Teaching Grapheme-Phoneme Correspondences using A Direct Mapping Approach for

At-risk Second Language Learners: A Randomized Controlled Trial

Abstract

Reading interventions developed to teach Grapheme-Phoneme Correspondences (GPCs) were evaluated among L2 at-risk readers. In the Direct Mapping of Graphemes (DMG) condition, children's attention was explicitly drawn to the application of a graphemes taught on that day to shared reading of words in authentic text. In the control condition there was no such systematic daily linkage of the GPCs and texts. The two reading interventions were otherwise identical. Two hundred and fifty-three Chinese Grade 1 and Grade 2 students were screened and those who scored in the bottom 30% of an English word-reading test were identified as L2 at-risk readers. Seventy-one L2 at-risk readers were thus randomly assigned to two conditions, both of which were small group reading interventions: 1) DMG or 2) taught control. We hypothesized a significant main effect of Intervention condition and significant interaction of Intervention by Phonological Awareness (PA) effects on word reading, word attack, spelling and sentence comprehension favoring the DMG intervention. Results showed that predicted interaction effects were significant for word reading, spelling and sentence comprehension. No other effects were significant. Results suggest that the daily Direct Mapping of taught GPCs to shared book reading promotes reading development in at-risk English L2 readers with stronger phonological skills.

Teaching Grapheme-Phoneme Correspondences using a Direct Mapping Approach

for At-risk Second Language Learners: A Randomized Controlled Trial

The global population of learners of English-as-a-second-language (ESL) is growing rapidly (Calderón, Slavin, & Sánchez, 2011). It is estimated that around 375 million people around the world are currently ESLs (Beare, 2017). A substantial proportion of these children may be at-risk for future reading difficulties at school entry (Swanson, Kudo, & Guzman-Orth, 2016). Innovative and evidence-based approaches to help these children who are at-risk of second language (L2) reading difficulties are thus needed to prevent future reading failure worldwide (Cirino et al., 2009; Gerber et al., 2004). In this line of study, well-evidenced reading intervention approaches for L1 learners have been cautiously applied to L2 populations (Linan-Thompson, Vaughn, Hickman-Davis, & Kouzekanani, 2003). Phonics approaches have received strong support from empirical research in helping at-risk L1 learners and L2 learners (Vadasy & Sanders, 2012). Reading researchers have further examined the differential contribution of key elements in phonics instruction, namely teaching of Grapheme-Phoneme Correspondences (GPCs) and Phonological Awareness (PA; Hatcher, Hulme & Snowling, 2004; Hulme, Bowyer-Crane, Carroll, Duff, & Snowling, 2012). Some recent studies have focused on the teaching of GPCs (Chen & Savage, 2014; Savage, Georgiou, Parrila, & Maiorino, 2018). The present study focused on the teaching of GPC to L2 learners, which, to our best knowledge, has not been investigated in previous work.

The current study built on a Randomized Controlled Trial (RCT) involving L1 learners of English conducted by Savage et al. (2018), which evaluated a novel and effective approach of preventative reading intervention for first language (L1) learners of

English. In their large-scale dual-site-matched control trial intervention study, a reading intervention teaching Direct Mapping of Grapheme (DMG) and Set-for-Variability (SfV) resulted in significant impacts on standardized word reading and spelling measures at immediate post-test and in word reading and sentence comprehension skills at a delayed post-test some 5 months after the intervention had finished. Their reported approach involved two processes (DMG and SfV) driving literacy gains. DMG in the context of reading intervention refers to explicitly and systematically applying the taught knowledge of GPC *of that day* to decode authentic text whereas SfV involves having a flexible mental process in matching pronunciations derived from decoding based on GPC knowledge to oral vocabularies that have similar pronunciations. The present study focused on the effect of DMG but not SfV for two reasons: (1) SfV requires a high level of lexical knowledge which may not be applicable to the target group of participants in this study who are young ESL learners from a low English immersion context and (2) Reading intervention with two teaching strategies precludes the identification of the specific effect of a teaching principle. To tease out the specific effects of DMG from SfV, and generalize such approaches to an L2 context, we designed a GPC teaching intervention based on the principle of DMG for Hong Kong Chinese ESL students in Grade 1 and Grade 2 and evaluated its impact on L2 literacy using RCT. We first describe the relevant research on what is known of effective practices of GPC teaching in L1 and L2 and then describe a CONSORT (Consolidated Standard of Reporting Trials)-quality RCT study that evaluates the impact of DMG in an L2 context.

Teaching Grapheme-Phoneme Correspondences (GPC) to At-risk Learners

Phonics instruction typically involves teaching GPCs (where graphemes are letters or letter clusters that correspond to phonemes) and the blending of those GPCs to derive word pronunciations. Phonics instruction varies in terms of instructional methods, but all involve the explicit teaching of pre-specified sequence of GPCs followed with various segmentation and blending activities which requires phonological awareness skills (De Graaff, Bosman, Hasselman, & Verhoeven, 2009).

Such systematic phonics approaches have been shown to be effective in improving literacy for both average and at-risk readers (De Graaff et al., 2009; Savage & Cloutier, 2017; Suggate, 2016). Emerging evidence has been reported that phonics is also effective in enhancing English literacy for L2 learners (Jamaludin et al., 2016; Li, Tao, Joshi, & Xu, 2018; Stuart, 2004; Vadasy & Sanders, 2010). However, fewer studies have involved students with L2 reading difficulties. Recently, Li et al. (2018) report the positive effects of explicit phonological awareness instruction coupled with teaching of letter knowledge (letter name and letter sound) for struggling ESL readers for L2 learners in China. We argue that intervention effects may however vary across diverse groups of learners from different socioeducational contexts and L2 proficiency level. Thus, more intervention research is needed to inform theory and practice in ESL reading difficulties.

One current line of research seeks to examine the components of distinct phonics programs that work best for at-risk readers, with an eye to optimizing practice (Graham et al., 2017; Henbest & Apel, 2017). While there exists evidence that intervention to teach phonological awareness (PA) is effective for young L2 readers with or without reading difficulties (Lesaux & Siegel, 2003; Yeung, Siegel, & Chan, 2013), much less research attention has been devoted to examine the specific effects of GPC teaching in promoting

literacy among ESL readers. Chen and Savage (2014) examined the effects of teaching the most frequently occurring complex GPCs (e.g., 'tch') on reading and reading motivation for at-risk Grade 1 and Grade 2 L1 children. The GPCs were selected based on an analysis of children texts in storybooks by Vousden, Ellefson, Solity, and Chater (2011). The results showed that the teaching of commonly encountered complex GPCs benefited at-risk children's spelling, word reading and reading motivation with large effect sizes of intervention. More recently, Savage et al. (2019) tested the teaching of complex GPCs to Grade 2 L1 at-risk children from Canada and showed advantages of GPC instruction adopting the simplicity principles over the control intervention for word reading, pseudoword reading, and reading comprehension moderated by pretest PA skills. To what extent this can be applicable to at-risk L2 children remains unanswered in the literature. Given the reading level of young at-risk L2 learners, the present study evaluated the teaching of simple GPCs (e.g., /t/) to Chinese ESL children on their word reading-related skills. A recent meta-analysis by Hall and Burns (2018) shows that small group phonologically based reading interventions positively impact a range of literacy outcomes including word recognition, word attack and spelling. In addition, interventions that involve shared book reading has shown positive impacts on vocabulary and listening comprehension (Dowdall et al., 2019; Neuman & Kaefer, 2018). Based on this prior work, we have included measures of word recognition, word attack, spelling, listening comprehension and vocabulary to evaluate our interventions.

Use of Storybook and Direct Mapping in teaching of GPCs

The teaching of GPCs and shared book reading are commonly co-occurring practices in many effective reading interventions (Shapiro & Solity, 2008, 2016). Prior

studies have demonstrated positive evidence of using storybooks in phonics teaching (Jenkins, Peyton, Sanders, & Vadasy, 2004; Tse & Nicholson, 2014). For example, Tse and Nicholson (2014) reported that using storybooks in phonics teaching resulted in greater gains in literacy skills of Grade 2 L1 learners of English than phonics teaching without storybook reading.

How storybooks are best utilized in the context of GPC teaching remains relatively under-explored. Savage et al. (2018) attempted to address this issue by introducing a novel approach to GPC teaching. They argued that the principle of DMG: applying the knowledge of a GPC *taught on that day* to decode real text, would be effective in enhancing literacy among at-risk readers. During the shared book reading, students' attention is drawn to the words with the GPCs taught on that day and students are asked to identify the grapheme and encouraged to read the words. Savage et al. (2018) were able to demonstrate that alongside a strategy for encouraging mental flexibility in using phonics, this approach was more effective than the current best practice of reading intervention (phonics and sight word teaching) in promoting literacy skills including word reading, spelling and sentence comprehension in at-risk L1 readers.

The DMG principle is grounded in the 'phonological linkage hypothesis', which proposes that graphemic, and phonemic skills and text reading experience should be linked to attain reading fluency (Hatcher et al., 1994, 2004, 2006). For children with English reading difficulties, prior intervention studies have shown that teaching of PA and GPC knowledge resulted in a promising improvement in nonword reading but less impact on real word reading (Compton, Miller, Elleman, & Steacy, 2014). We argue that some current reading interventions may not fully realize the connections among

graphemes, phonemes and authentic text reading suggested by the 'phonological linkage hypothesis'. The DMG principle addresses this issue. The formal reports of many existing phonics interventions do not reveal an emphasis on the linkage between the GPCs taught *on that day* and authentic text exposure. This practice was thus tested in the present intervention.

During the process of DMG, children need to use their PA skills to apply the knowledge of GPCs to decode the words. Conceivably, children with better phonological skills may be more able to apply the knowledge to decode the words. Previous interventions teaching PA and GPCs have strongly suggested that PA mediates the immediate and delayed intervention effects (Hulme et al., 2012). Therefore, in the present study, we will also evaluate the interaction of PA with GPC intervention on literacy skill. As DMG takes place in the context of shared book reading, we will also test the interaction effects of vocabulary and intervention on outcomes because oral vocabulary predicts reading performance (Dickinson, Golinkoff, & Hirsh-Pasek, 2010; Liu, Yeung, Lin, & Wong, 2017).

In sum, Savage et al., (2018) was the first study to explicitly test the DMG approach. However DMG was coupled with other strategies for at-risk L1 learners so it is currently unclear whether reading gains can be attributed to the DMG aspect of instruction specifically. Savage et al., (2019) report the effects of DMG training in L1 at-risk learners. There however exists no prior study that has specifically examined the use of DMG in teaching GPC for at-risk learners with English as a second language.

The Present Study

Past research involving L2 students have suggested positive effects of phonics instruction on reading and spelling in L2 (Li et al., 2018; Vadasy & Sanders, 2012). However, no study has examined the effects of shared book reading integrated with GPC instruction on L2 literacy skills among ESL students. Given more than 20% L2 children are identified as at-risk in various L2 learning contexts (Swanson et al., 2016; Yeung, 2018), research in preventative reading intervention for L2 learners is needed but currently very limited. The aim of the present study is to test whether the principle of DMG in GPC instruction is effective to enhance L2 literacy for Chinese students who struggle to learn to read English. Thus, we designed two reading interventions, with a single contrast in explicitly directing students to apply the taught GPC of the day to read authentic text in shared book reading activities. The intervention effects were examined in Grade 1 and Grade 2 readers because they started to receive formal reading instruction in Grade 1 and early identification and intervention are important for prevention of future reading failure (Lesaux & Siegel, 2003). We hypothesized that a GPC intervention with DMG will be more effective in improving literacy than a comparable reading intervention with typical GPC instruction but without systematic daily linkage between taught GPCs and texts. This hypothesis is based on the findings from Savage et al. (2018) on L1 at-risk learners and other studies (Jenkins et al., 2004; Tse & Nicholson, 2014) showing the positive impacts of phonics instruction with a linkage between storybook reading and GPC teaching. More specifically, Savage and colleagues' intervention of DMG has shown positive effects on word reading (real word and pseudoword) and spelling, so we also expected our DMG intervention would be more effective than non-DMG intervention in improving these literacy skills. To test intervention specificity, we also

examined the effects of the interventions on listening comprehension, vocabulary and PA. In other words, we hypothesized significant main effects of the DMG intervention on real word reading, pseudoword reading and spelling and non-significant main effects of the DMG intervention on listening comprehension, vocabulary and PA.

Past research has also suggested that PA mediates the effects of phonics instruction (e.g. Hulme et al., 2012), and moderates the effects of GPC instruction (Savage et al., 2019), because GPC knowledge and PA are dual foundations underpin decoding which is generative in alphabetic orthographies with self-teaching opportunities (Compton et al., 2014). Therefore, children with better PA should be more able to utilize the taught GPC knowledge in decoding authentic text. As the present intervention did not directly teach PA skills but only teaches GPCs, we hypothesized that children with better pre-test PA in the GPC intervention with DMG will improve more than those with weaker PA skill. An interaction effect of pre-test PA and intervention condition was thus predicted when pre-test of the literacy skill is controlled, reflecting a hypothesized *moderating* effect of PA on all literacy outcomes in the DMG condition. To further test the specificity of the moderating effect of PA, we would also examine the interaction effect of vocabulary and intervention on outcomes. This allows us to examine whether more verbally capable participants would benefit more in the DMG intervention than the control intervention in shared book reading activities. The testing of the two moderators in this study allowed us to test whether the effects of DMG are specifically tied to PA skills or related to more general language ability. Given existing evidence above on DMG and other evidence of positive impacts of phonology-based over vocabulary-based intervention on word reading (Bowyer-Crane et al., 2008), we predicted that the

vocabulary-as-moderator effect would not be a significant predictor of word reading. On the other hand, given the strong evidence of a role for vocabulary as a predictor of linguistic comprehension (e.g. Hjetland, Brinchmann, Scherer, & Melby-Lervåg, 2017), we did predict a specific effect of vocabulary on our comprehension outcome. In summary, we hypothesized a significant PA-as-a-moderator effect on word reading (real word and pseudo word) and spelling as well as a significant vocabulary-as-a-moderator effect on listening comprehension. In addition, we also hypothesized a non-significant vocabulary-as-moderator effects on all other literacy outcomes except listening comprehension.

Method

Research Design

The present study was a Randomized Control Trial that took place in the academic year 2017-2018 and used a pre-post-test experimental intervention design. Ethics approval was first obtained from the university ethics committee of the first author's institution. Informed consent was sought from parents of all participants. Participating students ($n = 71$) were randomized within class and thus teacher-effect on students' improvement on literacy during the intervention period was controlled. Stratified randomization with pretest reading as baseline was used. Students within the class were assigned into blocks based on the word reading at pretest and simple randomization was performed by flipping a coin to determine group membership.

The meta-analytic study by Suggate (2010) reported a moderate effect size (ES = .49) of reading intervention for children at-risk for reading difficulties. Similarly, Vadasy and Sanders (2012) found an effect size of 0.43 of phonics-based intervention for L2 children. Based on these previous studies, we predicted an effect size for intervention

of 0.40. Estimation of required sample size to detect this effect size was conducted using G-power (Faul, Erdfelder, Buchner, & Lang, 2009). The analysis showed that 64 students were required to attain conventional alpha, 2 covariates and power of .80 to detect an effect size of 0.40. All students were provided with small-group supplemental intervention by trained experimenters for 8 weeks, with three 30-minute sessions in a week, after class time in their attending school. Figure 1 shows the CONSORT flow diagram of all participants in the study.

Participants

Two hundred and fifty-three Grade 1 and Grade 2 Chinese ESL students from a local primary school in Hong Kong were screened using a word-reading test that combined a locally developed reading test and a standardized reading test from U.S. (Woodcock-Johnson III Test of Achievement) and those scored in the bottom 30% (*n* =76) were identified as at-risk readers. Out of the 76 students identified, 71 parents of the identified students consented to participate in the intervention study. The school where we recruited the participants used Chinese as the medium of instruction. All children from 10 classrooms received English instruction of around 3 to 4 hours in week as a subject in school. All pupils in the school were taught in Cantonese (the spoken language of Hong Kong Chinese) and all the students were Cantonese-speakers. These students used Cantonese as their daily communicative language. The school was government-subsidized and located in a low-income neighborhood. All parents of the participants were ethnic Chinese. There were 40 boys and 31 girls with a mean age of 82.25 months (range 72 – 91 months) at the commencement of the study. Information on maternal education levels of intervention participants was collected from a parent questionnaire.

English Language and Literacy Measures

The following measures were administered at both pretest and posttest by research assistants who had received a 2-hour training delivered by the first author before the commencement of the study. Given the English proficiency of the participants, test instructions were given in Cantonese, the spoken language of the participants, and items were presented in English.

Word reading. There is no validated measure of English word reading for the current population, children' so English word reading was assessed by the combination of two tasks, the Woodcock-Johnson III Test of Achievement Subtest (Woodcock, McGrew, & Mather, 2001; 2007) and the English word reading task developed by McBride-Chang and Kail (2002), who found that standardized reading tests developed in U.S. or England were not adequate to reflect children's reading ability due to differences in school curriculum. As the word reading task was both a screening task and outcome task, we combined a standardized reading test and a locally developed reading test so as to maximize the adequacy of the measurement. The first thirty words were selected from the Woodcock subtest, whereas the McBride-Chang and Kail's (2002) task contained thirty words that were commonly found in local kindergarten textbooks. A previous study has used the same test to identify Hong Kong Chinese students at-risk of English reading difficulties (Yeung, 2018). There are two words common in the two tasks and thus, in total, fifty-eight printed words were presented to the children, and they were asked to pronounce the words one-by-one. Each correct pronunciation was given one point, which added up to a maximum score of 58. Internal reliabilities assessed using Cronbach's alpha for this 58-item test were 0.88 for pretest and 0.91 for posttest.

Word attack. The word attack sub-task from the Woodcock Johnson III Test of Achievement form B (Woodcock, McGrew, & Mather, 2001; 2007) was administered to measure children's ability to decode pseudowords (e.g., "*nan*", "*rox*", "*tiff*"). All pronunciations were recorded and cross-marked by two trained experimenters and the agreement on scores given was 92%. In total, the first 10 items from the word attack task were used. These words increase in complexity. Prior to the test items, two practice items were given with feedback to ensure that children understand the purpose of the task. Due to low ability on pseudoword reading of participants at pretest, we scored individual phonemes correct within each pseudoword in addition to the total number of pseudowords read (The total number of psuedowords read at pretest for all participants was 0). There were 34 phonemes in the 10 pseudowords. The maximum score for total number of psuedoword read was 10 and the maximum score of the individual phonemes read was 34. Cronbach's alpha for reliabilities for the phoneme-level scores at pretest was 0.70 and that for the word-level score at posttest was 0.72.

Phonemic awareness. The phoneme elision task from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) was administered to measure children's phonological awareness. Children were presented with a list of words and required to say each word with a phoneme deleted. The deleted phoneme was always a consonant but varied in position of the words, ranging from the initial sound to the sound from the second or third syllables, to the final sound. For example, children had to say "bold" without saying /b/ with the correct production of "old"; to say "time" without saying /m/ with the correct production of "tie"; to say "cap" without say /p/ with the correct production of "ca". There were 2 practice items and 10

test items. Each correct response for the test items was given one point, adding up to a maximum score of 10. Reliabilities of internal consistency as indexed by Cronbach's alphas for pretest and posttest were 0.74 and 0.76 respectively.

Spelling. Children's spelling ability was assessed by a researcher-developed task. The task involved 8 regular and irregular words that were commonly found in local kindergarten textbooks (Yeung, 2006). Following each target word, there was a sample sentence given for children's reference. Children were asked to write down the target words one-by-one. Each correct response scored one point, which added up to a maximum score of 8. Cronbach's alphas for pretest and posttest reliability were 0.65 and 0.69 respectively.

Vocabulary. The Peabody Picture Vocabulary Test-4th edition (Dunn & Dunn, 2007), was administered to measure children's receptive vocabulary. Children were required to point to one of four given pictures that best corresponded to the word pronounced by an experimenter. There were 24 items in total. Each correct response was given one point. The maximum score was thus 24. Cronbach's alphas for pretest and posttest reliability were 0.70 and 0.65 respectively.

Sentence comprehension (Oral). The Sentence Comprehension subtest of the Woodcock Johnson III Test of Achievement subtest (Woodcock, McGrew, & Mather, 2007) was administered to measure children's listening comprehension ability. Children were required to point to one of four given pictures that best corresponded to a sentence presented orally by an experimenter. There were totally 20 items and each correct response was given one point. The maximum score was 20. Cronbach's alphas for pretest and posttest reliabilities were 0.65 and 0.66 respectively.

Non-verbal reasoning. Raven's Progressive Coloured Matrices (Raven, Raven, & Court, 2003) were administered only at pre-test, to serve as a proxy of children' general non-verbal intelligence. Here children were given a visual matrix with one missing part and six possible answers underneath, and asked to select one of the six answers that best matched the matrix. To avoid prolonged testing, only 2 sets were administered (24 test items). Each correct item was worth one score. The maximum score was 24. Reliability of internal consistency of this task indexed by Cronbach's alpha was 0.65.

Interventions

Two small-group intervention programs with detailed lesson plans were developed and implemented by trained experimenters. Interventions were run with groups of three to four children outside of class time for 8 weeks. A total of 12 hours of instruction were provided to these children. Research assistants with a Bachelor's degree in Psychology or Education were trained for at least 6 hours by the first author to run the intervention programs. Weekly or biweekly meetings were held with all the interventionists to discuss the implementation of the interventions. Interventionists for the two conditions met separately and their problems or concerns in running the interventions were addressed in the meetings. The first author also observed the lessons regularly and feedback was given to the interventionists as part of treatment fidelity process.

The two intervention programs (DMG intervention and control intervention) shared the same structure and teaching materials, with the sole exception that in the DMG condition, the GPCs first taught in isolation on any given day were systematically revisited in the shared reading authentic text on the same day, with children encouraged

to read words with taught GPCs in these texts. In the control condition the same GPCs were taught overall as in the DMG condition, but no systematic daily linkages were made between the GPC instruction and the words children were subsequently encouraged to read in shared texts.

Each week, 2 lessons were devoted to learning 4 GPCs and one further lesson was used to review the learnt GPCs of the week. In the first part of each lesson, participants learned GPCs through explicit teaching, and then engaging language games. The goals of the language games were to consolidate the learning of the GPCs and provide multiple exposures to the GPCs. The games also served to encourage participants to apply their GPC knowledge to decode printed words. An example of the language games was to ask students to respond with the correct phoneme associated with a grapheme presented by making body movement such as jumping to the left or right side as instructed. The second part of the lesson was devoted to shared storybook reading in which DMG was applied for the experimental condition and DMG was not present in the control condition. The same storybook was used in each lesson for the two conditions. For the DMG intervention, interventionists conducted shared book reading and read the text where necessary. During text reading, when the taught GPC of the day appeared, students' attention was drawn to the targeted GPC and students were invited to apply their GPC knowledge to decode the text. Feedback was given to children immediately after the application of knowledge of GPC taught on the day. In the control intervention, the same storybooks were read to the children and the text in the storybook was read out by the interventionist, but no application of GPC knowledge learned on that day was present. The interventionist did not stop at all on words in texts in the control condition.

The learning sequence of the GPCs followed the Ranked List of Grapheme-Phoneme Mappings developed by Vousden, Ellefson, Solity and Chater (2011). Vousden et al. (2011) sequenced GPCs based on lists of the most commonly occurring English words in contemporary children's books for young children in the United Kingdom. In the two intervention programs, the first 24 GPCs on the list were covered. All the storybooks chosen for each lesson contain a high density of the graphemes taught in that lesson. For both interventions, differentiation of curriculum was emphasized, meaning that weaker children were given activities or tasks that suited their current learning needs.

Intervention Fidelity

To assess intervention fidelity, both interventions were frequently observed by independent observers. A rubric was created to reflect to what extent (1) the specific content of each intervention session was taught as specified in the lesson plan; (2) the adherence to time spent on each activity as specified in the lesson plan was present and (3) the degree to which the teacher was responsive to students' needs and provided meaningful feedback. Each aspect of program fidelity was assessed with a 3-point scale with 0 indicating not done, 1 indicating partly done and 2 indicating fully done.

Approximately 25% of all of the teaching sessions were observed by two trained research assistants. Analyses of scores from the rubrics by the two observers showed 90% agreement in the experimental condition and 92% agreement in the control condition. Mean scores from each observer were then calculated. These ranged from 1.82 to 1.89 for all aspects of treatment observed, indicating a high overall intervention fidelity. There were also no significant differences between the fidelity of the two conditions ($p > 0.05$).

<p style="text-align:center">Results</p>

Preliminary Data Analyses

No significant deviations from normality of data distributions were detected from preliminary data screening. To evaluate the possible impact of outliers, all scores 2 standard deviations above or below variable means were considered as potential outliers. Analyses reported below were then conducted both with and without candidate outliers and the results were contrasted. There were no significant differences between results of analyses with and without the outliers, so analyses based on the full sample are presented below. There was no missing data in this sample. Mean, standard deviations and within-group effect sizes for all the measures at pretest and posttest are shown in Table 1 and Table 2. Inspection of means suggests, as expected, few overall differences between groups.

Results of Group Matching

Prior to conducting any analyses, we assessed the quality of match achieved across the two intervention conditions on a comprehensive range of candidate measures including pre-test attainment, non-verbal IQ, maternal education and age. To examine whether the two groups were similar before intervention, independent sample *t*-tests were conducted on pretest scores with intervention conditions as the independent variable. Results showed that the two groups were not significantly different on any measures at pretest ($p > 0.10$ in all cases) and thus indicated that the two groups were comparable before intervention. Inspection of these variables shows few differences, and none reached conventional significance, showing good overall matching by condition. Results of analyses are reported in Table 1.

Intervention Results

Anticipating dependencies within classrooms, our data were analyzed with HLM (e.g., Hox, 2010; Raudenbush & Bryk, 2002). Using standard procedures, the final HLM models were built in a stepwise fashion from preliminary analyses. Model 1, an unconditional one-way ANOVA model with random effects first confirmed that there was significant classroom variance at pre-test and post-test on primary achievement measures and spelling beyond the variance solely attributable to students. ICCs ranged between .0 and .50). As most outcome measures showed non-zero ICCs HLM was thus the appropriate analysis strategy to account for this nestedness.

In the final two-level hierarchical model presented here, we examined whether variance on posttest achievement outcome measures (after controlling pre-test classroom-level achievement variance at level 2), was explained by the Condition (DMG versus Control) factor, moderated by phonological awareness skill, at leve1 1. An ANCOVA model was appropriate because the most precise test our central hypothesis concerning the predicted GPC x PA interaction in modeling growth in attainment at post-test requires a control for pre-test reading ability. In addition, controls for nested pre-test attainment improve the power of hierarchical analyses even where the covariate is not statistically significant (e.g., Raudenbush et al., 2011). Equations 1, and 2 describe this final model at the student, and classroom levels, for student i in classroom j, respectively.

1) Equation for Student Level 1 Model:

$$Y_{ij} = \beta_{0j} + \beta_{1j} (\text{INTERVENTION}) + \beta_{2j}*(\text{INTERVENTION x}$$
$$\text{PHONOLOGICAL AWARENESS}) + e_{ij}$$

2) Equations for Classroom Level 2 Model:

$$\beta_{0j} = \gamma_{00} + _{01j}*(\text{PRETEST ATTAINMENT}) + r_{0j,}$$

In all analyses, predictor variables were grand mean-centered. Identical 3-level models were run for all variables with the corresponding pre-test measure as the covariate. Results of the HLM analyses are reported in Table 3 where researcher created dummy-coded DM condition is compared against the zero-coded Control condition.

To balance risk of type-1 error among our multiple outcome measures (MOMs) alongside our a priori identification of primary outcomes on conceptual grounds, where type-2 error risk exists if these are placed among MOMs and adjusted family wise, unadjusted alphas were used only primary outcome measures. Alphas corrected for multiple contrasts were used on all other secondary measures. The DMG intervention involves teaching GPC correspondences so that children can read the many words containing these correspondences. Reflecting the content of our intervention, our primary outcome measures were Word Reading and Word Attack, so unadjusted alpha was applied for these two measures. Other measures of Spelling, and Sentence Comprehension were secondary outcomes as they did not directly reflect the content of our intervention and adjustments to alphas were made for the total number of secondary outcomes ($\alpha = .05/ 2 = .025$).

The results showed that there was no significant main effect of Intervention for any measure. There was, however, a significant Intervention x Phonological Awareness interaction effect for Word Reading and Spelling at post-test ($p < .01$, and $p < .025$ respectively), as well as for Sentence Comprehension ($p < .025$). In each case, this interaction favored DMG over Control but only for the children with stronger

21

phonological awareness ability at pre-test. No other effects reached significance. As an additional check on the specificity of predicted PA x Intervention interactions we also ran Vocabulary x Intervention interaction analyses. These were non-significant except for the sentence comprehension measure ($p < .05$).

Discussion

The present study sought to evaluate the impact of teaching GPCs using the direct mapping of graphemes (DMG) to L1 Chinese children who were English second language learners. As far as we are aware, there exists no relevant data from such English L2 learners on this question (Yeung & Chan, 2013; Yeung et al., 2013). We hypothesized that GPC teaching with DMG would be more effective than without DMG and predicted pre-test PA would be a moderator of reading outcomes given that GPC and PA are joint foundation skills of reading (Hulme et al., 2012). Our data showed no overall main effects of the DMG condition over controls, but supported the 'PA-as-moderator' hypothesis. Previous research (Hulme et al., 2012) has demonstrated that phonological intervention has stronger impacts on reading outcomes when children had stronger PA. Existing theory and research evidence from L1 learners of English also points to the conjoint influence of PA and GPCs on reading development (e.g. Hulme et al., 2012). We have provided evidence consistent with such conjoint effect in L2 learners of English with our DMG intervention. Furthermore, we were able to find such a 'PA as a moderator' effect when we controlled for pre-test reading skill. These findings suggest that the effects of the DMG intervention on reading and spelling outcomes were highly specific to children who were more phonologically able. It is possible that the process of DMG required PA as a supporting skill when children's attention was explicitly drawn to the pronunciation of GPCs.

Interestingly, children with better PA skills in the DMG condition also benefitted more on a sentence comprehension measure (listening comprehension), possibly because of the significant correlation between PA and listening comprehension (Lepola, Lynch, Laakkonen, Silven, & Niemi, 2012; Senechal, Ouellette, & Rodney, 2006) or possibly because shared book reading with DMG indirectly improved L2 listening comprehension skills. Practically, the current findings suggest that shared book reading with DMG as a teaching strategy should be included in phonology-based reading intervention for ESL children who are at-risk of reading difficulties. During intervention, attention should be taught graphemes systematically linked to reading authentic texts, thereby allowing the students to also apply their phonological skills in reading real text.

The findings of the current study partially support the only previous studies of DMG by Chen and Savage, 2014 and Savage et al. (2018, 2019). Savage and colleagues (2018) have demonstrated both DMG and SfV supported literacy development among L1 at-risk children and Savage et al., (2019) showed moderating effects of PA on the impact of DMG intervention alone in L1 at-risk students. This study provides positive evidence of DMG alone in teaching of GPCs for children with better PA skill and the applicability of this approach to L2 children. However, we were not able to demonstrate the same effects on word attack as Savage and colleagues' study. It may be due to the difficulty of engaging in phonological decoding among Chinese learners whose L1 is logographic (Uchikoshi & Marinova-Todd, 2012). Our 8-week DMG intervention focused only on a limited number of GPCs and thus might not be intensive and comprehensive enough to induce positive gain in word attack. As the present study is the first examining the specific effects of the

teaching principle of DMG in L2, further research is needed to establish the generalizability of current findings to other ESL groups with varied proficiency levels.

Contrary to our predictions, we were not able to find a main effect of the DMG intervention on reading outcomes, although small effect sizes favoring DMG over controls across measures were evident. Considering the educational context, Hong Kong ESL children are in a low-immersion environment and their phonological skills are generally lower than ESL children from high-immersion environment (McBride-Chang, Bialystok, Chong, & Li, 2004; Yeong, Fletcher, & Bayliss, 2017). Therefore, well-designed interventions such as the DMG and control programs with much shared content may both produce comparable overall growth in reading outcomes. Nevertheless, nuanced advantages appear to be evident for children with stronger phonological abilities at pre-test.

The present study suggests that the combined contribution of GPC knowledge and PA to reading development demonstrated in L1 learners (Hulme et al., 2012; Schaars, Segers, & Verhoeven, 2017) also applies to L2 learners. Previous studies have addressed the causal role of GPC knowledge and PA through intervention study and mediation analysis (Hulme et al., 2012). Previous studies have also demonstrated positive impacts of PA intervention alone on ESL children's word reading (Yeung et al., 2013). Our PA-as-moderator effect has addressed the causal links of components in phonics instruction and reading in L2 populations. Our DMG intervention embodies a daily directly-mapped connection forming between GPC knowledge and PA and the consolidation of such a connections in the context of authentic text presented in engaging and age-appropriate reading materials, and as such offers support to the "phonological linkage hypothesis" (Hatcher et al., 1994, 2004, 2006). This approach to supporting reading development

among at-risk readers holds the promise that, alongside other aspects of effective instruction, that it can be generative of the richer orthographic lexicon associated with reading development in typical readers (e.g. Compton et al., 2014). More extended and multi-componential interventions are however needed to explore this claim.

Finally, our results also speak to the specificity of predicted interaction effects as while predicted interactions of PA and Intervention were evident on both primary and secondary outcome measures, we also tested the interaction effects of Vocabulary and Intervention. These were generally not significant for word-level outcomes, and thus suggest that it was not just the generally more verbally able children who benefited from the DMG approach. Rather, advantages were closely tied to phonological abilities more specifically. There was however an interaction effect of PPVT and intervention for listening comprehension. Such a pattern is predictable based on any view that shared book reading activities are beneficial to oral language development (Mol & Bus, 2011; Dowdall et al., 2019). The interaction effects further showed that our intervention benefited L2 children with better initial vocabulary more than children with weaker initial vocabulary. This interaction effect suggests that the DMG intervention also served to enhance listening comprehension skill of some L2 learners.

Limitations and future work

Several limitations of the present study should be noted. First, the identification of at-risk readers in the present study was not ideal. We have used a cut-off criteria of below the 30th percentile might include L2 at-risk English readers who were not at-risk. Recent studies have used other identification methods such as growth curve models (Yeung, 2018) or latent profile analysis (O'Connor, Geva, & Koh, 2019) to accurately identify at-risk ESL

children. Future research should adopt these identification methods and examine the effects of DMG interventions. Furthermore, the current sample was drawn from one school only. A more diverse sample and larger replication are needed to warrant the generalizability of the current findings. Second, given the low proficiency in pseudoword reading, we were forced to use psuedoword reading coded at an individual phoneme accuracy level as a pre-test score in the analysis of word attack and this also preluded us from testing the work attack as a moderator of intervention effects. Therefore, this may be an imperfect model in examining intervention effects of DMG intervention on decoding skill. In addition, because there exists no standardized literacy battery for L2 learners in Hong Kong, the tests used in the present study were not standardized measures. While the test reliabilities were acceptable, the measures used might not fully capture individual differences in L2 literacy skills. Third, our preventative interventions only focused on simple GPCs, mostly singletons. The effects of DMG intervention in teaching complex GPCs for more proficient at-risk L2 learners of English should be explored in future studies. Finally, we have not examined the effects of sustained intervention. Delayed post-tests should also be included in future work to entirely rule out Hawthorne effects in which participants may perform better when they participate in an experiment, and other short-term influences on performance. Inclusion of delayed post-test also allows examination of sustained effects of the intervention on literacy skills.

Nevertheless, practically speaking, this study has potentially demonstrated an effective way to help some at-risk L2 English learners in their L2 reading. Our findings also suggest that ESL children with low PA skills may need extra support in English

reading development. Future work may investigate the effects of PA plus GPC teaching with DMG on L2 reading development.

Conclusions

The present research provides preliminary evidence that our DMG procedure of explicitly linking the teaching of GPCs taught on that day to texts that richly represent them is more effective than a typical GPC teaching with generic shared book reading in ESL children with good PA. This is the first study to show the value of this DMG procedure for at-risk learners of English. Findings are consistent with other current reports in L1 (Savage et al., 2018, 2019) and much theorizing about PA and GPC knowledge as the dual foundations of word reading (e.g. Hulme et al., 2012), and suggest this holds in the English L2 context among Chinese students. Our findings also provide empirical support for the importance of 'phonological linkage' in reading intervention between taught GPCs and authentic text. This is the first demonstration of such a finding in the L2 populations and sets the stage for future work in this domain. Practically speaking, these results provide some preliminary guidance for how to utilize shared book reading in phonology-based reading intervention for at-risk ESL learners.

References

Beare, K. (2017). *How many people learn English.* Retrieved from

https://www.thoughtco.com/how-many-people-learn-english-globally-1210367

Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, K.,

Gotz, K., & Hulme, C. (2008). Improving early language and literacy skills:

Differential effects of an oral language versus a phonology with reading

intervention. *The Journal of Child Psychology and Psychiatry, 49*, 422-432. doi:

10.1111/j.1469-7610.2007.01849.x

Calderón, M., Slavin, R., & Sánchez, M. (2011). Effective Instruction for English

Learners. *Future of Children, 21*, 103-127.

Chen, V., & Savage, R. S. (2014). Evidence for a simplicity principle: Teaching common

complex grapheme-phonemes improves reading and motivation in at-risk readers.

Journal of Research in Reading, 37, 196–214. doi: **10.1111/1467-9817.12022**

Cirino, P., Vaughn, S., Linan-Thompson, S., Cardenas-Hagan, E., Fletcher, J., & Francis,

D. (2009). One-year follow-up outcomes of Spanish and English interventions for

English Language Learners at risk for reading problems. *American Educational*

Research Journal, 46, 744-781. doi: 10.3102/0002831208330214

Compton, D. L., Miller, A.C., Elleman, A. M., & Steacy, L. M. (2014). Have we forsaken

reading theory in the name of 'quick fix' interventions for children with reading

disability? *Scientific Studies of Reading, 18*, 55-73. doi:

10.1080/10888438.2013.836200

De Graaff, S., Bosman, A. M. T., Hasselman, F., & Verhoeven, L. (2009). Benefits of

    systematic phonics instruction. *Scientific Studies of Reading, 13*, 318-333. doi:

    10.1080/10888430903001308

Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for

    language: Why language is central to reading development. *Educational*

    Researcher, 39, 305-310. doi: 10.3102/0013189X10370204

Dowdall, N., Melendez-Torres, G. J., Murray, L., Gardner, F., Hartford, L., & Cooper, P.

    J. (2019). Shared picture book reading interventions for child language

    development: A systematic review and meta-analysis. *Child Development*,

    Advanced online publication. doi: **10.1111/cdev.13225**

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test 4th edition*.

    Pearson: San Antonio Texas.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.G. (2009). Statistical power analyses

    using G*Power 3.1: Tests for correlation and regression analyses. *Behavior*

    Research Methods, 41, 1149-1160. doi:10.3758/BRM.41.4.1149

Gerber, M., Jimenez, T., Leafstedt, J., Villaruz, J., Richards, C., & English, J. (2004).

    English reading effects of small-group intensive intervention in Spanish for K-1

    English learners. *Learning Disabilities Research & Practice, 19,* 239-251. doi:

    https://doi.org/10.1111/j.1540-5826.2004.00109.x

Graham, S., Liu, X., Aitken, A., Ng, C., Bartlett, B., Harris, K.R., & Holzapfel, J. (2017).

    Effectiveness of literacy programs balancing reading and writing instruction: A

    meta-analysis. *Reading Research Quarterly, 53*, 279-304. doi: **10.1002/rrq.194**

Hall, M. S., & Burns, M. K. (2018). Meta-analysis of targeted small-group reading interventions. *Journal of School Psychology, 66*, 54-66. Doi: 10.1016/j.jsp.2017.11.002

Hatcher, P. J., Hulme, C., & Snowling, M. J. (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *Journal of Child Psychology and Psychiatry, 45*, 338-358. doi: **10.1111/j.1469-7610.2004.00225.x**

Hatcher, P. J., Hulme, C., Miles, J. N. V., Carroll, J. M., Hatcher, J., Gibbs, S., . . . Snowling, M. J. (2006). Efficacy of small group reading intervention for beginning readers with reading-delay: A randomised controlled trial. *Journal of Child Psychology and Psychiatry, 47*, 820-827. doi: 10.1111/j.1469-7610.2005.01559.x

Hatcher, P., Hulme, C., & Ellis, A. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The Phonological Linkage Hypothesis. *Child Development, 65*, 41-57. doi: **10.1111/j.1467-8624.1994.tb00733.x**

Henbest, V., & Apel, K. (2017). Effective word reading instruction: What does the evidence tell us? *Communication Disorders Quarterly, 39*, 303-311. doi: 10.1177/1525740116685183

Hjetland, H. N., Brinchmann, E. I., Scherer, R., & Melby-Lervåg, M. (2017) Preschool predictors of later reading comprehension ability: a systematic review. Campbell Systematic Reviews, 14, 1-156. doi: **10.4073/csr.2017.14**

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. New York, NY: Routledge.

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological Science, 23*, 572-577. doi: 10.1177/0956797611435921

Jamaludin, K. A., Alias, N., Mohd K., Roselina J., DeWitt, D., & Kenayathula, H. B. (2016). The effectiveness of synthetic phonics in the development of early reading skills among struggling young ESL readers. *School Effectiveness and School Improvement, 27*, 455-470. doi: 10.1080/09243453.2015.1069749

Jenkins, J., Peyton, J., Sanders, E., & Vadasy, P. (2004). Effects of reading decodable texts in supplemental first-grade tutoring. *Scientific Studies of Reading, 8*, 53-85. doi: 10.1207/s1532799xssr0801_4

Lepola, J., Lynch, J., Laakkonen, E., Silven, M., & Niemi, P. (2012). The role of inference making and other language skills in the development of narrative listening comprehension in 4-6-year-old children. *Reading Research Quarterly, 47*, 259-282. doi: **10.1002/rrq.020**

Lesaux, N. K., & Siegel, L. S. (2003). The development of reading in children who speak English as a Second Language. *Developmental Psychology, 39*, 1005-1019. doi: 10.1037/0012-1649.39.6.1005

Li, S., Tao, S., Joshi, R. M., & Xu, Q. (2018). Second-Language Reading Difficulties Among Native Chinese Speaking Students Learning to Read English: The Roles

of Native and Second Language Skills. *Reading Research Quarterly, 53*, 423-441.

doi: **10.1002/rrq.204**

Linan-Thompson, S., Vaughn, S., Hickman-Davis, P., & Kouzekanani, K. (2003).

Effectiveness of supplemental reading instruction for second-grade English

language learners with reading difficulties. *The Elementary School Journal, 103*,

221-238. doi: https://www.jstor.org/stable/1002270

Liu, Y., Yeung, S.S., Lin, D., Wong, R.K.S. (2017). English expressive vocabulary

growth and its unique role in predicting English word reading: A longitudinal

study involving Hong Kong Chinese ESL children. *Contemporary Educational

Psychology, 49*, 195-202. doi: 10.1016/j.cedpsych.2017.02.001

McBride-Chang, C., & Kail, R. V. (2002). Cross-cultural similarities in the predictors of

reading acquisition. *Child Development, 73*, 1392-1407. doi: 10.1111/1467-

8624.00479

McBride-Chang, C., Bialystok, E., Chong, K. K. Y., & Li, Y. (2004). Levels of

phonological awareness in three cultures. *Journal of Experimental Child

Psychology, 89*, 93–111. doi: 10.1016/j.jecp.2004.05.001

Mol, S. E., & Bus, A. (2011). To read or not to read: A meta-analysis of print exposure

from infancy to early adulthood. *Psychological Bulletin, 137*, 267-296. doi:

10.1037/a0021890

Neuman, S. B., & Kaefer, T. (2018). Developing low-income children's vocabulary and

content knowledge through a shared book reading program. *Contemporary

Educational Psychology, 52*, 15-24. doi: 10.1016/j.cedpsych.2017.12.001

O'Connor, M., Geva, E., Koh, P. W. (2019). Examining reading comprehension profiles of grade 5 monolinguals and English language learners through the lexical quality hypothesis lens. *Journal of Learning Disabilities, 52*, 232-246. doi: 10.1177/0022219418815646

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (Second ed.).* Thousand Oaks, CA: Sage.

Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X. F., Martinez, A., Bloom, H. (2011). *Optimal design software for multi-level and longitudinal research (Version 3.01)* [Software].

Raven, J., Raven, J. C., & Court, J. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales.* San Antonio, TX: Harcourt Assessment.

Savage, R. S., & Cloutier, E. (2017). Early reading interventions: The state of the practice and new directions in building causal theoretical models. In K. Cain, D. Compton, & R. Parrila (Eds.), *Theories of reading development (pp.* 409-436). Amsterdam: John Benjamins.

Savage, R., Georgiou, G., Parrila, R., & Maiorino, K. (2018). Preventative reading interventions teaching direct mapping of graphemes in texts and set-for-variability aid at-risk learners. *Scientific Studies of Reading, 22*, 225-247. doi: 10.1080/10888438.2018.1427753

Savage, R., Georgiou, G., Parrila, R., Maiorino, K., Dunn, K., &Burgos, G. (2019). The effects of teaching complex grapheme-phoneme correspondences: Evidence from a dual site cluster trial with at-risk grade 2 students. *Scientific Studies of Reading*, doi: 10.1080/10888438.2019.1669607

Schaars, M. M. H., Segers, E., & Verhoeven, L. (2017). Word decoding development during phonics instruction in children at risk for dyslexia. *Dyslexia, 23*, 141-160. doi: **10.1002/dys.1556**

Sénéchal, M., Ouellette, G., & Rodney, D. (2006). The misunderstood giant: On the predictive role of early vocabulary to future reading. In D.K. Dickinson & S.B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 173-182). New York: Guilford.

Shapiro, L. & Solity, J. (2008). Delivering phonological and phonics training within whole class teaching. *British Journal of Educational Psychology, 78*, 597-620. doi: **10.1348/000709908X293850**

Shapiro, L. R., & Solity, J. (2016). Differing effects of two synthetic phonics programmes on early reading development. *British Journal of Educational Psychology, 86*, 182-203. doi: **10.1111/bjep.12097**

Stuart, M. (2014). Getting ready for reading: A follow-up study of inner city second language learners at the end of key stage I. *British Journal of Educational Psychology, 74,* 15-36. doi: 10.1348/000709904322848806

Suggate, S. P (2010). Why what we teach depends on when: Grade and reading intervention modality moderate effect size. *Developmental Psychology, 46*, 1556-1579. doi: 10.1037/a0020612

Suggate, S. P. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities, 49*, 77-96. doi: 10.1177/0022219414528540

Swanson, H. L., Kudo, M., & Guzman-Orth, D. (2016). Cognition and literacy in English

language learners at risk for reading disabilities: A latent transition

analysis. *Journal of Educational Psychology, 108*, 830-856. doi:

10.1037/edu0000102

Tse, L., & Nicholson, T. (2014). The effect of phonics-enhanced big book reading on the

language and literacy skills of 6-year-old pupils of different reading ability

attending lower SES schools. *Frontiers in Psychology, 5,* 1222. doi:

10.3389/fpsyg.2014.01222

Uchikoshi, Y., & Marinova-Todd, S. H. (2012). Language proficiency and early literacy

skills of Cantonese-speaking English language leaners in the U.S. and Canada.

Reading and Writing, 25, 2107-2129. doi: 10.1007/s11145-011-9347-2

Vadasy, P. F., & Sanders, E. A. (2010). Efficacy of supplemental phonics-based

instruction for low-skilled kindergarteners in the context of language minority

status and classroom phonics instruction. *Journal of Educational Psychology,

102*, 786-803. doi: 10.1037/a0019639

Vadasy, P.F. & Sanders, E.A. (2012). Two-year follow-up of a kindergarten phonics

intervention for English learners and native English speakers: Contextualizing

treatment impacts by classroom literacy instruction. *Journal of Educational

Psychology*, *104*, 987-1005. doi: 10.1037/a0028163

Vousden, J. I., Ellefson, M. R., Solity, J., & Chater, N. (2011). Simplifying reading:

Applying the simplicity principle to reading. *Cognitive Science, 35*, 34-78. doi:

10.1111/j.1551-6709.2010.01134.x

Wagner, R.K., Torgesen, J.K., & Rashotte, C.A. (1999). *Comprehensive Test of Phonological Processes.* Austin, TX: PRO-ED Inc.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001, 2007). *Woodcock-Johnson III*. Rolling Meadows, IL: Riverside

Yeong, S. H. M., Fletcher, J., & Bayliss, D. M. (2017). Impact of early home language exposure on phonological and orthographic skills and their contributions to English literacy abilities in English monolingual and Chinese-English bilingual adults. *Applied Psycholinguistics*, *38*, 181-210. doi: 10.1017/S0142716416000151

Yeung, P. (2006). Learning to read and spell in English among Chinese English-as-a-second-language learners in Hong Kong. *Unpublished Thesis.*

Yeung, S. S. (2018). Second language learners who are at-risk for reading disabilities: A growth mixture model study. *Research in Developmental Disabilities, 78*, 35-43. doi: 10.1016/j.ridd.2018.05.001

Yeung, S. S., & Chan, C. K. (2013). Phonological awareness and oral language proficiency in learning to read English among Chinese kindergarten children in Hong Kong. *British Journal of Educational Psychology*, *83*, 550-568. doi: **10.1111/j.2044-8279.2012.02082.x**

Yeung, S. S., Siegel, L. S., & Chan, C. K. K. (2013). Effects of a phonological awareness program on English reading and spelling among Hong Kong Chinese ESL children. *Reading and Writing, 26*, 681-704. doi: 10.1007/s11145-012-9383-6

Table 1. Matching characteristics of participants by intervention conditions

Variable	DMG intervention M (SD)	Control Intervention M (SD)	Significance t (1, 71)
Gender (% male)	59	54	0.4
Chronological Age in months	81.85 (5.51)	82.62 (5.31)	0.60
Maternal education#	2.20 (0.69)	2.19 (0.74)	0.10
Word Reading	8.88 (7.64)	9.86 (8.00)	0.53
Word Attack (Sound)	0.68 (1.34)	1.11 (1.74)	1.16
Phonemic Awareness	3.94 (1.11)	4.24 (1.42)	0.99
Spelling	0.18 (0.18)	0.11 (0.20)	1.51
Sentence Comprehension	9.88 (2.53)	10.59 (3.56)	0.96
Vocabulary	10.61 (1.67)	11.00 (1.45)	1.03
Non-verbal IQ	15.94 (2.55)	16.59 (2.91)	1.00

Note: #maternal education: 1 = primary education; 2 = secondary education; 3 = college degree; 4 = postgraduate degree.

Table 2. Means and Standard Deviations for Pretest and Posttest Measures by Intervention Group

Intervention	DMG			Control		
Measure	Pre-test M (SD)	Post-test M (SD)	Effect Sizes Cohen's D	Pre-test M (SD)	Post-test M (SD)	Effect Sizes Cohen's D
Word Reading	8.88 (7.64)	18.00 (7.64)	1.52	9.86 (8.00)	19.19 (8.82)	1.30
Word Attack (Item)	0 (0)	0.94 (1.59)	.59	0 (0)	1.29 (1.75)	.74
Phonemic Awareness	3.94 (1.11)	4.41 (2.35)	.20	4.24 (1.42)	4.81 (2.31)	.23
Spelling	0.18 (0.18)	0.41 (0.61)	.39	0.11 (0.20)	0.73 (0.99)	.72
Sentence Comprehension	9.88 (2.53)	11.68 (2.92)	.59	10.59 (3.56)	11.57 (3.43)	.26
Vocabulary	10.61 (1.67)	11.62 (3.40)	.31	11.00 (1.45)	12.32 (3.55)	.50

Table 3

HLM results for the effect of intervention condition by phonological ability on post-test attainment

_____

Variable	Student-Level Model		Classroom-Level Model	
	Coefficient	*SE*	Coefficient	*SE*

_____

Word reading Post = Dependent Variable

Variable	Student-Level Model		Classroom-Level Model	
Intercept	19.15	(1.45)***		
Intervention	-0.65	(2.12)		
Intervention x Phonological Ability	0.53	(0.19)**		
Intervention x PPVT	0.38	(0.38)		
Word reading pre-test			0.42	(0.17)*
Random effects				
Child	40.53	(8.32)***		
Class			10.77	(8.97)

Woodcock-Johnson Word Attack Post = Dependent Variable

Variable	Student-Level Model		Classroom-Level Model	
Intercept	1.20	(0.27)***		
Intervention	-0.12	(0.39)		
Intervention x Phonological Ability	0.04	(0.04)		

Intervention x PPVT	0.52	(0.40)	
Word reading pre-test			0.49
(0.23)*			
Random effects			
Child	2.53	(0.47)	
Class			0
(0)			

Spelling = Dependent Variable

Intercept	0.80	(0.17)***	
Intervention	- 0.42	(0.25)	
Intervention x Phonological Ability	0.04	(0.17)*	
Intervention x PPVT	-0.01	(0.04)	
Word reading pre-test			1.76
(0.64)*			
Random effects			
Child	0.35	(0.07)***	
Class			0.19
(0.11)			

Woodcock-Johnson Sentence Comp = Dependent Variable

Intercept	11.42	(0.51)	
Intervention	0.50	(0.75)	
Intervention x Phonological Ability	0.17	(0.07)*	
Intervention x PPVT	0.38	(0.16)*	
Word reading pre-test			0.43
(0.27)			
Random effects			
Child	9.19	(1.59)	

| Class | 0 |
| (0) | |

_____

_____

Note: HLM = hierarchical linear modeling; PPVT = Peabody Picture Vocabulary Test

*** *p* < .001, ** *p* < .01, * *p* <.05*

CONSORT diagram

Enrollment

Recruitment & Screening (n=253)

Excluded (n=182)

Randomized (n=71)

Baseline assessment (n=71)

Allocation

DMG intervention (n=34)

Control intervention (n=37)

Post-intervention assessment (n=34)

Post-intervention assessment (n=37)

Analysis

Analysed (n=71)