, E., Brehm, L., Brick, C., Brown-Schmidt, S., Fischer, K., and Wagner, K. (2010). Carpet or Cárcel: the effect of age of acquisition and language mode on bilingual lexical access. *Lang. Cogn. Process.* 25, 669–705. doi: 10.1080/01690960903474912
- Colomé, À (2001). Lexical activation in bilinguals' speech production: language-specific or language-independent? *J. Mem. Lang.* 45, 721–736. doi: 10.1006/jmla.2001.2793
- Colomé, À, and Miozzo, M. (2010). Which words are activated during bilingual word production? *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 96–109. doi: 10.1037/a0017677
- Costa, A., Caramazza, A., and Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1283–1296. doi: 10.1037/0278-7393.26.5.1283
- De Groot, A., Delmaar, P., and Lupker, S. (2000). The processing of interlingual homographs in translation recognition and lexical decision: support for non-selective access to bilingual memory. *Q. J. Exp. Psychol.* 53, 397–428. doi: 10.1080/713755891
- Dijkstra, T. (2005). "Bilingual visual word recognition and lexical access," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. Kroll and A. de Groot (New York, NY: Oxford University Press), 179–201.
- Dijkstra, T., De Bruijn, E., Schriefers, H., and Ten Brinke, S. (2000). More on interlingual homograph recognition: language intermixing versus explicitness of instruction. *Biling. Lang. Cogn.* 3, 69–78. doi: 10.1017/S136672890000146
- Dunn, A., and Fox Tree, J. (2014). More on language mode. *Int. J. Biling.* 18, 605–613. doi: 10.1177/1367006912454509
- Duyck, W., Van Assche, E., Drieghe, D., and Hartsuiker, R. (2007). Visual word recognition by bilinguals in a sentence context: evidence for nonselective lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 663–679. doi: 10.1037/0278-7393.33.4.663
- Elston-Güttler, K., Gunter, T., and Kotz, S. (2005). Zooming into L2: global language context and adjustment affect processing of interlingual homographs in sentences. *Cogn. Brain Res.* 25, 57–70. doi: 10.1016/j.cogbrainres.2005.04.007
- Festman, J., and Schwieter, J. W. (2015). "Behavioural measures of language control: production and comprehension," in *The Cambridge Handbook of Bilingual Processing*, ed. J. W. Schwieter (Cambridge: Cambridge University Press), 527–547. doi: 10.1017/CBO9781107447257.023
- Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81. doi: 10.1017/S1366728998000133
- Green, D. (2011). Language control in different contexts: the behavioral ecology of bilingual speakers. *Front. Psychol.* 2:103. doi: 10.3389/fpsyg.2011.00103
- Green, D., and Abutalebi, J. (2013). Language control in bilinguals: the adaptive control hypothesis. *J. Cogn. Psychol.* 25, 515–530. doi: 10.1080/20445911.2013.796377
- Green, D. W., and Wei, L. (2014). A control process model of code-switching. *Lang. Cogn. Neurosci.* 29, 499–511. doi: 10.1080/23273798.2014.882515
- Grosjean, E. (1998). Studying bilinguals: methodological and conceptual issues. *Biling. Lang. Cogn.* 1, 131–149. doi: 10.1017/S136672899800025X
- Grosjean, F. (2001). "The bilinguals' language modes" in *One Mind, Two Languages: Bilingual Language Processing*, ed. J. Nicol (Oxford: Blackwell), 1–22.
- Hermans, D., Ormel, E., van Besselaar, R., and van Hell, J. (2011). Lexical activation in bilinguals' speech production is dynamic: how language ambiguous words can affect crosslanguage activation. *Lang. Cogn. Process.* 26, 1687–1709. doi: 10.1080/01690965.2010.530411
- Hilchey, M., and Klein, R. (2011). Are there bilingual advantages on non-linguistic interference tasks? Implications for the plasticity of executive control processes. *Psychon. Bull. Rev.* 18, 625–658. doi: 10.1075/lab.15041.fri
- Incera, S., and McLennan, C. (2018). The time course of within and between-language interference in bilinguals. *Int. J. Biling.* 22, 88–99. doi: 10.1177/1367006916644688
- Jared, D., and Kroll, J. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *J. Mem. Lang.* 44, 2–31. doi: 10.1006/jmla.2000.2747
- Khachatryan, E., Camarrone, F., Fias, W., and Van Hulle, M. (2016). ERP response unveils effect of second language manipulation on first language processing. *PLoS One* 11:e0167194. doi: 10.1371/journal.pone.0167194
- Lemhöfer, K., and Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: effects of code similarity in language-specific and generalized lexical decision. *Mem. Cogn.* 32, 533–550. doi: 10.3758/BF03195845
- Lemhöfer, K., and Radach, R. (2009). Task context effects in bilingual nonword processing. *Exp. Psychol.* 56, 41–47. doi: 10.1027/1618-3169.56.1.41
- Luk, G., and Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *J. Cogn. Psychol.* 25, 605–621. doi: 10.1080/20445911.2013.795574
- Marian, V., and Spivey, M. (2003). Competing activation in bilingual language processing: within-and between-language competition. *Biling. Lang. Cogn.* 6, 97–115. doi: 10.1017/S1366728903001068
- Paap, K., Johnson, H., and Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex* 69, 265–278. doi: 10.1016/j.cortex.2015.04.014
- Prior, A., and Gollan, T. (2011). Good language switchers are good task-switchers: evidence from Spanish-English and Mandarin-English bilinguals. *J. Int. Neuropsychol. Soc.* 17, 682–691. doi: 10.1017/S1355617711000580
- Schwieter, J. W., and Ferreira, A. (2016). "Effects of cognitive control, lexical robustness, and frequency of codeswitching on language switching," in *Cognitive Control and Consequences of Multilingualism*, ed. J. W. Schwieter (Amsterdam: Benjamins), 193–216.
- Soares, C., and Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: the effect on lexical access. *Mem. Cogn.* 12, 380–386. doi: 10.3758/BF03198298
- Spivey, M., and Marian, V. (1999). Cross talk between native and second languages: partial activation of an irrelevant lexicon. *Psychol. Sci.* 10, 281–284. doi: 10.1111/1467-9280.00151
- Van Hell, J., and Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychon. Bull. Rev.* 9, 780–789. doi: 10.3758/BF03196335
- van Heuven, W., Conklin, K., Coderre, E., Guo, T., and Dijkstra, T. (2011). The influence of cross-language similarity on within-and between-language Stroop effects in trilinguals. *Front. Psychol.* 2:374. doi: 10.3389/fpsyg.2011.00374

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Yu and Schwieter. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership