Comparing the development of mathematical abilities in individuals with different neurodevelopmental conditions

Erica Ranzato

Department of Psychology and Human Development

IOE, UCL's Faculty of Education and Society

Submitted for the degree of Doctor of Philosophy

UCL

Declaration

‘I, Erica Ranzato confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.'

Signature ______________________

Abstract

This thesis adopts a neuroconstructivist approach and uses the multi-level framework of mathematical cognition by Gilmore (2023) to investigate and compare the mathematical profiles of three neurodivergent populations: Down syndrome (DS), Williams syndrome (WS), and autism. The two overarching aims of this thesis are to investigate syndrome specificity in mathematical profiles of neurodivergent populations and to explore their mathematical learning experiences. The thesis presents four studies including three experimental studies which employed both quantitative and qualitative methodologies and a systematic review. The first study explores and compares enumeration skills in children and adults with DS and WS, while the second study examines functional and structural indicators of the home learning environment of these populations. The third study presents a systematic review of mathematical abilities in autism. The final study explores the teaching strategies employed by educators to support mathematical abilities of primary school students with DS in the inclusive classroom. The last chapter presents a general discussion which highlights the key findings and implications of the research presented in this thesis and future directions. The findings from this thesis highlight similarities in the mathematical profiles of individuals with DS and WS, particularly in relation to the performance level in enumeration and in relation to their home mathematical environment. Moreover, they show that the mathematical learning experiences of individuals with DS and with WS are rich and individualised. Results from this work emphasise a gap in the literature investigating mathematical development of autistic individuals with intellectual disabilities. Finally, this thesis addresses some misconceptions concerning mathematical abilities in neurodivergent populations. In particular it confirms the multi-component and developing nature of mathematical profiles in these populations and challenges the prejudice which associates autism with exceptional mathematical abilities.

Impact statement

In an increasingly technologically-driven and complex societal framework, mathematical abilities have emerged as crucial for the promotion of social inclusion, independence, and quality of life. The past two decades have seen an increase in research on mathematical cognition. However, despite substantive advancements, significant gaps remain in our understanding of how mathematical abilities develop in neurodivergent populations. This is particularly startling given the increasing number of neurodivergent students in mainstream educational settings and given the impact that mathematical abilities have on their day-to-day.

This thesis was set up to investigate and compare the mathematical profiles of three neurodevelopmental conditions (NDCs): Down Syndrome (DS), Williams Syndrome (WS), and autism. It made three significant contributions:

1) Provide further insights into the mathematical profiles of these populations.
2) Extend the knowledge around the syndrome specificity of mathematical profiles.
3) Provide insights into the mathematical learning environments of DS and WS populations.

The first study focused on enumeration skills in individuals with DS and WS using eye-tracking data. Results revealed that the performance and strategies used by these individuals did not differ from the typically developing (TD) control group. Some differences were observed in relation to their eye movements. However, these did not affect their performance. The second study investigated the home learning environment of primary school children with DS and WS and showed that the type and frequency of home-based activities, parental expectations, as well as the parental and child's attitudes towards learning were similar between the two groups. Instead, the type and frequency of support changed based on the level of general functioning of the child. The third study presented a systematic review of mathematical abilities in autism and reported a partial profile for autistic individuals with intellectual disabilities (ID) and one for autistic individuals without ID. The final study explored experiences and teaching strategies employed by educators supporting the mathematical skills of primary school students with DS in the inclusive classroom. The study identified the main challenges faced by educators and showed that the teaching strategies employed were aligned with the general cognitive profile of students with DS and that educators adapted learning materials to their students' profile.

The studies included in this thesis address some misconceptions about mathematical abilities in neurodivergent populations and about mathematical development in general. The findings from the systematic review challenges the prejudice which associates autism with exceptional mathematical abilities. The findings from the studies comparing specific components of the mathematical profiles of DS and WS report some similarities between these populations. The studies investigating mathematical learning experiences in neurodivergent populations show that these are rich and individualised. Finally, the adoption of a neuroconstructivist approach and of the multi-level framework of mathematical cognition by Gilmore (2023) spotlights the developing and multi-component nature of mathematical profiles in neurodivergent populations. Ultimately, the findings from these studies will assist researchers, educators, and decision-makers in identifying effective educational interventions for individuals with NDCs and in promoting inclusion in the classroom and in society.

UCL Research paper declaration forms

UCL Form referencing the doctoral candidate's own published work(s)

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)

   a) **What is the title of the manuscript?**

   Subitizing in Williams syndrome and Down syndrome: Insights from eye movements.

   b) **Please include a link to or doi for the work.**

   https://doi.org/10.1016/j.ridd.2020.103746

   c) **Where was the work published?**

   Research in Developmental Disabilities.

   d) **Who published the work?** (e.g. OUP)

   Elsevier.

   e) **When was the work published?**

   The work was made available online in August 2020 and was published in November 2020.

   f) **List the manuscript's authors in the order they appear on the publication.**

   Erica Ranzato, Andrew Tolmie, Jo Van Herwegen.

   g) **Was the work peer reviewed?**

   Yes, the work was peer reviewed.

   h) **Have you retained the copyright?**

   As an author of this article, I retain the right to include it in a thesis or dissertation, provided it is not published commercially.

   i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

   No, earlier versions of the manuscript were not uploaded to a preprint server.

   If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

   ☒ *I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)

   a) **What is the current title of the manuscript?**

   n/a

b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

n/a

c) **Where is the work intended to be published?** (e.g. journal names)

n/a

d) **List the manuscript's authors in the intended authorship order.**

n/a

e) **Stage of publication** (e.g. in submission)

n/a

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4)

J.V.H. designed and directed the study. E.R. collected the data, prepared, and analysed eye tracking data. E.R. wrote the manuscript in consultation with A.T. and J.V.H.

4. In which chapter(s) of your thesis can this material be found?

This material can be found in Chapter 2.

5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Erica Ranzato

Date:

15/03/2024

Supervisor/ Senior Author (where appropriate)

Jo Van Herwegen

Date:

15/04/2024

UCL Form referencing the doctoral candidate's own published work(s)

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)

   a) **What is the title of the manuscript?**

      The home learning environment of primary school children with Down Syndrome and those with Williams Syndrome.

   b) **Please include a link to or doi for the work.**

      https://doi.org/10.3390/brainsci11060733

   c) **Where was the work published?**

      The work was published in Brain Sciences.

   d) **Who published the work?** (e.g. OUP)

      MDPI.

   e) **When was the work published?**

      The work was published on 31 May 2021.

   f) **List the manuscript's authors in the order they appear on the publication.**

      Erica Ranzato, Andrew Tolmie, Jo Van Herwegen.

   g) **Was the work peer reviewed?**

      Yes, the work was peer reviewed.

   h) **Have you retained the copyright?**

      Yes, I have retained the copyright.

   i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

      No, earlier versions of the manuscript were not uploaded to a preprint server.

   If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

   ☒ *I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)

   a) **What is the current title of the manuscript?**

      n/a

   b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

      n/a

c) **Where is the work intended to be published?** (e.g. journal names)

n/a

d) **List the manuscript's authors in the intended authorship order.**

n/a

e) **Stage of publication** (e.g. in submission)

n/a

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

Conceptualization: E.R. and J.V.H.; Methodology: E.R. and J.V.H.; Formal analysis: E.R.; Investigation: E.R.; Writing - original draft preparation: E.R.; Writing - review and editing: E.R., A.T. and J.V.H.; Visualization: E.R.; Supervision, J.V.H. and A.T.

4. **In which chapter(s) of your thesis can this material be found?**

This material can be found in Chapter 3.

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Erica Ranzato

Date:

15/03/2024

Supervisor/ Senior Author (where appropriate)

Jo Van Herwegen

Date:

15/04/2024

UCL Form referencing the doctoral candidate's own published work(s)

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)

   a) **What is the title of the manuscript?**

      n/a

   b) **Please include a link to or doi for the work**

      n/a

   c) **Where was the work published?**

      n/a

   d) **Who published the work?** (e.g. OUP)

      n/a

   e) **When was the work published?**

      n/a

   f) **List the manuscript's authors in the order they appear on the publication.**

      n/a

   g) **Was the work peer reviewed?**

      n/a

   h) **Have you retained the copyright?**

      n/a

   i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If 'Yes', please give a link or doi)

      n/a

      If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

      ☐

      *I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)

   a) **What is the current title of the manuscript?**

      OSF public registration: Mathematical abilities in the autistic populations: A systematic review (https://osf.io/q3d7c).

   b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if 'Yes', please give a link or doi)

No, the manuscript has not been uploaded to a preprint service.

c) **Where is the work intended to be published?** (e.g. journal names)

The study has been registered in the Open Science Framework website (https://osf.io/q3d7c). The manuscript is intended to be published in the peer-reviewed journal Autism Research.

d) **List the manuscript's authors in the intended authorship order.**

Erica Ranzato, Andrew Tolmie, Jo Van Herwegen.

e) **Stage of publication** (e.g. in submission)

Not submitted yet.

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

Conceptualization: E.R. and J.V.H.; Methodology: E.R. and J.V.H.; Data extraction and data analysis: E.R.; Writing - original draft preparation, E.R.; Writing - review, E.R., A.T. and J.V.H.; Data visualisation, E.R.; Supervision, J.V.H. and A.T.

4. **In which chapter(s) of your thesis can this material be found?**

This material can be found in Chapter 4.

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Erica Ranzato

Date:

3/04/2024

Supervisor/ Senior Author (where appropriate)

Jo Van Herwegen

Date:

15/04/2024

UCL Form referencing the doctoral candidate’s own published work(s)

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)

   a) **What is the title of the manuscript?**

      n/a

   b) **Please include a link to or doi for the work**

      n/a

   c) **Where was the work published?**

      n/a

   d) **Who published the work?** (e.g. OUP)

      n/a

   e) **When was the work published?**

      n/a

   f) **List the manuscript’s authors in the order they appear on the publication.**

      n/a

   g) **Was the work peer reviewed?**

      n/a

   h) **Have you retained the copyright?**

      n/a

   i) **Was an earlier form of the manuscript uploaded to a preprint server?** (e.g. medRxiv). If ‘Yes’, please give a link or doi)

      n/a

      If ‘No’, please seek permission from the relevant publisher and check the box next to the below statement:

      ☐

      *I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)

   a) **What is the current title of the manuscript?**

      OSF project: Teaching mathematics to primary school students with Down syndrome (https://osf.io/z7m8r/).

   b) **Has the manuscript been uploaded to a preprint server?** (e.g. medRxiv; if ‘Yes’, please give a link or doi)

No, the manuscript has not been uploaded to a preprint service.

c) **Where is the work intended to be published?** (e.g. journal names)
The study has been published in the Open Science Framework website (https://osf.io/z7m8r/). The manuscript is intended to be published in the peer-reviewed journal British Journal of Learning Disabilities.

d) **List the manuscript's authors in the intended authorship order.**

Erica Ranzato, Unta Taiwo, Andrew Tolmie, Jo Van Herwegen.

e) **Stage of publication** (e.g. in submission)

Not submitted yet

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)

Conceptualization: E.R.; Funding acquisition: ER; Methodology: E.R.; Resources: ER and UT; Data collection: ER and UT; Data analysis: E.R.; Writing - original draft preparation, E.R.; Writing - review, E.R., A.T. and J.V.H.; Data visualisation, E.R.; supervision, J.V.H. and A.T.

4. **In which chapter(s) of your thesis can this material be found?**

This material can be found in Chapter 5.

5. **e-Signatures confirming that the information above is accurate** (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate

Erica Ranzato

Date:

3/04/2024

Supervisor/ Senior Author (where appropriate)

Jo Van Herwegen

Date:

15/04/2024

Acknowledgements

I owe a debt of gratitude to a few wonderful colleagues, whom I met at a late stage of my PhD journey, at a time when I was struggling with my research work, and I had lost the enthusiasm and curiosity that originally drew me towards research. They helped me rediscover that passion and gave me confidence in my own research. Thanks, Unta and Angie, for your time and for bringing fresh energy into what at times felt like an overwhelming and thankless experience.

Finally, I would have not been able to consider undertaking this journey, let alone completing it without the immense support of my husband. Thank you, Lorenzo, for your patience and for your understanding. I feel this project has taken so much of our time, but it made me realise what really matters to me.

Table of Contents

List of Abbreviations

AF	Activity frequency score
AoI	Area of interest
AS	Asperger syndrome
ASD	Autism spectrum disorder
BPVS	British picture vocabulary scale
CA	Chronological age
CDR-LMC	Cognitive developmental skills in arithmetic – Linguistic scale, Mental representation scale and Contextual scale
CDR-P	Cognitive developmental skills in arithmetic – Procedural calculation scale
CEQ	Close-ended question
ChA	Child attitude score
DAS	Differential ability scales
DD	Developmental dyscalculia
DS	Down syndrome
DSM	Diagnostic and statistical manual of mental disorders
ES	Expectation score
ET	Eye-tracking
FSIQ	Full scale IQ
HFA	High functioning autism
HLE	Home literacy environment
HME	Home mathematics environment
ID	Intellectual disability
IS	Importance score
KBIT	Kaufman brief intelligence test
KeyMath	KeyMath diagnostic assessment
KS	Key stage
KTEA-MA	Kaufman test of educational achievement – Math applications scale
KTEA-MC	Kaufman test of educational achievement – Math computation scale
MA	Mental age
MD	Mathematical difficulties
MPI	Mathematical problem instrument
MPWS	Mathematical word problem solving
NDC	Neurodevelopmental condition
NL	Number line

NVIQ	Non-verbal IQ
OEQ	Open-ended question
PA	Parent attitude score
PAE	Percent absolute error
PIAT	Peabody individual achievement test
PPVT	Peabody picture vocabulary test
RCPM	Raven's coloured progressive matrices
RQ	Research question
RT	Reaction time
SES	Socioeconomic status
TA	Teaching assistant
TD	Typically developing
TEDI-MATH 1	Test for the diagnosis of mathematical competencies – Sub-test 1
TEDI-MATH 2	Test for the diagnosis of mathematical competencies – Sub-test 2
TEDI-MATH 5	Test for the diagnosis of mathematical competencies – Sub-test 5.1
TEMA	Test of early mathematics ability
TOMA	Test of mathematical abilities
TOMA-C	Test of mathematical abilities – Computation scale
TOMA-SP	Test of mathematical abilities – Story problem scale
TTR	Arithmetic number facts test
VABS	Vineland adaptive behaviour scale
VIQ	Verbal IQ
w	Weber fraction
WIAT	Wechsler Individual Achievement Test
WIAT-MR	Wechsler Individual Achievement Test – Mathematical reasoning scale
WIAT-NO	Wechsler Individual Achievement Test – Numerical operation scale
WRAT	Wide range achievement test
WJ	Woodcock–Johnson test of achievement
WJ-AP	Woodcock–Johnson test of achievement – Applied problems scale
WJ-C	Woodcock–Johnson test of achievement – Calculation scale
WJ-F	Woodcock–Johnson test of achievement – Fluency scale
WS	Williams syndrome

List of Tables

List of Figures

Chapter 1: General introduction

Mathematical skills are fundamental for independent life in our technologically-driven and complex society (Butterworth, 2005). How and whether these skills are acquired, or fail to be acquired, is of great importance not only to the individual and their close community but also to the organisation of formal education (Butterworth, 2005) and to the broader societal and economic context. This is especially relevant for individuals with neurodevelopmental conditions (NDCs) and their families in the context of their inclusion in society, level of independence, and quality of life.

There are still many gaps in our understanding of how mathematical abilities of individuals with NDCs develop and this hinders the impact of the efforts of parents[1] and family members, of educators supporting these individuals and of the whole education system. For instance, if we (academics) do not know how the development of mathematical abilities in children with Down syndrome (DS) unfolds compared to the one of the typically developing (TD) peers that sit in the same classroom, how can we expect teachers to support the learning and inclusion of these students? If we do not have a clear picture of the mathematical profile of an autistic child, how can we ask educators to set achievable but ambitious targets and have realistic expectations for them?

The studies included in this thesis were conducted during my posts as part-time doctoral student at Kingston University and at UCL (University College London) in the past eight years and are driven by my commitment to stay true to the experiences of individuals with NDCs and their families and to not lose sight of their priorities and daily lives. The main aspiration of this research, which is rooted in cognitive and developmental psychology, is to contribute to the field of mathematical education by providing more clarity and a better understanding of the processes and components underpinning the development of mathematical skills in individuals with NDCs.

The following two sections present the theoretical perspective and the cognitive framework of mathematical cognition that have shaped this doctoral

[1] In this thesis the terms “parent”, “carer”, and “caregiver” are used interchangeably.

thesis. Then, a description of the populations investigated in the studies included in this thesis is provided, followed by a brief overview of the educational system in England, where all the studies included in this thesis were based. The final part of this chapter presents the aims and the outline of the current thesis.

1.1 Theoretical and methodological considerations on development

A researcher's theoretical perspective on development deeply affects the way in which NDCs are described and investigated (Karmiloff-Smith, 1998). In fact, the use of different theoretical approaches may influence many aspects of their work, such as the language used to describe key-constructs, the type and the focus of the research questions (RQs) asked, the type of study carried out in terms of design and methodology used, and how findings are interpreted (D'Souza & Karmiloff-Smith, 2017). Differences in research outcomes affect the design of policies and educational practices such as the assessment tools and the interventions employed in school, which in turn have an impact on the educational outcomes of students with NDCs. Different levels of educational outcomes in turn affect the level of independence of individuals with NDCs and their quality of life, as well as the one of their families, and of the larger community that supports them (D'Souza & Karmiloff-Smith, 2017).

Given the impact of the theoretical perspective on key decisions-making points which deeply influence the lives of individual with NDCs, the first part of this section describes neuroconstructivism, that is the theoretical stance adopted in this thesis. The second part of this section focuses on the different methodologies employed in developmental research and explains the language used to describe development (i.e., in line, not in line, typical, atypical, delayed) in the current thesis.

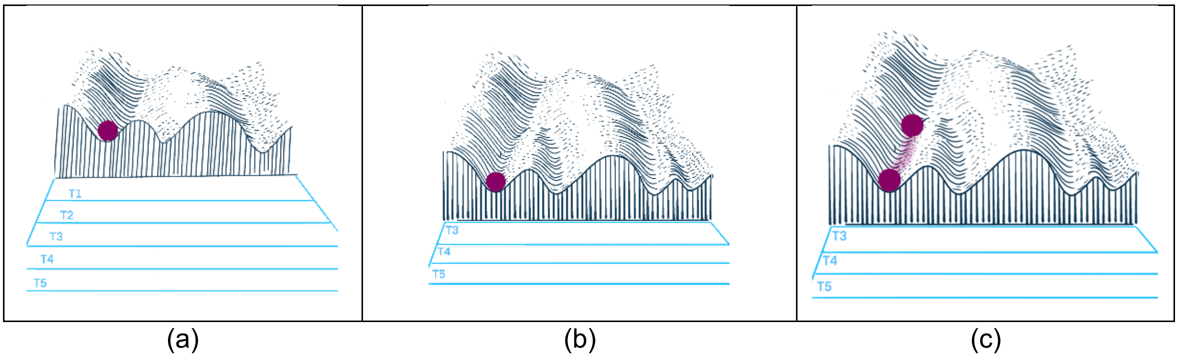
1.1.1 Neuroconstructivist theory of development

Neuroconstructivism views development as a complex and dynamic process involving multi-way interactions between different levels of description (Mareschal, 2007). The complexity of this process arises from acknowledging the system's multiple levels of description, in that the brain is situated within a body which itself is situated within the physical and social environment. Moreover, because development happens within a dynamic system, neuroconstructivism describes development as the product of multi-way interactions between all the different levels of description, rather than as a linear succession of events (Mareschal, 2007). As described in the book edited by Mareschal (2007), this approach leads to a view where gene expression, cell development, brain region development, and individual development occur and are influenced through multi-way interactions occurring at various levels. In fact, at the level of genes, gene expression may affect cellular processes and may lead to changes in brain structures or functions and affect cognitive and behavioural outcomes. At the level of the cells, the individual neuron will adapt to a special function (that is, it will acquire progressive specialisation) depending on the chemical and cellular context it finds itself in (that includes both genetic and physical environmental factors). Changes in the quality and quantity of this context and of the interactions of the neuron with the elements of the physical environment (that includes other brain cells and brain areas) will have an influence on the functionality of the neuron itself as well as on the connectivity it develops with neighbouring cells, triggering a multi-way process where cascading effects may also influence other levels. At the level of the brain areas, changes in the nature of the input to a brain area or changes in the functionality of the neighbouring areas may lead to changes in the functionality of the target brain area. At the level of the environment, the multi-way interactions can be observed, for example, in an infant's active and progressive selection and processing of different kinds of input, as described by Karmiloff-Smith (1998). In this scenario, the child does not passively internalise information from the environment, but they are actively involved in the exploration of the environment, and they play a key role in the selection of the information and in shaping the experiences that they will encounter. These processes influence the feedback loop between child and environment as they enable the child to identify aspects of the environment that are relevant and rewarding to them. The infant will then spend more time interacting with these components, and thus they will develop those specific representations which reflect relevant and rewarding information in the environment. This, in turn, may also shape the social environment in terms of parental interactions, as parents may prioritise the stimuli identified by the child as relevant and rewarding over other stimuli.

Neuroconstructivism provides a unifying framework for the understanding of the development of neurotypical and neurodivergent populations, in that both are viewed as unfolding developmental trajectories within a constrained space of possibilities (Mareschal, 2007). Indeed, the neuroconstructivist perspective describes a child with NDCs as an adaptive brain-body system with a different initial state than the neurotypical individual (D’Souza & D’Souza, 2020). In fact, the effects of differences in gene expressions are likely to originate basic-level deficits that generate multi-level interactions and cascading effects on multiple cognitive functions (Mareschal, 2007). According to this view, the emerging characteristics of NDCs are adaptations to a series of atypical constraints that deflect the normal path of development (Mareschal, 2007). It is important to stress the use of the adjective “emerging” when defining the profile of an individual, as their profile does not appear full-blown at birth, but, rather, develops gradually and sometimes in transformative ways with age (Thomas et al., 2009). These adaptations serve a functional purpose as they are the solution that works for that particular brain-body system at that specific time in development, but they might later constrain the emergence and the development of cognitive skills (D’Souza & D’Souza, 2020). As a result of cascading effects and multilevel interactions, children with NDCs are likely to develop along pathways that differ from the norm, with widespread repercussions at the broader cognitive level (Karmiloff-Smith, 1998) and on multiple domains. For example, cascading effects may impact on an individual’s development across the social, educational, and vocational domains and may reduce their level of independence and quality of life.

In summary, the neuroconstructivist approach seeks to explain development by considering the impact of low-level processes (which are affected by both internal and external constraints), the impact of cascading effects operating on a multi-way interaction basis, and the role of time (Karmiloff-Smith, 1998). In particular, the central role of time in the study of development is encapsulated in the concept of developmental trajectory, which is illustrated here using an amended version of Waddington's epigenetic landscape[2] (Figure 1.1).

Figure 1.1. Waddington's epigenetic landscape.



The dot shown in Figure 1.1 represents the brain-body system with its genetic information, where individual differences might be represented, for example, as differences in the size of the dot or the speed of the dot along the horizontal axis. The landscape represents the dynamic environment and all the external factors impacting on the individual's performance and development. The hills represent the effort required to move from one path to another and to potentially reach different outcomes. The position of the dot at different points in time can be viewed as a snapshot of the individual's developmental trajectory and corresponds to the observed behaviour, that is the outcome of an adaptive process at a given point in time in their development. Figure 1.1 (a) shows the behavioural outcome of the individual at time T0. Figure 1.1 (b) shows the behavioural outcome of the individual at time T3. The vertical section shows the space of possibilities, that is all the different outcomes that the individual could report on a specific ability. For example, this could be the reaction time (RT) measured when counting dots shown on a screen. The purple trail on Figure 1.1 (c) shows the developmental pathway, that is how the developmental trajectory unfolded as time passed from T0 to T3.

An important stance of the neuroconstructivist approach is that the study of developmental trajectories reveals the emergence of different routes through development. This can lead to the following scenarios:

[2] Waddington's epigenetic landscape was originally conceived to depict embryonic development. In the past, it has been used to represent development in different cognitive areas such as motor development, development of emotion, and language development (Baedke, 2013).

- Different pathways can lead to the same behavioural outcomes because different processes may lead to similar behavioural outcomes. If applied to the TD population and a population with NDCs, this scenario shows how a behavioural performance falling within the neurotypical range may be supported by different processes in different populations and it highlights the importance of investigating the routes to development (Karmiloff-Smith, 1998).
- The same constraint may have different cascading effects on the development of different individuals. This scenario can explain levels of variability within the same population and shows how two individuals may start with only slightly differing parameters but, with development, the effects of this small difference might influence different domains of their developing system and lead to different developmental trajectories and outcomes (D'Souza & Karmiloff-Smith, 2017).

1.1.2 Methods used to investigate development

Research in neurodivergent populations, such as populations with NDCs commonly involves the use of matched-group designs to determine whether the observed level of performance of individuals with NDCs differs from the performance of a control group. However, matched-group designs do not provide information about the mechanistic explanation of the behaviour observed (Thomas et al., 2009), in that they do not allow any conclusion to be drawn about the processes that underpin the behaviour observed. This limitation is overcome by the developmental trajectories approach. Both approaches are adopted in this thesis and are described in the following sections.

1.1.2.1 Matched-group designs

Matched-group designs are used to compare those with NDCs to either another clinical group (cross-syndrome comparison) or a group of TD individuals. Both approaches are adopted in this thesis. Cross-syndrome comparisons are a useful tool for unpicking syndrome specificity. This approach is used in the study in Chapter 2 to compare enumeration abilities in DS and in WS, and in the study in Chapter 3 to compare the home learning environments of primary school children with DS and WS. On the other hand, the use of TD control groups allows insight into whether the behaviour observed is in line with neurotypical outcomes and development. This approach is also adopted in the study in Chapter 2, where the performance of the two clinical groups is compared to the one of TD participants matched on developmental level, also labelled as mental age (MA)[3], measured through the RCPM assessment[4].

As described by Thomas et al. (2009), there are two methods in which the experimental group can be matched to the control group: individual matching and group matching. Individual matching is achieved through the pairing of each individual in the experimental group with an individual of the control group that has the same chronological age (CA) or the same score on the standardised test used for matching for MA. Group matching pairs the experimental and the control groups based on their average CA and / or their average scores on a chosen standardised test selected as measure of MA. All the studies included in this thesis employ a group matching method.

For both methods the comparison groups can be matched on different measures such as CA or the developmental level derived on relevant standardised tests (MA). The methodological choices made by the researcher in relation to the comparison group(s) employed, the matching factor(s) selected, and the measure(s) used are based on their theoretical assumptions and have important implications on the inferences that can be drawn from the findings of the study (Jarrold & Brock, 2004).

The aim of matching is to rule out “non-central” explanations of group differences (Jarrold & Brock, 2004) and it is based on theoretical assumptions that the matching factor does not play a critical role in explaining variations between the groups, if any. By doing so, the use of a matched-group allows the experimenter to assess differences and similarities between groups in relation to the specific factor that formed the basis for matching (Ansari, 2003), and limits the

[3] This thesis adheres to the common practice in the field of developmental psychology of using the term 'mental age' (MA) instead of the newer term 'developmental level'. It is important to acknowledge that language is evolving within the field, and this is shown by the coexistence of both old and new terms in this thesis.

[4] In this thesis, when a measure of developmental level is used to match the experimental group and the control group, the name of the assessment used to measure developmental level is always reported in brackets. For example, in this case the matching criteria would be reported as MA (RCPM).

implications of the findings by excluding that specific factor from the interpretation of the findings. For example, a study with an experimental design which compares counting abilities between two groups matched for language skills excludes the language scores from the inferential analyses based on the theoretical assumption that language abilities are a critical factor for both the investigated groups when it comes to counting. As such, this design cannot provide insights on whether the language domain is a crucial factor to explain differences or similarities in counting between the two groups. Moreover, the choice of the standardised test used to measure language skills is based on the theoretical assumption that such assessment is the correct measure for the language domain (Thomas et al., 2009). To follow the example above, the choice of using a measure of receptive language rather than, for example, a measure of expressive language to match the two groups, not only assumes that receptive (rather expressive) language skills play a crucial role for the development of counting for both groups, but it will also lead to different interpretations of the findings. In fact, as explained by Ansari (2003), if there is a difference between the groups, the investigator can discuss differences over and above receptive language abilities. Conversely, if the counting performance of the two groups do not differ significantly, the experimenter will be able to state that the experimental group's counting performance is at the level expected given their receptive language ability. However, this may not be the case for their expressive language skills, and this is where the initial assumptions made by the researcher (i.e., when choosing the matching cognitive skill and the assessment tool) limit the interpretation of the findings of the study.

When compared with a TD control group, the investigated ability of the experimental group can be found to be either in line or not in line with the control group. It is worth highlighting that even if the performance is reported to be in line with the control group this does not imply that the development is typical, even if the groups have been matched on both CA and MA. In fact, as discussed in the paragraph above, the lack of significant differences between groups in a specific point in time, does not imply typicality because performance scores within the typical range are not necessarily the results of the same cognitive processes (Karmiloff-Smith, 1998). Indeed, a performance can be defined as being typical only when the typical behaviour observed is driven by the same mechanisms

(Karmiloff-Smith, 1998). Hence, the assumption that similar outcomes are driven by the same cognitive processes needs to be validated by further analyses, such as correlational or regression analyses. For example, the use of correlational analyses separately for the experimental group and for the control group allows for the identification of similar correlational patterns between groups. Similarly, the use of regression analyses separately for the experimental group and for the control group allows for the comparison of the predictors of the investigated ability. In case of similar correlational patterns and of similar patterns of predictors the performance is reported to be typical. Conversely, in cases in which the relationships between the investigated variables are different, or the performance of the two groups is explained by different factors, the performance is reported to be atypical. In the instances where the performance of the experimental group is at the same level of a TD group matched for MA but significantly younger and present similar correlational or regression patterns to the control group, then the performance is defined as delayed.

While matched-group designs can provide information on whether the abilities of the experimental group are in line or not with the control group, this design does not provide the researcher with the temporal perspective. In fact, the analysis of change over time is often lost in the matched-group designs where age might be factored out or might not be adequately considered in the discussion of the findings. This is especially common in studies investigating rare conditions where CA is used as a matching factor and / or where wide age ranges are used to obtain larger sample sizes to increase the power of the statistical analyses conducted. This limitation is overcome by the use of the developmental trajectories approach proposed by Thomas et al. (2009).

1.1.2.2 Developmental trajectories method

According to neuroconstructivism, understanding development implies the investigation of developmental pathways. Moreover, the study of development requires not only the individuation of routes of development, but also understanding how these routes emerge over time (D'Souza & Karmiloff-Smith, 2017).

The developmental trajectories approach proposed by Thomas et al. (2009) requires the researcher to describe behaviour in a way that focuses on change over time. This method uses data collected at multiple points either longitudinally, from cross-sectional samples, or from both designs and uses correlations to explore relationships between different cognitive abilities and time. It involves two steps 1) the construction of a function linking the performance on a specific experimental task with CA or MA and, 2) the assessment of whether this function differs between groups.

The studies included in this thesis use cross-sectional developmental trajectories. The study in Chapter 2 investigates changes across development of reaction times (RTs) and eye movements in children and adults with DS and WS. The study in Chapter 3 investigates whether the distribution of the frequency of home-based learning activities and of parental expectations of primary school children with DS and WS change with the year group of the child. The study in Chapter 4 builds cross-sectional developmental trajectories using standardised scores from studies investigating arithmetic word problems, calculation, and overall mathematics achievement in autism.

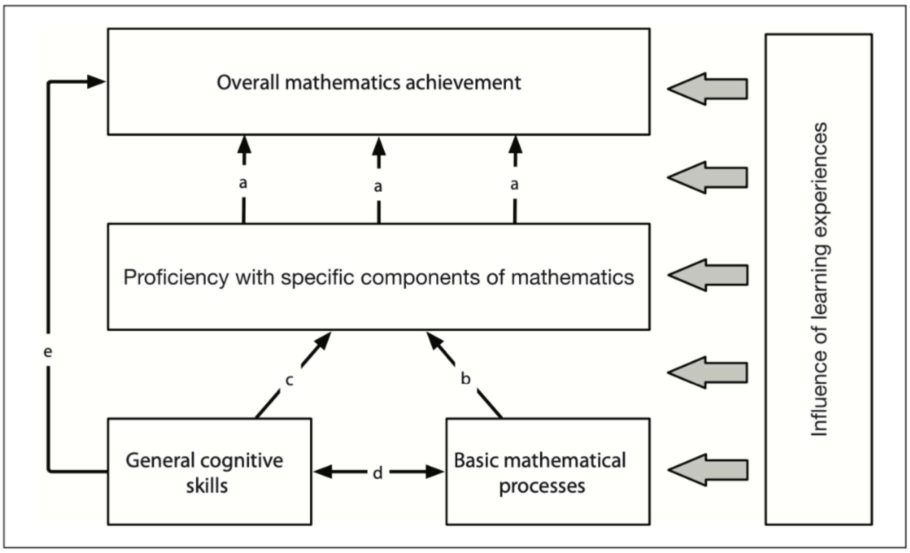
1.2 Mathematical cognition

In this thesis, Gilmore (2023)'s multi-level framework of mathematical cognition is used to describe the components and processes that underpin mathematical achievement. This section describes the multi-level framework of mathematical cognition proposed by Gilmore (2023) and provides an overview of the elements of the framework which are investigated in the studies included in the current thesis.

1.2.1 Multi-level framework of mathematical cognition

Figure 1.2 shows the multi-level framework of mathematical cognition proposed by Gilmore (2023). The framework has emerged from existing empirical evidence on TD population, and it provides a three-level hierarchy structure to describe how different processes and components fit within the broader concept of achievement in mathematics. The framework only includes cognitive and contextual factors and uses the distinction between domain-specific and domain-general factors to categorise these variables. As explained by De Smedt (2022), domain-specific factors are variables that are specific for learning in a particular domain, in this case mathematics. Domain-general factors are instead variables that are applied to learning on a broad level, regardless of the type of information being learned or the specific domain of learning (De Smedt, 2022).

Figure 1.2. Multi-level framework of mathematical cognition by Gilmore (2023).



Note: From “Understanding the complexities of mathematical cognition: A multi-level framework” by C. Gilmore, 2023, Quarterly Journal of Experimental Psychology, 76(9), p. 2 (https://doi.org/10.1177/17470218231175325). CC BY 4.0 DEED.

According to the framework proposed by Gilmore (2023), overall mathematics achievement emerges from proficiency with specific components of mathematics, which in turn recruit basic mathematical processes. Overall mathematics achievement refers to an individual’s overall attainment in mathematics and is measured using broad measures. Instead, specific components of mathematics capture an individual’s performance in distinct mathematical sub-components for which it may be anticipated that they will use a more-or-less consistent set of mathematical knowledge and skills. According to this definition, examples of specific components of mathematics are count sequence knowledge and arithmetic word problem skills (Gilmore, 2023). The nature and content of these components may change over development, and depending on the researcher’s aim it is possible to define them more or less broadly. For example, sometimes it may be appropriate to consider calculation as a single component, while other times it may be preferable to consider addition, subtraction, multiplication, and division as separate components. Finally, basic mathematical processes are described by the author as the low-level processes that cannot be easily subdivided and measured in a meaningful mathematical fashion but form the basis of all other mathematical abilities described above. Non-symbolic magnitude comparison and number line estimation are some examples of basic mathematical processes (Gilmore, 2023).

The multi-level framework recognises the involvement of a set of domain-general skills, that are independently related to each one of the levels mentioned above. A wide range of domain-general skills have been investigated with regards to mathematical learning and development, such as working memory, spatial skills, language, attention, and inhibitory control (for a discussion see Gilmore, 2023). While the body of evidence describing the role of domain-general skills on mathematics learning and development is growing, there are still a few open debates that reflect the relative novelty of this field. For example, it is still unclear to what extent some basic mathematical processes are distinct from domain-general skills and whether the impact on the overall mathematics achievement should be attributed to specific basic mathematical processes or to domain-general skills. In fact, as discussed in Gilmore et al. (2018), while in some cases the association between basic mathematical processes and overall mathematics achievement still holds after controlling for domain-general skills, there are also some exceptions. For example, several studies have demonstrated that the relationship between non-symbolic magnitude comparison (a basic mathematical process) and mathematics achievement is explained by domain-general skills such as inhibitory control, visual processing, and executive functions rather than by specific mathematical representations and processes (Coolen et al., 2022). Similarly, Simms et al. (2016) found that visuospatial skills, a domain-general skill, explained the relationship between performance on a number line task (basic mathematical process) and mathematics achievement. A parallel open debate addresses the domain-specificity of domain-general skills and whether we should think about domain-general skills differently, based on the specific domain that we are investigating – for example, see Wilkey (2023). The debates surrounding the role of various domain-general cognitive skills in mathematical development are outside the scope of this thesis, which instead focuses on the other components of the framework, i.e. basic mathematical processes, domain-specific components of mathematics and learning experiences.

Finally, the multi-level framework recognises the role of overall learning experiences, which influence the development of each domain-specific level as well as the links between the different components of the framework. As reported by Gilmore (2023), these include informal learning experiences, such as the home learning environment, and more specifically the home mathematics environment (HME), as well as more formal mathematics education, such as pedagogies, learning strategies, and resources used at school.

The arrows between the different elements of the framework represent the relationships which have been investigated so far. The most studied and recognised relationships are the ones between different specific components of mathematics and overall mathematics achievement (Gilmore, 2023). Although it is well known that the set of specific components which are associated with overall mathematics achievement changes over development and evolves with education, there is still debate around which specific components are most strongly associated with mathematics achievement and what the causes of individual differences are at different ages (De Smedt, 2022; Gilmore, 2023). Moreover, where associations have been found between basic processes and measures of overall mathematics achievement, these tend to be explained via specific components of mathematics (Gilmore, 2023). Hence, the framework does not show a direct link between the first level and the third level. As for the domain-general skills, there is evidence supporting relationships between domain-general skills and all the levels of domain-specific components (for a discussion see Gilmore, 2023). Finally, it is important to consider that it is likely that different sets of domain-general skills are related to different domain-specific components, that these relationships may change over development, and that the direct path from domain-general cognitive skills to overall mathematics achievement may exist for some cognitive skills but not for others (Gilmore, 2023).

The following subsections present an overview of the four basic mathematical processes and specific components of mathematics investigated in the studies included in the current thesis, with a specific focus on (a) their influence on specific components of mathematics and overall mathematics achievement, (b) how they change throughout development in TD populations, and (c) a brief description of the assessments that have been used to measure them in the studies cited in the following chapters. The last two subsections describe mathematics achievement and overall learning experiences.

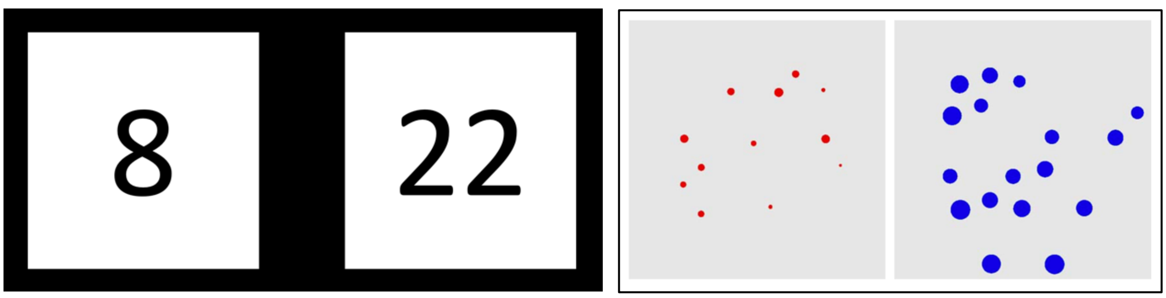
1.2.2 Numerical magnitude processes

Numerical magnitude processes refer to the representation of the magnitude of large numerosities, that could be either non-symbolic (quantities) or symbolic (numerals). These processes involve the ability to perceive, estimate, compare, manipulate quantities or numerals, and determine their order or relative size (Gilmore et al., 2018). Numerical magnitude processes are assumed to be supported by the innate Approximate Number System (Feigenson et al., 2004). However, there is an open debate on whether the Approximate Number System exists and how it should be measured (Gilmore et al., 2018).

Researchers have used several tasks to measure these processes (Gilmore, 2023; Schneider et al., 2018). The findings presented in this section are limited to the tasks used within the studies cited in the current thesis. These include symbolic and non-symbolic magnitude comparison tasks and the symbolic number line task.

In the magnitude comparison task, two numbers either in the form of Arabic digits (symbolic format) or two arrays of dots (non-symbolic format) are presented to the participant who is asked to indicate the one with the larger numerical magnitude (Figure 1.3). The difficulty of the task depends on the ratio between the numerosities presented to the participant. As the ratio between the pair approaches 1 (that is the closest the numbers are), the comparison becomes more difficult.

Figure 1.3. Example of symbolic comparison task and non-symbolic comparison task.



Using the habituation-dishabituation paradigm[5], researchers found that 6-month-old infants can discriminate between large sets of dots, provided that the sets to be discriminated differ by a large enough ratio. For example, Xu (2003) used six displays showing either 8 or 16 dots (1:2 ratio) that varied in size and position. The findings showed that 16 infants (CA range: 5 months and 20 days to 6 months and 15 days) looked longer at the novel stimulus regardless of whether they had been habituated to 8 or 16, suggesting successful discrimination of large numerosities with a 1:2 ratio. However, infants (CA range: 5 months and 17 days to 6 months and 15 days) failed to discriminate between 8 and 12 dots (2:3 ratio), a ratio that 10 months infants can discriminate (Xu, 2003). Performance of children and adults on this task is assessed through different measures that are based either on accuracy rates, with more precise representations associated with higher accuracy rates, or RTs, with more precise representations associated with faster responses. Accuracy rates and speed improve over development, with most adults being able to discriminate numerosities with a 9:10 ratio (Halberda et al., 2008). These findings should be interpreted considering the open debate discussed above and suggesting that the performance on this task might be supported by domain-general factors, such as inhibition visual processing, and executive functions (Coolen et al., 2022).

In addition to accuracy and RTs, measures derived from these may be computed to account for the ratio effect[6] or the distance effect[7] – for a discussion refer to Schneider et al. (2017). Finally, some studies report the Weber fraction (*w*), which measures the smallest numerical change to a stimulus that can be reliably detected, and in this case corresponds with the smallest ratio between two

---

[5] The habituation–dishabituation paradigm derives from the observation that infants get bored with looking at the same thing, so that when the same stimulus is repeatedly presented, this causes them to lose interest (habituation). When a stimulus that they perceive as new is presented, they get interested and dishabituation occurs. When this paradigm is used in mathematical cognition, infants are presented with the same stimulus containing, for example, a particular number of dots until they meet the habituation criterion (that usually is a set % decline in looking time, or a set number of trials), and then infants are presented with a post-habituation stimulus containing a different number of dots. Dishabituation occurs when the duration of the infant's first fixation at the post-habituation stimulus is significantly longer than the duration fixation during the habituation phase.

[6] The closer the ratio of the compared dot arrays is to 1, the more difficult it is to discriminate the dot arrays (Schneider et al., 2017).

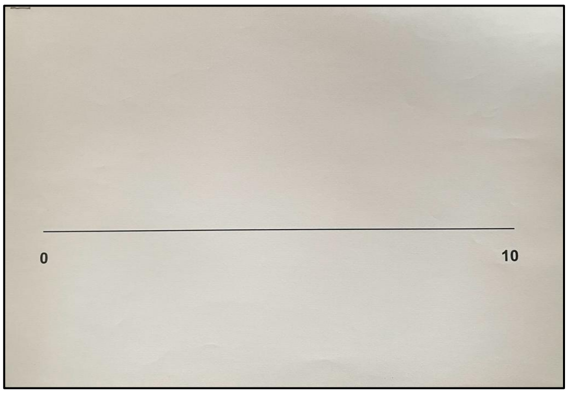
[7] Accuracy increases and RT decreases as the difference between the two numbers increases (Schneider et al., 2017).

numerosities for which a person can reliably judge the larger one (Dietrich et al., 2016). There is an open debate on whether the non-symbolic comparison task is a valid measure for assessing numerical magnitude processes due to several methodological challenges related to this task . For example, Dietrich et al. (2016) investigated whether accuracy-based measures and RT-based measures assessed performance on this task to the same extent and reported that these measures yield different results and could not be used interchangeably. Additional challenges arise from the need to control for those properties of the non-symbolic stimulus which change with numerosity, such as density and size of the dots on the screen. In fact, participants could base their response on the density of the dots rather than on their numerical magnitude and, for example, report that the larger set is the one with the higher density of dots. These challenges have drawn criticism over the last decades towards the validity of the task and the interpretation of its results (Gilmore et al., 2018). The discussions regarding the validity of this task are beyond the scope of this thesis.

Although findings vary across studies, a meta-analysis by Schneider et al. (2017) on TD population reported a reliable relationship between mathematics achievement and symbolic magnitude comparison (r = .30, 95% CI [.243, .361], k = 89) and non-symbolic magnitude comparison (r = .24, 95% CI [.198, .284], k = 195). This study found that the association between magnitude comparison skills and mathematical competence was weakly moderated by age and strongly moderated by the choice of measure of magnitude comparison and by the choice of the specific component of mathematics measured. Specifically, studies using accuracy rates (r = .316, 95% CI [.245, .387], k = 72), RTs rates (r = .269, 95% CI [.216, .322], k = 79), and w (r = .315, 95% CI [.248, .382], k = 69) lead to stronger effects than the ones reporting measures derived from RTs such as ratio effect on RT (r = .142, 95% CI [.030, .254], k = 11) and distance effect on RT (r = .135, 95% CI [.080, .190], k = 41). Moreover, stronger effects were found for mental arithmetic (r = .378, 95% CI [.321, .435], k = 52) than for written arithmetic (r = .281, 95% CI [.189, .373], k = 81) as well as the residual category which included a variety of specific components of mathematics such as number decomposition, mathematical reasoning, and geometry (r = .210, 95% CI [.159, .261], k = 52).

There are different versions of the number line task. This section focuses on the symbolic, number-to-position version, as this is the task used in the studies cited in this thesis. In the number-to-position number line task, the participant is asked to locate the position of a given number on a "number line" (Siegler & Opfer, 2003). The task can be presented in different formats such as pen-and-paper and computer-based, and stimuli can be presented using different formats, such as Arabic digits, analogue magnitude (e.g., dots), or verbally. Figure 1.4 shows the stimulus of a paper-based number-to-position task with a 0-10 bounded line.

Figure 1.4. Example of number line task.

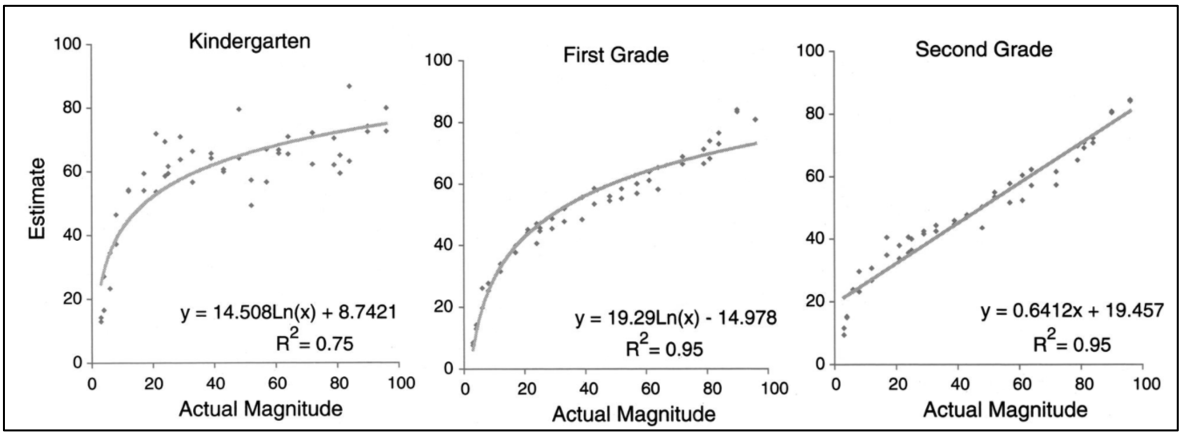


Note: 0 -10 number-to-position task with a bounded line. Starting point (0) and ending point (10) of the bounded line are labelled using Arabic digits.

Proficiency on the number line task can be measured using the percentage of correct trials, where an answer is defined as correct if it lies within a predefined interval. Alternatively, proficiency on this task can be measured using deviations of the participant's estimate from the correct position, that is based on the mean absolute difference between the correct position and the estimated position and can be expressed either in absolute terms or in terms of ratio against the number line length (Percent Absolute Error, PAE) (Schneider et al., 2018). As the PAE is a measure of error, higher values of PAE indicate poorer performance. PAE is the most frequently used measure, as it scores performance as a continuous score, compared to the percentage of correct trials, which instead is based on a dichotomous coding of correct and incorrect (Schneider et al., 2018).

A study by Siegler and Booth (2004) on TD populations showed that while younger children tend to overestimate smaller numbers and to compress larger numbers at the end of the line, older children tend to produce more accurate estimates. Figure 1.5 shows the developmental shift from a logarithmic to a more linear representation in precision of the number line responses. The chart plots the correct position on the x axis and the estimated positions on the y axis and shows that the gain in precision of number line judgments is characterised by a transition from a logarithmic representation (for younger children) to a linear one (for older children), which provides an adequate reflection of the actual numbers. These findings should be interpreted considering the open debate presented above and suggesting that the performance on this task could be explained by domain-general cognitive abilities, such as visuospatial skills (Simms et al., 2016), and considering that the development of such specific components of mathematics might be supported by a domain-general factor.

Figure 1.5. Progression of number line representation from logarithmic to linear pattern.



Note: From "Development of numerical estimation in young children" by Siegler R.S. and Booth J.L., 2004, Child development, 75(2), p. 433 (https://doi.org/10.1111/j.1467-8624.2004.00684.x). Licence number: 1460789-1.

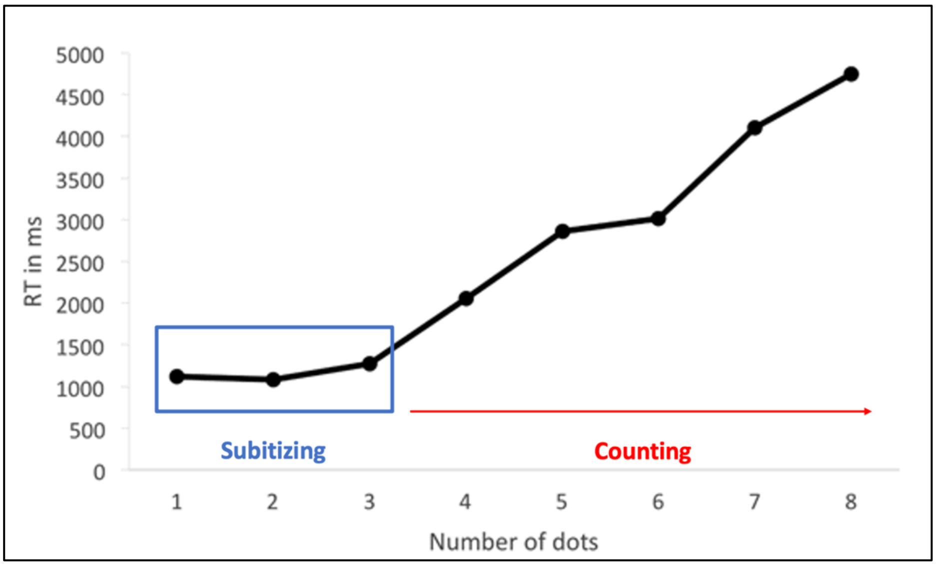
A meta-analysis by Schneider et al. (2018) showed that the performance on the number line task is associated with a broad range of specific components of mathematics. Specifically, the study reported a relationship between number line performance and counting (r = .369, 95% CI [.265, .473], k = 10) between number line performance and mental arithmetic (r = .382, 95% CI [.274, .490], k = 16) and between number line performance and written arithmetic (r = .466, 95% CI [.405, .527], k = 25).

1.2.3 Enumeration

As described by Gilmore et al. (2018), when individuals are presented with a set of items and are asked about the number of items in the set, responses tend to be fast and accurate for small numbers of items (typically 1 to 4) and to take longer and present more errors as the numerosity of the set increases. Figure 1.6 shows a typical performance on an enumeration task with primary school children

(6 to 10 years). This is characterised by small or no increase in RTs for each additional item in the so-called subitizing range (1– 3, in this example), and a sharp increase in the RTs when individuals start engaging in counting (more than 3, in this example). This pattern was described for the first time by Jevons (1871) and has been replicated since then.

Figure 1.6. Typical performance on an enumeration task.



Note: Adapted from "An introduction to mathematical cognition" by Gilmore, Camilla; Göbel, Silke M.; Inglis, Matthew. 2018, Routledge. P. 8. Licence number: 1453064.

Subitizing is the ability to enumerate quickly and effortlessly small sets of items without having to count (Kaufman et al., 1949). The term "subitize" was coined by Kaufman et al. (1949) and stems from the Latin adjective "subitus" that means sudden. Several theoretical models of subitizing have been proposed to describe this process. Some of these models rely on parallel pre-attentive mechanisms for object tracking, while others suggest that subitizing is supported by visuo-spatial working memory (Gilmore et al., 2018). Currently there is no consensus on this topic, which needs to be further investigated. A contribution to this debate is outside of the scope of this thesis.

Subitizing seems to emerge before verbal counting and to develop with age both in terms of RTs, which show increased speed, and of range, which extends from 1-3 to 1-4 or 1-5, depending on the studies (Gilmore et al., 2018).

An important distinction needs to be made between perceptual subitizing and conceptual subitizing (Clements, 1999). In fact, the process described so far refers to perceptual subitizing. However, when individuals are presented with more than 3 items arranged in a familiar pattern (e.g., dice pattern or shown of fingers), they are still able to apprehend the numerosity simultaneously and effortlessly through conceptual subitizing. For instance, Krajcsi et al. (2013) found that in TD participants (mean CA = 23.7 years) the RTs for enumerating canonical patterns up to 6 dots were similar to the RTs for enumerating 1 to 3 dots. Furthermore, Jansen et al. (2014) reported that when 4 to 5 years old TD children were presented with different dots configurations, they were more accurate in the dice pattern condition than in the random display condition. Conceptual subitizing involves higher-level abilities than perceptual subitizing, as it requires the individual to perceive the quantities as groups and to perform mental processes on them (Clements, 1999). The ability to perform conceptual subitizing makes it possible, for example, to see 6 dots in 2 groups of 3 (Özdem & Olkun, 2021). Because individuals engaging with conceptual subitizing use pattern recognition and knowledge of numbers, it has been proposed that conceptual subitizing develops with age and experience with numbers (Sarama & Clements, 2009).

Typically perceptual subitizing and conceptual subitizing are investigated using the computer tachistoscopy methodology. Computer tachistoscopy is a method used to present visual stimuli such as text or images to participants on a computer screen for a very brief duration, often measured in milliseconds (ms). This technique allows researchers to study perceptual and cognitive processes by controlling the duration of the stimulus presentation and by assessing participants' responses.

A large body of studies has reported that perceptual subitizing and conceptual subitizing are important factors for mathematical development, as they serve as a scaffold for the development of counting, calculation skills, and overall mathematics achievement (Butterworth, 2005; Clements, 1999; Kroesbergen et al., 2009; Özdem & Olkun, 2021; Reigosa-Crespo et al., 2013). For example, Reigosa-Crespo et al. (2013) assessed perceptual subitizing in 49 primary school children (mean CA = 9.3 years) and one year later they assessed students' mathematical fluency using a form with 100 arithmetic operations to be completed in 3 minutes, and overall mathematics achievement using a researcher-developed test based on the mathematical Cuban curriculum and including calculation and arithmetic word problems tasks. Their findings showed that perceptual subitizing was a significant predictor of both mathematical fluency and overall mathematics achievement and that, after controlling for other variables, perceptual subitizing predicted 7.1% of the variance observed in mathematical fluency and 9.5% of individual variability in mathematical achievement. The intervention study conducted by Özdem and Olkun (2021) involved 737 primary school students who underwent a 8-week training programme aimed at enhancing their conceptual subitizing skills. The results of this study showed a significant improvement for the experimental group from pre-test to post-test across various areas. These included enumeration skills measured using a task where participants were asked to enumerate 1-9 black dots arranged either in a random or a dice pattern as fast as they could, calculation skills measured with the TTR test (refer to Table 1.2 for a description) and mathematics achievement assessed using the Mathematics achievement test, a measure based on the mathematical Turkish curriculum and including counting, number patterns and basic calculation. Furthermore, the authors reported a long-term effect of the conceptual subitizing training on mathematics achievement, which was assessed 6 months after the conclusion of the intervention.

Counting is a sequential enumeration process characterised by a linear increase in RTs (Figure 1.6) and a decrease in accuracy rate as the number of items to be enumerated increases. Although it may seem an easy task, counting is far from trivial, and it is characterised by a high level of complexity. Learning to count takes about 4 years, from 2 to 6 (Butterworth, 2005). It has been suggested that children learn the procedural aspects of counting first, and only after that they develop the conceptual understanding of counting. Gelman and Gallistel (1978) argued that the child's development of counting involves five principles. Four of these principles focus on procedural aspects of counting and include stable order, one-to-one correspondence, abstraction, and order irrelevance. The last principle is the cardinality principle and involves the understanding that the last number word in a counting sequence represents the total number of items in the set. It is only when children follow the cardinality principle that they grasp why they are counting and they can be said to have understood the point of it, that is to determine the total numerosity of a set. Table 1.1 presents a brief description of the tasks used to assess counting abilities, which are used in the studies cited in this thesis.

Table 1.1. List and description of the tasks used in the studies cited in this thesis to assess counting abilities.

Assessment	Scale	Test description
Give a number (Wynn, 1990)	n/a	Give a number assesses one-to-one correspondence and cardinality principles. Individuals are asked to give a specific number of objects to the experimenter.
How many? (researcher-developed task)	n/a	Individuals are asked to tell how many items in a set.
Test for the diagnosis of mathematical competencies (Gregoire, Noel, & Van Nieuwenhoven, 2004)	Sub-test 1 (TEDI-MATH 1)	TEDI-MATH 1 assesses the stable order principle of counting. Individuals are asked to count forward to an upper bound, and to count backwards.
Test for the diagnosis of mathematical competencies (Gregoire, Noel, & Van Nieuwenhoven, 2004)	Sub-test 2 (TEDI-MATH 2)	TEDI-MATH 2 assesses one-to-one correspondence and cardinality principles. Individuals are asked to count both linear and non-linear sets of objects and to tell how many items they counted.
Verbal counting (researcher-developed task)	n/a	Individuals are asked to verbally count as far as they could, to count starting from a given number, and / or to count backwards.

Note: When available, the description of each assessment was retrieved from the studies which used them. Otherwise, information about the assessment was retrieved from the publisher's website. n/a: not applicable.

A large body of evidence has indicated the central role of counting and of the understanding of the cardinality principle on the development of calculation skills and early mathematics achievement (Aunola et al., 2004; Butterworth, 2005; Krajewski, 2009; Marcelino et al., 2017). For example, in their longitudinal study Aunola et al. (2004) assessed counting abilities of 194 pre-school Finnish students (mean CA = 6.25 years at T1) who were asked to count as far as they could, to count starting from a given number, to count backwards, and to solve a simple addition task using the "count on" strategy (e.g., what is the number you get when you count 5 numbers on from 2?). Mathematical performance was measured using a test which assessed knowledge of ordinal and cardinal numbers, number identification, simple word problems and basic arithmetic skills 6 times across 3 school years (preschool, year 1, and year 2). Their findings showed that the initial level of counting skills was a strong predictor of the level of mathematical performance and of its development. In fact, the results showed that the higher the level of counting abilities exhibited by the children at the beginning of the pre-school year, the higher their level of maths performance and the faster their rate of growth in mathematical performance. However, while counting is crucial in early stages of developing mathematical skills, some studies highlighted that as individuals progress and acquire more advanced and complex mathematical skills, its influence on mathematical development diminishes. For example, Marcelino et al. (2017) measured mathematical abilities of 123 children (mean CA = 6.37 years at T1) at the beginning and at the end of the school year and reported that although counting skills were highly correlated with calculation skills, they did not appear to be a good predictor of mathematical achievement measured with a test based on the Portuguese mathematics curriculum which included place value, applied problems and basic geometry components.

1.2.4 Calculation

Calculation is the ability to solve arithmetic operations (Gilmore et al., 2018). Despite the apparent simple definition, this specific component of mathematics has been defined using a vast range of terms and labels in the literature. For example, some studies use general labels such as “computation”, “arithmetic fluency”, or “arithmetic operations”, while others provide a more detailed definition, such as “written multi-digit arithmetic” or “mental calculation”. While there is conflicting evidence regarding infants’ capacity to understand and perform arithmetic operations, the evidence indicating that pre-schoolers can perform simple additions and subtractions before receiving formal arithmetic instruction in school is robust (Gilmore et al., 2018). Levine et al. (1992) developed a non-verbal calculation task to assess children’s calculation skills (refer to Table 1.2 for a description). The study assessed non-verbal calculations of 60 American children aged between 4 and 6.5 years. Their findings revealed that participants were able to perform the calculation task, with increasing accuracy as age increased. This was interpreted as an indication of children’s ability to perform non-verbal calculations. These results were replicated with participants as young as 2 years old (Huttenlocher et al., 1994). Furthermore, Jordan et al. (2007) reported that children’s performance on this task in 277 American pre-schoolers was moderately correlated ($r = 0.52$) with their mathematical achievement measured at the end of the first year of school with the Woodcock–Johnson Test of Achievement (refer to Table 1.4 for a description) and it also emerged as a significant predictor of mathematical achievement.

Depending on the aim of the study and on the age of the participants, calculation skills can be assessed through different tasks and assessments. This component has been measured using a wide range of standardised assessments, curriculum-based assessments, and research-made tests in different studies. Furthermore, while some researchers are interested in the procedural aspect of calculations, others are interested in measuring the conceptual understanding and focus on recording the strategy used by participants when performing the task. Table 1.2 presents a list and brief description of the assessment tools used to assess calculation abilities in the studies cited in this thesis.

Table 1.2. List and description of the assessment tools used in the studies cited in this thesis to assess calculation abilities.

Assessment	Scale	Test description
Arithmetic number facts test (TTR; De Vos, 1992)	n/a	TTR measures the ability to use number facts. It consists of five subtests, each one containing 40 items: addition, subtraction, multiplication, division, and mixed exercises (time: 1 min).
Cognitive developmental skills in arithmetic (Desoete & Roeyers, 2006)	Procedural calculation (CDR-P)	CDR-P consists of 90 items presented in a number problem format, such as number splitting and addition / subtraction by regrouping – for example, “12 – 9 =__”.
Kaufman test of educational achievement (Kaufman & Kaufman, 1985)	Math computation (KTEA-MC)	KTEA-MC measures the ability to solve written computational problems including addition, subtraction, multiplication, division, and algebra.
Non-verbal calculation task (Levine et al., 1992)	n/a	The non-verbal calculation task involves presenting to the participant sets of items that are then transformed either by adding or removing elements. The participant sees the initial set and the number of elements that are added or removed, but not the final set. The participant is asked to construct an array that contains the number of elements in the final set. One addition problem (1 + 1) and one subtraction problem (2 - 1) are used as demonstration.
Test of mathematical abilities (Brown et al. multiple versions)	Computation (TOMA-C)	TOMA-C consists of 30 items that range in difficulty from one-digit addition to writing in scientific notation. The tasks assess knowledge of basic operations, advanced fractions, decimals, and percent.
Wechsler individual achievement test (Wechsler, multiple versions)	Numerical operations (WIAT-NO)	WIAT-NO consists of 54 items. It measures the ability to identify and write numbers, to count, and to solve written calculation tasks including addition, subtraction, multiplication, division, and simple equations.
Wide range achievement test (Jastak & Wilkinson, multiple versions)	Computations (WRAT)	WRAT measures an individual’s ability to perform basic mathematics computations through tasks involving identifying numbers, counting, and solving written calculations.
Woodcock–Johnson test of achievement (Woodcock et al., multiple versions)	Calculation (WJ-C)	WJ-C measures the ability to perform mathematical computations such as addition, subtraction, multiplication, division, as well as geometric, trigonometric, logarithmic, and calculus operations.
Woodcock–Johnson test of achievement (Woodcock et al., multiple versions)	Fluency (WJ-F)	WJ-F measures the ability to solve simple addition, subtraction, and multiplication facts quickly (time: 3 mins).

Note: When available, the description of each assessment was retrieved from the studies which used them. Otherwise, information about the assessment was retrieved from the publisher’s website. n/a: not applicable.

1.2.5 Arithmetic word problems

Arithmetic word problem skills have frequently been the focus of research studies on mathematical abilities (Gilmore et al., 2018). To correctly perform this task participants must create a representation of the problem, extract the relevant information, choose the appropriate operation and finally, correctly perform the calculation (Gilmore et al., 2018). Consequently, participants' performance results from a combination of the mental model used to construct the problem situation, their procedural skills in performing the relevant operation, and the influence of their domain-general skills such as inhibition of non-relevant information or working memory skills (Gilmore et al., 2018).

Table 1.3 provides an overview of all the assessment tools used in the studies cited in this thesis to assess arithmetic word problem skills.

Table 1.3. List and description of the assessment tools used in the studies cited in this thesis to assess arithmetic word problem abilities.

Assessment	Scale	Test description
Cognitive developmental skills in arithmetic (Desoete & Roeyers, 2006)	Linguistic Mental representation Contextual (CDR-LMC)	CDR-LMC includes word problem tasks such as "1 more than 5 is __" and "Wanda has 47 cards. Willy has 9 cards less than Wanda. How many cards does assessment Willy have?"
Kaufman test of educational achievement (Kaufman & Kaufman, 1985)	Math applications (KTEA-MA)	KTEA-MA measures the ability to solve story problems presented orally and to interpret tables and graphs.
Mathematical problem instrument (MPI; Mulligan & Mitchelmore, 1997)	n/a	MPI consists of eight arithmetic word problem tasks which focus on multiplication and division, such as: "There are two tables, and four people at each table, how many people are there in total?". Participants are provided with a booklet and with cubes they could use to solve the tasks.
Mathematical word problem solving (MWPS; Griffin & Jitendra 2009)	n/a	MPWS consists of 12 arithmetic word problem tasks.
Test for the diagnosis of mathematical competencies (Gregoire, Noel, & Van Nieuwenhoven, 2004)	Sub-test 5.1 (TEDI-MATH 5)	TEDI-MATH 5 includes six visually supported addition and subtraction word problems (e.g., ''Here you can see two red balloons and three blue balloons. How many balloons are there altogether?'')
Test of mathematical abilities (Brown et al., multiple versions)	Story problem (TOMA-SP)	TOMA-SP consists of 25 arithmetic word problem tasks arranged in an easy to difficult order.
Wechsler (Wechsler, multiple versions)	Arithmetic	Arithmetic scale consists of 22 arithmetic word problem tasks.
Wechsler individual achievement test (Wechsler, multiple versions)	Mathematical reasoning (WIAT-MR)	WIAT-MR measures the ability to count, identify geometric shapes, and solve single-step and multistep word problems.
Woodcock–Johnson test of achievement (Woodcock et al., multiple versions)	Applied problems (WJ-AP)	WJ-AP measures the ability to analyse and solve maths problems of increasing difficulty. Most items require a student to listen to the problem that is verbally presented, recognise the mathematical procedure that must be followed, and perform the appropriate calculations.

Note: When available, the description of each assessment was retrieved from the studies which used them. Otherwise, information about the assessment was retrieved from the publisher's website. n/a: not applicable.

1.2.6 Overall mathematics achievement

Overall mathematics achievement represents the highest level of the hierarchical framework developed by Gilmore (2023). Overall mathematics achievement refers to an individual's overall attainment in mathematics and requires an understanding of how different components relate to each other and the ability to integrate new knowledge with existing knowledge and skills (Gilmore, 2023). Typically, overall mathematics achievement is measured by broad measures based on the school curriculum, by composite standardised assessments, or by standardised assessments including a wide range of mathematical tasks. These tools usually incorporate a variety of specific mathematical components such as enumeration, calculation, and arithmetic word problem skills. Table 1.4 provides an overview of all the assessment tools used in the studies cited in this thesis to measure overall mathematics achievement.

Table 1.4. List and description of the assessment tools used in the studies cited in this thesis to assess overall mathematics achievement.

Assessment	Scale	Test description
Differential ability scales (Elliott, C. D., 1990)	Basic number skills (DAS)	DAS measures number concepts, calculation, and arithmetic word problems.
KeyMath diagnostic assessment (KeyMath; Connolly, 2007)	n/a	KeyMath measures basic number concepts, arithmetic operations, and arithmetic word problems.
Peabody individual achievement test (Dunn & Markwardt, 1970)	Mathematics (PIAT)	PIAT consists of 100 multiple choice items which test knowledge and application of mathematical concepts and facts.
Star maths test (STAR; Renaissance, 2019)	n/a	STAR is a computer adaptive assessment which assesses various mathematical concepts, including number sense, operations, algebra, geometry, measurements, data analysis, and probability.
School performance test (SPT; Knijnik, Giacomoni, & Stein, 2013)	Math	SPT is a Brazilian standardised test to measure school performance. The Math scale measures calculation and arithmetic word problem skills.
Test of early mathematics ability (TEMA; Ginsburg & Baroody, 2003)	n/a	TEMA consists of 72 items assessing counting, reading numbers, writing numbers, comparing numbers and quantities, calculation, and arithmetic word problems abilities.
Test of mathematical abilities (Brown et al., 2013)	Maths ability index (TOMA)	TOMA consists of 145 items assessing number concepts, computation, mathematics in everyday life, word problems, and the attitude toward maths.
Wechsler individual achievement test (Wechsler, multiple versions)	Mathematics composite (WIAT)	WIAT score is a combination of the total scores on Numerical operations and Mathematical reasoning scales.
Woodcock–Johnson test of achievement (Woodcock et al., multiple versions)	Broad math (WJ)	WJ score is a combination of the total scores on the Calculation, Math Fluency, and Applied problems scales.

Note: When available, the description of each assessment was retrieved from the studies which used them. Otherwise, information about the assessment was retrieved from the publisher's website. n/a: not applicable.

1.2.7 Learning experiences

The multi-level framework of mathematical cognition proposed by Gilmore (2023) recognises the influence of learning experiences on all the domain-specific components included in the framework. Overall learning experiences include any child-carer interaction with mathematical content, also known as home mathematics environment (HME) which is investigated in the study presented in Chapter 3, as well as more formal aspects of mathematical education, such as pedagogical methods, teaching strategies and learning resources used at school (Gilmore, 2023), some of which are investigated in the study presented in Chapter 5.

HME has started to receive increasing attention in the field of mathematical cognition over the last decade (Daucourt et al., 2021). Most of the HME studies focus on TD populations and either describe the HME or explore its relationship with mathematical development. The results reported by these studies are often conflicting. This is due to the fact that, as reported in the meta-analysis conducted by Daucourt et al. (2021), the HME literature is characterised by a lack of standardisation in how the HME is defined and measured. In fact, while all the studies conceptualise HME as multifaceted construct consisting of various components, there is a lack of consensus on the specific components that should be included in the definition of this construct, as well as on how they should be measured, with different studies employing different methodologies, such as direct observations, interviews or surveys (Daucourt et al., 2021). Additionally, contradictory findings on the role of the HME on the overall mathematical development may be explained by the assessment of different specific components of mathematics in order to measure overall mathematics achievement and by the use of different assessment tools to measure such components (Mutaf-Yildiz et al., 2020).

When defining HME, a common thread is the emphasis on the child-carer interaction with mathematical content (Levine et al., 2010), and most studies include both structural and functional indicators (Mutaf-Yildiz et al., 2020). Structural indicators include factors such as the availability of learning resources, the frequency of home learning maths experiences, and parents “maths talk”, that is the maths language used by the carers. Functional indicators include factors such as family characteristics, quality of the carer-child interactions such as the participation in the learning activity, and parental socioemotional factors about maths, such as parent's maths-related attitudes and levels of anxiety, beliefs, and expectations for their child's maths achievement. With reference to structural indicators, the taxonomy proposed by LeFevre et al. (2009), is commonly used to describe the type of mathematical learning activities that happen at home. This categorization suggests that children at home can be exposed to both formal maths instruction activities, and informal maths experiences. Formal activities explicitly focus on mathematical abilities and are used by carers for the specific purpose of developing mathematical skills, such as practicing number names or simple sums. On the other hand, informal activities consist of real-world tasks during which maths teaching happens without an explicit purpose and the acquisition of number skills is likely to be incidental, such as playing card or board games which involve numbers or the use of dice, reading clocks, and applying maths concepts to everyday activities like cooking and shopping.

The meta-analyses conducted by Daucourt et al. (2021) involved 51 studies and reported a small and positive significant correlation between HME and overall mathematics achievement (r = .13, 95% CI [.09, .17], k = 64). When investigating whether different structural and functional components of HME moderated the strength of this relationship, the meta-analysis reported similar correlations for formal activities (r = .10, 95% CI [.03, .17], k = 29) and informal activities (r = .12, 95% CI [.08, .16], k = 28), and a slightly higher correlational value for parental expectations (r = .22, 95% CI [.09, .35], k = 14). The study reported similar correlational values for different specific components of mathematics identified by the authors as numbering, which included activities supporting counting, subitizing, and estimation (r = .12, 95% CI [.09, .15], k = 12) and arithmetic operations (r = .14, 95% CI [.07, .20], k = 21). Conversely, the findings did not show a significant effect size for the correlation between HME, and the specific component of mathematics identified by the authors as "relations", that referred to activities including quantity comparison, number comparison, number naming, ordinality, and number line sequencing (r = .07, 95% CI [-.002, .14], k = 16). The results showed that chronological age (CA) was a significant moderator, with younger (i.e., preschool/kindergarten) samples showing a stronger association between the HME and children's math achievement (r = .15, 95% CI [.11, .19], k = 50) than primary/secondary grade samples (*r* = .06, 95% CI [-0.01, .13], *k* = 22). Finally, the study reported non-significant correlations for socioeconomic status (SES) and for parental education. The authors stated that the absence of a statistically significant correlation for SES could be attributed to the limited number of studies involving high- and/or low-SES samples.

1.3 Populations studied

This thesis investigates mathematical abilities and development of individuals with Down syndrome (DS), Williams syndrome (WS), and autism. All these conditions are included in the category "neurodevelopmental disorders" of the Diagnostic and Statistical Manual of Mental Disorders (DSM), text revision of fifth edition (American Psychiatric Association, 2022). The category "neurodevelopmental disorders" is used to describe a group of conditions which typically manifest early in development, often before the individual reaches school age (American Psychiatric Association, 2022). While autism is described separately under its own distinct diagnostic sub-category "Autism Spectrum Disorder (ASD)"[8], DS and WS fall under different sub-categories of the DSM, depending on the individual's specific profile.

All these conditions might be associated with intellectual disability (ID), a NDC which includes both intellectual and adaptive functioning deficits in conceptual, social, and practical domains (American Psychiatric Association, 2022). Hence, the definition of ID provided by the DSM considers IQ scores in combination with other person-specific factors such as the level of support needed by the individual to function in their daily life (Fodstad et al., 2020). In contrast, in psychology research, ID is typically defined based on IQ scores only and is generally defined with IQ scores lower than 70 (Spaniol & Danielsson, 2022). However, the specific threshold used to define ID can vary from one study to another.

[8] The term "autism" has been chosen over the diagnostic term "autism spectrum disorder" wherever possible to reflect the view that autism is a natural part of variation in the human population rather than a disfunction (Fletcher-Watson & Happé, 2019). For the same reason the term "neurodevelopmental condition" is used in place of the term "neurodevelopmental disorder". The identity-first phrasing "autistic people" is also used, as it is reported to be the preferred term by the autism community (Kenny et al., 2016).

DS is the most common genetic syndrome associated with the neurodevelopmental disorders of ID (Bull, 2020). Studies considering IQ scores show a slower development in children with DS than in TD peers, with the gap between the two populations increasing with age (Onnivello et al., 2022). The severity of ID among individuals with DS falls on a wide spectrum, with IQ scores that range from 30 to 70 and average IQ scores of 50 (Grieco et al., 2015). Moreover, as reported by Thomas et al. (2020), there is marked individual variability with around 80% of individuals reporting a moderate ID, some falling in the severe range, and others within the normal range.

WS is a genetic syndrome associated with developmental delay[9] which typically leads to mild-to-moderate ID (Kozel et al., 2021). In fact, although a few individuals have severe ID or, at the other extreme, average intellectual ability, 75% of older children and adults have mild-to-moderate ID (Lashkari et al., 1999). The systematic review conducted by Martens et al. (2008) on 47 studies investigating overall intelligence in WS reported that across 46 of the included studies the average Full Scale IQ (FSIQ) score ranged from 42 to 68 (mean = 55). The remaining study reported a FSIQ pertained to a single 4-year-old participant who obtained a FSIQ score of 82 after receiving two years of intensive language therapy.

Finally, autism is a neurodevelopmental disorder which commonly co-occurs with other conditions, that, among others, include ID (Lord et al., 2020). As reported by Fodstad et al. (2020), through the years, the incidence of co-occurrence of autism and ID has changed from more than 70% in the early 2000s to approximately 30%. The decline in co-occurring diagnoses has been attributed to several factors including clearer diagnostic criteria for both NDCs and the development of more effective early-age assessment methods (Fodstad et al., 2020).

In summary, while some degree of ID, typically defined with IQ scores lower than 70 and with deficits in the level of adaptive functioning, is prevalent among

[9] The diagnosis of developmental delay is reserved for individuals under the age of 5 years when the clinical severity level cannot be reliably assessed due to, for example, the child being too young to participate in a standardised assessment. This category is diagnosed when an individual fails to meet expected developmental milestones in several areas of intellectual functioning, and requires reassessment after a period of time (American Psychiatric Association, 2022).

individuals with DS and WS, autistic individuals can be diagnosed with the co-occurring condition of ID, but this is not typically the case.

The following sections describe the main characteristics of each condition, including causes, prevalence, diagnosis, and overall cognitive profile. Moreover, each section provides a synthesis of the literature investigating mathematical abilities and mathematical development for each population.

1.3.1 Down syndrome

Down syndrome (DS), also called Down's syndrome, is a genetic syndrome that arises from extra chromosomal material. The most common type of DS is trisomy 21 and is present in 95% of individuals with DS. It occurs when the entire chromosome 21 has an extra copy (Antonarakis et al., 2020). However, DS also occurs when only a segment of chromosome 21 has three copies (partial trisomy 21, rare type), when chromosomal material from chromosome 21 is rearranged (translocation DS, present in ~5% of individuals with DS), and when only a proportion of the cells in the body have an additional copy of chromosome 21 (mosaic DS, present in ~2% of individuals with DS) (Antonarakis et al., 2020). Mosaic DS and partial trisomy 21 are usually associated with fewer clinical features of DS (Bull, 2020). DS was first described by John Langdon Down in 1866, who identified the syndrome as a distinct and unique medical condition. The discovery of a link between chromosome 21 and the DS phenotype was first reported in 1959 by Jérôme Lejeune and relied on the work initiated and directed by Marthe Gautier (Gautier & Harper, 2009).

DS occurs in approximately 1 of 800 births worldwide (Bull, 2020), and according to the NHS England (2022) the prevalence in England was approximately 1 in 873 live births in 2020.

DS can be identified during pregnancy with diagnostic tests through amniocentesis or chorionic villus sampling, or after birth by direct observation and genetic testing on a blood sample from the baby (Down's Syndrome Association, 2020) by mean of the fluorescence in situ hybridization of chromosome 21. Also, in developed countries, laboratory-based screening tests are offered to pregnant women as part of routine antenatal care. Screening is a way to identify the chance that a baby may have DS, which is calculated on the basis of the measurement of defined markers and of specific mother's details, such as the mother's age (Down's Syndrome Association, 2020). The tests use blood samples taken from the mother and measurements taken from ultrasound scans at the first and at the second trimester of pregnancy. More recently the use of sequencing of cell-free DNA in maternal serum has been introduced as an additional diagnostic test (Antonarakis et al., 2020).

Even if each individual with DS has a distinct set of strengths and challenges which requires different levels of medical input and social care throughout their life, there are several health problems that are more common in individuals with DS than in the general population (Antonarakis et al., 2020). These include congenital heart defects, obstructive sleep apnoea, thyroid disease, dementia (which is the proximal cause of death on 70% of older adults with DS), epilepsy, gastrointestinal disease, hearing and vision problems, and haematological disorders including leukaemia (Antonarakis et al., 2020). Moreover, individuals with DS show a predisposition to autoimmune diseases, such as coeliac disease and Type 1 diabetes (Antonarakis et al., 2020). While some of these medical conditions require immediate intervention at birth and others require lifelong surveillance (Bull, 2020), they all affect the quality of life of individuals with DS as well as development and cognitive functions, such as learning.

Generally, the cognitive profile of DS has been associated with better non-verbal than verbal skills, with language considered to be the greatest domain of vulnerability (Karmiloff-Smith et al., 2016). In fact, individuals with DS are most likely to have more pronounced language and verbal memory challenges, and relatively stronger non-verbal abilities and implicit memory skills (Onnivello et al., 2022). However, high interindividual variability has been registered in the IQ scores (Karmiloff-Smith et al., 2016; Onnivello et al., 2022). Moreover, recent studies have begun to emphasise that the DS cognitive profile seems to present greater complexity than previously depicted by the literature. For example, a recent study by Onnivello et al. (2022) identified three different cognitive profiles in DS, indicating multiple profiles within the DS population. This study explored the cognitive profile of 77 children and adolescents with DS (mean CA = 11.2 years) and investigated verbal and non-verbal intelligence and interindividual variability. The findings showed that the cognitive profile of the whole sample was characterised by similar scores in the verbal and non-verbal domains, which by itself was a surprising finding. Moreover, a cluster analysis revealed three different cognitive profiles, highlighting the great variability within the syndrome, and suggesting that DS can express multiple cognitive profiles (Onnivello et al., 2022). The group showing the lowest IQ scores, had the typical profile described in the literature, with higher non-verbal than verbal intelligence. The group reporting intermediate IQ scores showed greater verbal than non-verbal intelligence. Finally, the group showing the highest IQ scores reported equally high scores in the verbal and the non-verbal domains.

Similar high interindividual variability has been registered in the syndrome for other cognitive domains. For instance, a study by Deckers et al. (2019) reported that the language domain, usually described as a domain of vulnerability, seems to be marked by within-domain strengths and within-domain weaknesses. In fact, this study reported that 36 children with DS aged between 2 and 7 years understood more words than they were able to pronounce, indicating better receptive skills than expressive skills.

Table 1.5 summarises the literature reviews that have been conducted on mathematical abilities in DS. These include one unpublished doctoral thesis, two narrative reviews and two systematic reviews. The reviews that reported the age range of the participants who took part in the studies cited (*n* = 2) show a wide range (1 – 35 years).

The findings reported on numerical magnitude processes measured through the non-symbolic magnitude comparison task in DS yielded contrasting results. The narrative review by Brigstocke et al. (2008, p. 78) reported that at the time of writing there were “suggestions that pre-verbal number sense system may show atypical development in DS” but that this conclusion needed to be treated with caution due to the small sample size (*n* = 16) of the only study the authors were referring to. The literature review by Onnivello et al. (2019, p. 266) concluded that because of methodological differences between the studies it was “not clear whether approximate number system works in line with mental or chronological age in individuals with DS”. The systematic review by Porter (2019, p. 145) reported that although variability between studies and within the same sample need to be taken into consideration, “there is a great degree of agreement in the reviewed studies that the approximate number system of children with DS is relatively intact”. Moreover, the literature review by Onnivello et al. (2019) reported findings of studies on numerical magnitude processes measured through the number line task. The review highlighted that in DS the development of number line skills “seems to be more correlated with MA” than with CA (Onnivello et al., 2019, p. 270).

Regarding enumeration skills, most of the reviews focused on counting abilities and all of them reported that individuals with DS show some understanding of counting principles. Paterson (2000, p. 188) reported that “children with DS could master the principles of counting”. Brigstocke et al. (2008, p. 78) reported that even if findings from research showed that learning to count was difficult for children with DS, there was “no evidence that the development of counting followed a qualitatively different path to that seen in younger TD children”. Moreover, the authors argued that the poor counting performance in DS reported by initial studies should be attributed to the level of complexity of the language used by the experimenter when presenting the task and when providing feedback to the participants. Finally, Onnivello et al. (2019, pp. 271-272) concluded that individuals with DS “perform poorly in counting tasks, show some understanding of the counting principles, and show difficulties with subitizing”.

Lastly, findings reported on calculation skills in DS were consistent. Paterson (2000, p. 188) reported that “children with DS may have difficulties with more complex arithmetic tasks”, while Onnivello et al. (2022, p. 270) stated that calculation is “a particularly difficult area for individuals with DS”.

The systematic review by King et al. (2017) summarised findings from studies which compared mathematical abilities of children and adolescents with DS with another group of individuals without DS, with or without learning disabilities. The authors reported that the included studies (*n* = 8) mainly focused on early numeracy skills of young children. Moreover authors reported that “outcomes from the comparison studies revealed few differences that would indicate a distinct mathematical phenotype for individuals with DS” (King et al., 2017, p. 217). However, the authors also emphasised that the limited number of studies included in the review affected their ability to make meaningful interpretations from the findings.

Some of these reviews considered the role of domain-general skills on mathematics achievement, such as, for example, executive functions (Onnivello et al., 2022) and language (Brigstocke et al., 2008). However, these reviews made no mention of studies reporting findings related to either the investigation of the learning environment or the investigation of the impact of learning experiences on mathematics achievement in DS.

Table 1.5. List of literature reviews on mathematical abilities in the DS population.

Author	Type of review	Studies included	Main Findings	Age range
Paterson (2000)	Unpublished narrative review (PhD thesis)	5 studies (*)	Could master counting. May have difficulty with complex calculation.	n/r
Brigstocke et al. (2008)	Narrative review	20 studies (*)	Magnitude comparison may show atypical development. No evidence that the development of counting followed a qualitatively different path to that seen in younger TD children.	n/r
King et al. (2017)	Systematic review	8 studies published between 1989 and 2016	Could not identify any pattern of differential mathematics performance for DS.	3 – 20 years
Onnivello et al. (2019)	Narrative review (book chapter)	19 studies (*)	Not clear if magnitude comparison in line with MA or CA. Development of number line skills is more correlated to MA than to CA. Some understanding of counting principles, with poor performance. Show difficulties with subitizing. Calculation is a difficult area.	n/r
Porter (2019)	Systematic review	8 studies published between 2001 and 2018	Approximate number system of children with DS is relatively intact.	1 – 35 years

Note. n/r: not reported. (*) The count of the studies was not explicitly reported in the review, but it was manually computed by the author of this thesis based on the in-text citations.

In summary, findings summarising performance on basic mathematical processes and specific mathematical components appear to be conflicting. However, this may be attributed to either the different RQs of the authors when synthesising previous literature or to the inconsistent language used in different reviews to describe both the mathematical components investigated and the findings of the studies. Moreover, the lack of details reported in the reviews of literature when describing different basic processes and specific components of mathematics may lead to a misinterpretation of the information conveyed. These inconsistencies and vagueness of the language used reduces clarity, complicates the comparison of results, and limits the identification of genuine trends or patterns across reviews.

1.3.2 Williams syndrome

Williams syndrome (WS), also known as Williams-Beuren Syndrome, is a rare genetic syndrome caused by the deletion of one copy of 26–28 genes on chromosome 7q11.23 (Morris et al., 2020). WS was first identified in 1961 by the cardiologists Williams, Barrat-Boyes, and Lowe, through a study of 4 children with aortic stenosis (a specific heart defect involving narrowing of the arteries), who also had learning disabilities and distinctive facial appearances. A year after this report, German physician A. J. Beuren described 3 new patients with the same presentation. The genetic basis of WS was first uncovered in 1993 (Ewart et al., 1993).

As reported by Kozel et al. (2021), the most widely cited epidemiological study for WS is one conducted in Norway which reports a prevalence of 1 in 7,500 live births (Stromme et al., 2002), a higher prevalence than that cited in many non-epidemiological sources which instead report a prevalence between 1 in 10,000 and 1 in 20,000 (Martens et al., 2008; National Organization for Rare Disorders, 2006). Lashkari et al. (1999) suggested that the differences in prevalence estimates might be explained by the substantial minority of individuals with the genetic markers of WS which lacks the characteristic facial features, or the intellectual disabilities, and which are not immediately recognised as having the syndrome. In fact, as there is no new-born screening for WS, clinical consideration of the diagnosis is prompted by the presence of suggestive signs and / or symptoms and confirmed by genetic testing. The most widely used laboratory methods available to detect the 7q11.23 microdeletion include fluorescence in situ hybridization and chromosomal microarray analysis, which are widely used in the high-income countries due to their high cost (Kozel et al., 2021). The age of diagnosis of WS has trended towards younger ages over the past decades, especially in high-income countries with greater availability of molecular diagnostic testing. For example, in cohorts from the USA and Australia, the median age of diagnosis decreased by more than 2 years to around 1 year of age since the

1980s (Kozel et al., 2021). However, series from other countries indicate that diagnosis is often still established during childhood rather than in infancy, even with access to molecular confirmation (Kozel et al., 2021).

As reported by Morris et al. (2020), there is a high prevalence of a variety of health problems that affect the quality of life of individuals with WS as well as their development and cognitive and social function through their lifespan. Most children with WS have cardiovascular anomalies (which are the major source of mortality in WS), one third of which requires surgical correction. Infants and toddlers with WS often have infantile hypercalcemia and difficulty with feeding (Morris et al., 2020). Chronic constipation and abdominal pain are common lifelong issues (Morris et al., 2020). Mental health and behavioural problems may include hypersensitivity to sound, hyperactivity and attention disorders, difficulties with emotional regulation, non-social anxiety, and sleep disorders, that affect more than 50% of individuals with WS (Morris et al., 2020).

The cognitive profile of WS is often described at group level as characterised by stronger verbal than non-verbal abilities (Kozel et al., 2021; Martens et al., 2008). A recent study by Farran et al. (2024) reported legacy data from cross-sectional and longitudinal experiments conducted in seven different laboratories in the United Kingdom (UK) that assessed in children and adults with WS verbal abilities using as a proxy scores from the British Picture Vocabulary Scale and non-verbal abilities using as a proxy scores from the Raven' s Coloured Progressive Matrices and from the Pattern Construction subtest of the British Ability Scales . The findings from this study confirmed the characteristic cognitive profile of stronger verbal than non-verbal abilities in WS. However, the study also reported high levels of individual differences, which were attributed to the verbal cognitive abilities. A significant variability across all cognitive domains in WS has also been reported in the review by Kozel et al. (2021), but the authors could not identify the contributing factors. Moreover, there are a few studies that identified within-domain strengths and weaknesses in WS. For example, Mervis and John (2010) identified a distinctive profile in this population within the language domain, with a relative strength in concrete vocabulary and grammatical abilities and considerable weakness in both relational language and pragmatics.

The investigation of mathematical abilities in WS is an under-researched but growing area. A good indicator of this is the absence of the description of the mathematical profile and specific recommendation for interventions in the comprehensive reviews cited so far. For example, the article by Mervis et al. (2000) provides a thorough characterization of the WS cognitive phenotype with a special focus on ID, language and literacy, memory, and executive function along with suggested interventional approaches, without mentioning mathematical skills. Moreover, the comprehensive review by Kozel et al. (2021), only reported that little is known about mathematical abilities in WS, and only cited one study to support this statement. The clinical report by Morris et al. (2020) provides recommendations suggesting that children with WS should be referred to an early intervention program for physical, occupational, and speech therapy evaluation and treatment. However, the report does not explicitly mention mathematical support. It also emphasises the need to develop an appropriate educational plan for school-age children and to provide vocational training and social skills training for adults, yet mathematics support is not specifically mentioned as a critical factor to support their level of independence.

Table 1.6 summarises the literature reviews that have been conducted on mathematical abilities in WS. So far, this topic has been reviewed in two unpublished doctoral thesis, two narrative reviews and one systematic review.

The findings reported on numerical magnitude processes yielded contrasting results. The literature review conducted by O'Hearn and Luna (2009, p. 11) summarised findings on numerical magnitude processes as “atypical in WS, throughout development”, while Van Herwegen et al. (2020, p. 9) reported that the “development of larger magnitude discrimination is impaired[10] and remains low across development”. A few reviews reported findings from studies investigating the discrimination of small numerosities. Ansari (2003) reported the findings from a study that used the habituation–dishabituation paradigm[5] to investigate discrimination of small numerosities (2 vs 3) in infants with WS. The findings from this study were interpreted by the author as suggesting that infants with WS “seem to possess some number-relevant processing capacities” (Ansari, 2003, p. 65).

[10] In this study the term “impaired” is used to describe skills of individuals with WS that are below CA-matched controls. The term “spared” is used to describe abilities of individuals with WS in line with CA-matched controls or with those of a similar overall cognitive ability such as individuals with DS. The term “atypical” is used to describe abilities of individuals with WS that are not in line with the abilities measured in TD population (Van Herwegen & Simms, 2020, p. 4).

The literature review conducted by O'Hearn and Luna (2009, p. 11) reported that "the precise representation of small numbers appears to be relatively typical in infancy, but limitations become evident by maturity, suggesting a truncated developmental trajectory". Moreover, the literature review by Van Herwegen et al. (2020) cited one study on numerical magnitude processes measured through the symbolic number line task (0 – 1,000). The findings showed that performance in WS was best predicted by a logarithmic function, hence suggesting inaccurate representations, which was interpreted by the authors as impaired[10] (Van Herwegen et al., 2020, p. 9). The authors commented that because of the wide range of numbers used in this task, it was "difficult to conclude whether participants (with WS) have impaired number line performance per se, or that the task in this study was beyond the boundary of their numerical knowledge" (Van Herwegen & Simms, 2020, p. 7).

Regarding enumeration skills, most of the reviews focused on counting abilities. The literature review included in the unpublished doctoral thesis by Paterson (2000) and Ansari (2003) cited a study where children and adults with WS (CA range: 8 – 33 years) were asked to count objects and sounds. The findings showed no differences in the counting performance between the clinical and comparison groups (TD group aged 4-5 years old) and were interpreted as suggesting that individuals with WS "seem to follow all the counting principles used by TD children" (Paterson, 2000, p. 184). Moreover the reviews by Paterson (2000) and by Ansari (2003) reported that individuals with WS failed to understand number conservation, that refers to the understanding that change in the physical attribute of an array (that is, for example, the space between the items) does not affect numerosity. The review by Van Herwegen et al. (2020, p. 9) concluded that "counting skills and understanding of the counting principles are not really a strength for people with WS but counting abilities are delayed and may develop atypically[10]".

Finally, a few reviews cited studies that investigated calculation skills. Both the reviews by Paterson (2000) and by Ansari (2003) reported a longitudinal study investigating calculation abilities measured with two versions of the Wechsler assessment. This study indicated that calculation performance did not improve between childhood and adulthood in individuals with WS but instead remained at a plateau, in contrast to improvements in FSIQ scores. These findings, although limited, were interpreted as suggesting that “calculation abilities of individuals with WS do not exceed those expected for TD children aged 8 years” (Ansari, 2003, p. 65). Van Herwegen et al. (2020, p. 9) concluded instead that because of the small number of studies investigating this specific component (n = 2) and their small sample size (n_1 = 1, n_2 = 8) “the true arithmetical ability of people with WS is currently unknown”.

Several reviews investigated the role of domain-general skills on mathematics achievement. For example, the literature review conducted by O'Hearn and Luna (2009) highlighted how relatively strong verbal abilities, and memory skills could support mathematical development in WS. The narrative review by Dowker (2020) emphasised that in WS the mathematical abilities related to a spatially represented mental number line are more markedly impaired than those aspects of mathematics more dependent on verbal recall. Finally, the systematic review conducted by Van Herwegen and Simms (2020) suggested that impaired attention and impaired visuo-spatial abilities early on in life could impact the development of domain-specific abilities, such as numerical magnitude processes and counting skills. None of these reviews cited studies reporting findings investigating learning experiences in WS.

Table 1.6. List of literature reviews on mathematical abilities in the WS population.

Author	Type of review	Studies included	Main findings	Age range
Paterson (2000)	Unpublished narrative review (PhD thesis)	4 studies (*)	Seem to follow all counting principles used by TD. Issues with number conservation and arithmetic skills.	n/a
Ansari (2003)	Unpublished narrative review (PhD thesis)	3 studies (*)	Seem to possess some number-relevant processing capacities in infancy. Present significant problems in grasping fundamental numerical concepts, such as conservation of number and seriation. Calculation abilities do not exceed those expected for 8 years old TD.	n/a
O'Hearn and Luna (2009)	Narrative review	10 studies (*)	Magnitude representation is atypical in WS throughout development. Representation of small numbers appears to be relatively typical in infancy, with a truncated developmental trajectory.	n/a

			Strong verbal and memory skills may support mathematical development.	
Dowker (2020)	Narrative review	4 studies (*)	Mathematical abilities appear to be severely impaired. Mathematical abilities related to spatial representations are more impaired than the ones related to verbal recall.	n/a
Van Herwegen and Simms (2020)	Systematic review	27 studies published between 1961 and 2019	Overall mathematical abilities, except for simple counting and subitizing, were reported to be impaired but in line with overall mental-age abilities. Could not conclude that number line skills are impaired. Cardinality is delayed and atypical. Calculation skills are unknown.	n/r

Note. n/a: not applicable; n/r: not reported. The count of the studies was not explicitly reported in the review, but it was manually computed by the author of this thesis based on the in-text citations.

In summary, the use of different terms to define mathematical abilities in the different reviews, the different terminology used to interpret the findings, the insufficient information provided in the reviews, and the lack of indication of the age of the participants in the different studies makes it difficult to make judgements on the overall findings and to provide a precise account of mathematical abilities in this population, in terms of basic processes, specific components, as well as overall mathematics achievement.

1.3.3 Autism

Autism is a common, highly heritable, and heterogenous NDC (Lord et al., 2020). Twin and genetic studies consistently demonstrate that autism has a particularly large genetic contribution with estimated heritability from 40% to 90% (Lord et al., 2020). Despite this, autism is not currently defined by any specific biological features that provide a distinctive marker or a specific cause (Muhle et al., 2018). In fact, currently only 15% of cases of autism appear to be associated with a known genetic mutation. However, even when a known genetic mutation is associated with autism, it appears that not all individuals with that same genetic mutation will develop autism (American Psychiatric Association, 2022). A list of environmental risk factors for autism has been identified by the literature. However, these factors cannot be considered causal but reactive, independent, or contributory to autism (Lord et al., 2020). Hence, autism is currently diagnosed on the basis of a set of behaviours observed from early years (American Psychiatric Association, 2022).

Autism was described for the first time in the 1940s by two scientists on the basis of a series of case studies; the Austrian psychiatrist Hans Asperger and the Austrian-American psychiatrist Leo Kanner (Asperger, 1991; Kanner, 1943). For a long time, autism research moved slowly and the attention to what was considered a rare condition was limited until a few decades later, when the diagnosis started to enter the general use. In 1979 Wing and colleagues published a seminal paper where the concept of "triad of impairments" was introduced for the first time (Wing & Gould, 1979). The paper reported on a large epidemiological survey and grouped the features of autism in 3 categories: social interaction, communication, and imagination. This study marked the recognition of heterogeneity within autism and the triad was then represented in the DSM-3 (1980) as a triad of social interaction, communication, and restrictive and repetitive behaviours and interests. Since then, the clinical model of autism has gone through several changes that were reflected both in the diagnostic criteria and in the focus of academic research. The current DSM-5-TR (2022) defines autism as characterised by persistent impairment in reciprocal social communication in multiple contexts, and restricted and repetitive patterns of behaviour, interest, or activities. These symptoms are present from early childhood and limit or impair everyday functioning. In addition to these core symptoms, co-occurring genetic, psychiatric, neurological and medical conditions are more common in autistic population than in the general population (Al-Beltagi, 2021). These include, epilepsy (10% to 30% autistic children have epilepsy), anxiety disorders, depression, hyperactivity and attention disorders, ID, sleep disorders (present in about 80% of the autistic population), gastrointestinal disorders (occurring in 46% to 84% of autistic population), metabolic disorders and autoimmune disorders (Al-Beltagi, 2021; Lord et al., 2020). These co-occurring conditions further impact the quality of life of individuals with autism, as well as their development.

As described by Lord et al. (2020), the then-current formulation of the diagnostic criteria for autism contained several changes from previous editions to better reflect clinical consensus and practice. For example, the sub-type of "Asperger's disorder" (AS) was combined under the unitary term "autism spectrum disorder" as the AS diagnosis was inconsistently applied by clinicians, and the sub-type of Pervasive Developmental Disorder – Not Otherwise Specified was removed. Moreover, this version of the diagnostic manual recognised the developmental nature of autism and accepted symptom onset during the early developmental period, rather than only during the first three years of life. By doing this, the DSM-5 (2013) – as well as the latest DSM-5-TR (2022) – acknowledges that symptoms might not fully manifest until social demands exceed limited capacities of the individual, and accepts that, for example, onset might not be noticed until the child reaches school-age or later. Symptoms of autism have a gradual developmental onset. Indeed, although the average age of autism diagnosis remains at ~4–5 years of age, parents typically report first concerns to health professionals in early childhood, at ~2 years of age (Lord et al., 2020). However, symptoms may be seen earlier than 12 months if developmental delay is severe or may be noted later than 24 months if symptoms are more subtle (Lord et al., 2020). Several diagnostic and screening instruments for autism exist, including structured diagnostic interviews, observational assessments and standardised screening tools, but only a limited number have been rigorously tested for diagnostic accuracy of the expert clinician judgement (Lord et al., 2020). The best validated instruments are the Autism Diagnostic Interview (Le Couteur et al., multiple versions) and the Autism Diagnostic Observation Schedule (Lord et al. multiple versions). However, the high cost of the instrument and training, the time required to complete the assessment and the need of substantial training to use them reliably pose a challenge on their widespread adoption (Lord et al., 2020). Table 1.7 lists the diagnostic and screening tools used in the studies cited in this thesis.

Table 1.7. List of diagnostic and screening tools for autism.

Name	Rater	Description
Autism Diagnostic Interview (ADI; Le Couteur et al., multiple versions)	C	Developmental interview completed with parents.
Autism Diagnostic Observation Schedule (ADOS; Lord et al., multiple versions)	C	Semi-structured assessment of the child.
Autism Spectrum Quotient (ASQ; Wakabayashi et al., 20006)	S, P	Screening tool assessing autistic traits in the general population on a 3-points Likert scale (from definitely agree to definitely disagree)

Autism Spectrum Screening Questionnaire (ASSQ, Ehlers et al., 1999)	P, T	Screening tool assessing the presence of specific behaviours on a 3-points Likert scale (from no to yes)
Childhood Autism Rating Scale (CARS; Schopler et al. 1986)	C	Behavioural rating scale based on the direct observation of the child.
Gilliam Autism Rating Scale (GARS; Gilliam, 2006; Abdel Rahman & Hassan,2004)	C, P, T	Standardised screening tool assessing the frequency of behaviours on a 4-points Likert scale (from never observed to frequently observed).
Repetitive Behaviour Questionnaire (Turner, 1997)	P	Questionnaire measuring the presence, frequency, and duration of repetitive behaviours.
Social Communication Questionnaire (SCQ; Rutter et al., 2003)	P	Screening tool assessing the presence of specific behaviours on a yes/no format.
Social Responsiveness Scale (SRS; Constantino & Gruber, multiple versions)	P, T	Standardised assessment which identifies the presence and severity of social impairments within the autism spectrum on a 4-points Likert scale (from not true to almost always true).
Social Skills Rating System (SSRS, Gresham and Elliott, 1990)	P, T, S	Standardised assessment which assesses on a 3-points Likert scale the frequency of occurrence (from never to very often) and importance of a specific behaviour (from not important to critical).

Note: C: clinician or trained professional; P: parent / caregiver; T: teacher; S: self-rating.

Estimates of the prevalence of autism vary considerably between studies, with figures ranging from 0.28% to 2.64% (Lord et al., 2020). This large variation arises from the different methodologies employed for calculating estimated prevalence and from the different level of strictness of the diagnostic criteria (Lord et al., 2020). Noteworthy, after accounting for methodological variation it has been shown that there is no clear evidence of a change in the prevalence of autism based on geographic region or ethnicity, across child and adult samples (American Psychiatric Association, 2022) and across time, specifically between 1990 and 2010 (Lord et al., 2020). In the 2010 Global Burden of Disease study, an estimated 52 million people had autism globally, equating to a prevalence of 1 in 132 individuals, that was just under 1% (Lord et al., 2020). However, the prevalence of autism in mental health inpatient settings is estimated to be higher than in the general population, ranging from 4% to 9.9% (Lord et al., 2020).

Autistic individuals often display a wide range of cognitive abilities and behavioural profiles and can have a range of IQ scores from above average intelligence to scores within the range of ID (Masi et al., 2017). It has been a long-held belief that autistic individuals exhibit a distinctive cognitive profile at a group level, characterised by a higher Non-verbal IQ (NVIQ) scores compared to the

Verbal IQ (VIQ) scores (Fodstad et al., 2020). This heterogeneity was not limited to individuals with autism who have intellectual disabilities but it was also observed among those with average or high IQ scores (American Psychiatric Association, 2022) and across various cognitive domains (Chen et al., 2019). As a result, there has been a general tendency of academic research to identify cognitive patterns and different subtypes of autism (such as, Asperger; AS, and High Functioning Autism; HFA) to guide diagnostic criteria and better inform clinicians, while the interest in the academic performance of autistic population has grown more recently (Whitby & Mancil, 2009). Still the major focus remains the language and social-communication domains (Fletcher-Watson & Happé, 2019), while studies investigating mathematical development are fewer and their findings are fragmented and conflicting (e.g., Dowker, 2020).

On one hand, the public notion of mathematical ability in autism is of a relative strength (Aagten-Murphy et al., 2015), possibly reinforced by the representation of autism in TV and films (Nordahl-Hansen et al., 2018). In fact, mathematics represents one of the skills reported as a savant talent in autism (Howlin et al., 2009) and there are a number of case studies describing autistic individuals with excellent exact quantification skills (Sacks, 1986), mental calculation skills (Kelly et al., 1997), and memorization of mathematical patterns such as calendrical calculation (Thioux et al., 2006) or detection of prime numbers (Hermelin & O'Connor, 1990). Also, a few studies suggested that individuals with autism are better at some aspects of mathematics than TD individuals (Iuculano et al., 2014). On the other hand, there is a growing body of research suggesting that mathematics is an area of difficulty for quite a large number of autistic students (Eaves & Ho, 1997; Myles et al., 2001). Moreover, mathematical difficulties have been reported to be more common in autistic than non-autistic individuals (Dowker, 2020). Finally, the impairment of social and communication abilities in young autistic children may hamper learning (including mathematical learning) especially when learning happens in the classroom through social interactions with peers (American Psychiatric Association, 2022).

Table 1.8 summarises the findings from the reviews which have been conducted so far on mathematical abilities in autism. These include one systematic review, one narrative review, and one meta-analysis. The three reviews span across a wide age range (3 – 51 years).

In contrast to what was observed for DS and WS, the studies investigating mathematical abilities in autism focused on the specific mathematical components of calculation and arithmetic word problems. Only the narrative review by Dowker (2020) mentioned a few studies investigating early numerical abilities such as numerical magnitude processes, subitizing, and counting and reported conflicting findings from the literature.

As for arithmetic word problem skills, the systematic review conducted by Chiang and Lin (2007) reported that most of the participants with AS/ HFA[11] included in the studies performed on the Arithmetic subtest of the Wechsler assessment at an average level (mean: 92.5; SD: 7.1). This was interpreted by the authors as indicating that "the majority of students with AS/HFA have average mathematical abilities" (Chiang & Lin, 2007, p. 551). However, in most studies the score on the Arithmetic subtest was significantly lower than the score that would be predicted from their FSIQ. This was interpreted by the authors as indicating that "the majority of individuals with AS/HFA autistic have a significant but clinically modest mathematical weakness" (Chiang & Lin, 2007, p. 547). Further, a few participants obtained extremely high scores. This was interpreted by the authors as suggesting that "some individuals with AS/ HFA have mathematical giftedness" (Chiang & Lin, 2007, p. 547). Similarly, the meta-analysis by Tonizzi and Usai (2023) reviewed studies investigating arithmetic word problem skills and calculation skills on autistic individuals without ID using two subscales of the WIAT assessment, and showed that generally the autistic group fell within the mean of the normative sample. This result was interpreted by the authors as indicating that "most students with ASD have average mathematical ability" (Tonizzi & Usai, 2023, p. 7). However, the autistic group (n = 533) reported lower scores than the TD group (n = 525) with a small-to-medium effect (g = 0.49), which highlighted that "people with ASD have poorer maths skills that their TD peers" (Tonizzi & Usai, 2023, p. 1). Moreover, the analysis of the mean difference between the autistic and the TD group that used the two subscales of the WIAT assessment as moderator showed that "students with ASD show similar performance on different

[11] Chiang and Lin (2007) define individuals with a diagnosis of AS as individuals with a diagnosis of autism who have normal language development, and individuals with a diagnosis of HFA as individuals with a diagnosis of autism who have average and above IQ.

maths tasks", indicating that their performance on calculation skills and on arithmetic word problems were similar (Tonizzi & Usai, 2023, p. 8). Finally, the narrative review by Dowker (2020) reported that even if a few studies suggested that autistic individuals are better at some aspects of mathematics than TD individuals, this was not found in most studies, and when this was found, it was limited to calculation abilities.

The meta-analysis by Tonizzi and Usai (2023) investigated the role of domain-general skills on mathematics achievement. The findings from the moderator analyses showed that the effect size was moderated by CA, VIQ, and working memory. Specifically, as age and the standardised mean difference on working memory measures between autistic and TD groups increased, effect sizes in mathematical measures also increased. Conversely, higher VIQ scores were associated with smaller effect sizes in mathematical performance.

Table 1.8. List of literature reviews on mathematical abilities in the autistic population.

Author	Type of review	Studies included	Main Findings	Age range
Chiang and Lin (2007)	Systematic review	18 (of which 8 studies were examined and discussed) published between 1986 and 2006	Most individuals with AS/HFA demonstrated average mathematical ability. Most individuals diagnosed with AS/HFA have a significant but clinically modest math weakness. Some individuals with AS/HFA have mathematical giftedness.	3 – 51 years
Dowker (2020)	Narrative review	15 (*)	Mathematical difficulties seem to be more common in autistic than non-autistic individuals. Conflicting findings on early mathematical abilities. Autistic individuals show better mathematical abilities than TD only for calculation.	n/a
Tonizzi and Usai (2023)	Meta-analysis	13 published between 2020 and 2022	The autistic group generally fell within the mean of the normative sample. Autistic participants reported a lower performance than the TD group on tasks assessing calculation and arithmetic word problem skills (significant small-to-medium effect). VIQ, WM and CA were found to be significant moderators of the differences on maths performance between TD and autistic groups.	6-16 years

Note. n/a: not applicable. (*) The count of the studies was not explicitly reported in the review, but it was manually computed by the author of this thesis based on the in-text citations.

In summary, the scarcity of the information provided on the studies investigating numerical magnitude processes and enumeration skills make it difficult to draw any conclusion on the development of these basic processes and specific components in autism. On the other hand, the reviews unanimously indicate that autistic individuals possess average arithmetic word problem skills and calculation skills but perform poorer than their TD peers. It remains uncertain whether this observation holds true for the entire autistic population, that includes individuals with and without ID. In fact, the limited information provided on the characteristic of the samples of the studies cited in the reviews by Chiang and Lin (2007) and by Dowker (2020), and the fact that the review by Tonizzi and Usai (2023) only focus on autistic individuals without ID limits the conclusions that could be drawn in relation to the existence of different mathematical profiles for autistic individuals with ID and autistic individuals without ID.

1.3.4 Overview of the educational system in England

The experiences and outcomes of students are influenced by various factors which are unique to each educational system, such as its structures, policies, and practices. This section offers a brief overview and provides the reader with some context of the educational system in England, where all the studies included in this thesis were conducted.

Education in England is overseen by the Department for Education. The education system provides both inclusive education through mainstream schools and specialised support and provisions for students with more complex or severe learning difficulties through specialist schools. The educational system is structured in several stages and provides full-time compulsory education to students between ages of 5 to 18. The educational system is typically divided into the following stages:

- Early Years (EY): Ages 3 – 5. This stage includes Nursery and Reception classes and focuses on learning through play and social development.
- Key Stage 1 (KS1): Ages 5 – 7. This stage includes Year 1 and Year 2 of primary school, during which core subjects such as English, mathematics, and science are introduced, and emphasis is made on foundational skills.

- Key Stage 2 (KS2): Ages 7 – 11. This stage includes Year 3 to Year 6 of primary school. It builds on earlier learning, further develops core subjects, and introduces additional topics like history, geography, and languages.
- Key Stage 3 (KS3): Ages 11 – 14. This stage includes Year 3 to Year 9 of secondary school.
- Key Stage 4 (KS4): Ages 14 – 16. This stage includes Year 10 and Year 11 of secondary school. It includes the General Certificate of Secondary Education examinations.
- Post-16 Education: Ages 16 – 18. After completing compulsory education at age 16, students have the option to continue their studies in post-16 education, which includes further education colleges, sixth form colleges, and other educational institutions. This stage may include qualifications such as A-levels, vocational qualifications, and apprenticeships.

Within the school environment, various professional roles contribute to a student's support. These include class teacher, Teaching Assistant (TA), Special Educational Needs Coordinator, and the Senior Leadership Team. While the Special Educational Needs Coordinator and the Senior Leadership Team focus on overseeing the overall school environment, setting school policies, and coordinating interventions and programs, the class teacher and TA are responsible for delivering the curriculum. While the class teacher is a fully qualified educator who holds teaching qualifications, there are no formal qualification requirements for the TA role. TAs have been systematically employed in schools since the 1980s and since then their potential areas of responsibility have steadily expanded (Fritzsche & Köpfer, 2022), especially when it comes to supporting students with special educational needs. Consequently, their involvement in direct instruction for both assigned individual students and entire classes has grown, with studies noting that TAs spend a significant portion of their time engaging in direct pedagogical interactions with pupils (Wren, 2017). Additionally, the study recent by Hargreaves et al. (2021), which used an online survey to explore parent views of the educational experiences of pupils with DS attending primary mainstream schools in the UK, showed that teachers and TAs were perceived by parents to share many of the teaching responsibilities in supporting students with DS. In fact, a high proportion of the parents indicated that they viewed TAs as primarily responsible for activities including delivering teaching, preparing teaching materials, managing behaviour, and motivating the student.

1.4 Aims and outline of doctoral thesis

The first aim of this thesis is to investigate the syndrome specificity of mathematical profiles in different NDCs. Diagnostic categories are widely used by clinicians to categorise difficulties, to establish who receives additional support and to inform intervention decisions (Astle et al., 2022). These categories also provide a frame of reference for research and affect study design, analyses, and theory. However, as reported by Astle et al. (2022, p. 397), the DSM system and its use of discrete diagnostic categories "appears to be straining at its limits". The question on the relevance of using diagnostic categories to differentiate cognitive profiles and the consequent definition of the support needed has been discussed in the past, for example, in the context of the language profiles (Bradshaw et al., 2020), to evaluate the speech and language provision offered (Dockrell et al., 2019), and in relation to numerical impairments (Paterson et al., 2006). The aim of this thesis is to further investigate syndrome specificity of mathematical profiles of different neurodivergent populations (DS, WS, and autism) through the use of cross-syndrome comparisons conducted by using experimental studies and systematic review methodology.

The second aim of this thesis is to describe the mathematical learning experiences of primary school children with NDCs, with a focus on DS and WS. As discussed in the first part of this Chapter, the neuroconstructivist approach views the environment as a crucial factor for development (Mareschal, 2007). However, when it comes to mathematical development in NDCs the role of the environment has been scarcely investigated. Despite the findings from literature in TD populations showing a positive correlation between HME and the child's mathematical achievement (Daucourt et al., 2021), there are no studies that have investigated structural and functional indicators of the HME or its relationship with mathematical development in DS, WS and autism. This thesis addresses this aim through the use of quantitative and qualitative studies involving the participation of parents and educators.

The four studies included in the current thesis investigate the syndrome specificity of mathematical profiles and the mathematical learning experiences of neurodivergent children by asking the following research questions (RQs):

1) What processes (perceptual subitizing, conceptual subitising or counting) do individuals with DS and WS use to perform an enumeration task and what is their performance level? Do DS and WS exhibit similar eye gaze behaviours during the enumeration task and do those impact their performance? (Chapter 2)
2) What are the structural and functional indicators of the HME of primary school children with DS and WS? Are they similar for the two groups? (Chapter 3)
3) What are the mathematical profiles of autistic individuals with ID (i.e., with IQ scores below 75) and of autistic individuals without ID (i.e., with IQ scores above 75)? Are they similar? (Chapter 4)
4) What are the challenges faced and the teaching strategies employed by educators in the maths inclusive classroom to support primary school students with DS? (Chapter 5)

Chapter 2 presents the investigation and cross-syndrome comparison of the performance of children and adults with DS and WS matched for MA (RCPM) to a TD control group in an enumeration task, and it covers for the first time the investigation of conceptual subitizing skills in these populations. The use of the eye-tracking methodology allowed the investigation of the participants' eye movements while performing the task and the exploration of underlying syndrome-specific mechanisms. Cross-sectional developmental trajectories were used to investigate changes across development in these populations.

Chapter 3 explores structural and functional indicators of the home learning environment of primary school children with DS and with WS, with a focus on the HME. Moreover, this study compares the HME of the two groups and investigates whether the type and frequency of home-based activities and the parental expectations changed on the basis of the diagnostic category or the general level of functioning of the child.

Chapter 4 presents a systematic review on mathematical abilities in autism. This study offers a comprehensive overview of the current state of literature on mathematical abilities in autism, covering all basic mathematical processes and specific components of mathematics investigated so far. The study aims to determine the mathematical profiles of autistic individuals with ID and without ID to enable their comparison and the comparison with different diagnostic categories associated with ID (i.e., autistic individuals with ID, individuals with DS, and individuals WS).

Chapter 5 focuses on more formal aspects of the overall learning experiences and explores the experiences of educators supporting students with DS in mainstream English primary schools and the mathematics education practices implemented in the classroom through a series of focus group interviews.

The final chapter, Chapter 6, provides a general discussion of the wider implications and impact for research and practice arising from the findings presented throughout this thesis with reference to its two aims, outstanding questions, and future directions.

Chapter 2: Enumeration skills in Down syndrome and in Williams syndrome: Insights from eye movements

This study was conducted with the aim to investigate enumeration skills of individuals with Down syndrome (DS) and individuals with Williams syndrome (WS). Perceptual subitizing and counting skills have been investigated in these populations before (Ansari et al., 2007; O'Hearn et al., 2011; Sella et al., 2013; Zimpel & Rieckmann, 2022). However, because these studies investigated enumeration skills in the two groups separately, it is challenging to compare their findings as different designs and methodologies were used. The current study addressed this gap by investigating enumeration skills and by conducting a comparison between the two groups within the same experiment. This is the first study examining conceptual subitizing in these populations.

Furthermore, this was the first study investigating enumeration skills in DS and in WS by means of eye-tracking (ET). This methodology has been used in the past to examine eye movements when performing the enumeration task, but only in typically developing (TD) populations (Schleifer & Landerl, 2011; Watson et al., 2007) and in individuals with mathematical difficulties (MD) (Schindler et al., 2020) and developmental dyscalculia (DD) (Moeller et al., 2009). The investigation of fixation count and fixation duration in DS and WS while performing an enumeration task provided insights into the impact of eye movements on the overall performance and into the presence of syndrome-specific mechanisms.

Finally, in line with the neuroconstructivist approach, this study takes a developmental perspective and explores changes across development of both enumeration skills and eye movements observed during the enumeration task in DS and in WS.

The findings from this study have been published in Subitizing in Williams syndrome and Down syndrome: Insights from eye movements. *Research in Developmental Disabilities*, *106*, 103746. https://doi.org/10.1016/j.ridd.2020.103746; which was available online since 20

August 2020. This chapter presents a more detailed and expanded version of this work.

2.1 Background and rationale of the study

Enumeration skills encompass an individual's ability to identify the number of items in a set and comprise perceptual subitizing, conceptual subitizing, and counting, as detailed in Section 1.2.3. The upcoming section presents findings from studies exploring enumeration skills in individuals with DS and WS. This is followed by a section that provides an overview of the eye-tracking (ET) methodology and of the findings from studies which employed ET to investigate enumeration skills. The last section describes the aims and hypotheses of the current study.

2.1.1 Enumeration skills in Down syndrome and in Williams syndrome

There are a few studies that use the habituation–dishabituation paradigm[5] to investigate the ability of infants with DS and infants with WS to discriminate small numerosities, i.e. 2 vs 3 dots (Paterson et al., 2006; Van Herwegen et al., 2008). The study by Patterson et al. (2006) found that infants with WS (n = 11; mean CA = 30 months) displayed a performance pattern similar to the TD control groups matched for chronological age (CA) and mental age (MA) measured using the Bayley scale of infant development. This similarity was characterised by a significant difference in looking time between the habituation and post-habituation stimuli. Such a difference was not reported for infants with DS (n = 18; mean CA = 30 months). In line with the findings by Patterson et al. (2006), the study by Van Herwegen et al. (2008) found that infants and toddlers with WS (n = 9; mean CA = 35 months) displayed a significantly longer looking time at the post-habituation stimulus than the habituation one. The evidence from these studies suggests that the ability to discriminate small numerosities is present in infants with WS, whereas it is not observed in infants with DS.

Additionally, enumeration skills have been investigated in children and adults with WS (Ansari et al., 2007; O'Hearn et al., 2011) and with DS (Sella et al., 2013; Zimpel & Rieckmann, 2022) using tasks where participants were asked to enumerate the items in a set or to compare sets of items.

The study by Ansari et al. (2007) investigated perceptual subitizing skills in 18 children with WS (mean CA = 9.70 years), 13 adults with WS (mean CA = 28.90 years) and in four groups of TD individuals (age groups: 4 – 5 years, 6 – 7 years, 9 – 10 years, and adult group with mean CA = 30.80 years). Participants were sat in front of a screen and were presented with a display showing 2, 3, 5, 7, 9, or 11 black dots on a white background for 250 milliseconds (ms), a duration too brief to allow verbal counting of the items but sufficient for subitizing (O'Hearn et al., 2011). Participants were asked to tell the researcher “how many dots” were shown on the screen (Ansari et al., 2007, p. 761). Participants were told by the experimenter that they did not have to count, but instead to quickly estimate or guess how many dots were presented each time, and they were given the choice of providing their answer by saying the number aloud or by pointing to a 1-20 number line mounted underneath the laptop screen in front of them. The accuracy rates reported for arrays included in the subitizing range (that is 2 and 3 dots) was 100% for both WS groups as well as for the TD groups.

The study by O'Hearn et al. (2011) included two experiments where enumeration skills in WS were investigated and compared with a TD group matched for MA (KBIT). In the first part of Experiment 1, 15 participants with WS (mean CA = 20.33 years) were presented with a display showing 1 to 8 dark squares on a light grey background for 5 seconds (s) and were asked to report how many squares they saw (24 trials). Participants were encouraged to count out loud. Then, in the second part of Experiment 1, participants were presented with the same stimuli for 250 ms and were asked to perform the same task (48 trials). Findings were based on accuracy only and showed that participants with WS were highly accurate to enumerate up to 8 objects when stimuli were presented for 5 s. However, when stimuli were presented for 250 ms, participants could accurately enumerate up to 3 items, suggesting that individuals with WS showed a limited subitizing range – 1 to 3 instead of the typical 1 to 4 or 1 to 5 (Gilmore et al., 2018). In Experiment 2, 12 participants with WS (mean CA = 18.50 years) were presented with the same stimuli used for Experiment 1. However, this time each stimulus was presented for an unlimited duration of time (48 trials). This allowed researchers to measure both accuracy and reaction times (RTs), which were measured via a microphone that detected the participant’s response. Participants were given the following instructions: “When I hit the button, some squares are going to appear on the screen. Your job is to tell the computer how many there were. The computer will be listening for you, and it only wants to hear numbers so be careful and don't say things like ‘‘um’’. Sometimes you will just know how many squares there are really fast and if that happens, I want you to tell the computer the answer as fast as you can! Sometimes you won't know the answer so quickly, and then I want you to count in your head and tell me the answer when you know it. I want you to do your best and try to tell me the exact number each time. Ready to try?" (O'Hearn et al., 2011, p. 299). Findings for the WS group showed a relatively flat RT function for the limited subitizing range (1 to 3) along with an accurate performance, and greater RT slopes and lower accuracy for numerosities higher than 3. These results were interpreted by the authors as individuals with WS engaging with two separate processes, subitizing for small numbers (up to 3) and counting for larger numbers. Except for the width of the subitizing range, this behaviour was reported by the authors as being in line with the one observed in the TD population.

The study by Sella et al. (2013) investigated enumeration skills of 21 participants with DS (reduced to 14 for the dots-to-dots task and to 12 for the digit-to-dots task after data cleaning) with mean CA of 14.16 years. Participants with DS were matched with two TD control groups, one matched on CA and one matched on MA (PPVT). Participants were presented for 200 ms with a sample numerosity on the screen (that could be either a white set of dots from 1 to 9 or a white Arabic digit from 1 to 9 on a grey background) and, after a 1.1 s, with a target numerosity (a set of black dots on a grey background that could either match the sample numerosity – 50% of trials – or differ for one dot from the sample numerosity – 50% of trials). Participants were asked to compare the sample numerosity and the target numerosity and to report whether they were the same by pressing a button on the keypad (dots-to-dots task: 108 trials; digit-to-dots task: 108 trials). The time allowed to provide a response was 8 s, otherwise the response was categorised as missed. Findings from the dots-to-dots task showed that the accuracy of participants with DS decreased as numerosity increased, whereas the accuracy of both TD control groups remained at ceiling. Findings from the digit-to-dots task showed that participants with DS reported increasing RTs as the target numerosity increased. These findings were interpreted by the authors as suggesting that participants with DS did not use perceptual subitizing but instead relied on counting to perform the enumeration task.

More recently, Zimpel and Rieckmann (2022) investigated enumeration skills of 175 individuals with DS with mean CA of 19.50 years and of 276 TD individuals. Participants were presented with 1 to 8 black dots arranged in a dice pattern on a white background for 250 ms. They were asked to identify the number presented on the screen and either say the number name, type the number into the computer, or point to the number shown on a table (24 trials). The results of the study reported high accuracy rates for 1-6 dots for both the DS participants and the TD group (higher accuracy rates than 85%). These results seem to suggest that individuals with DS can perform both perceptual subitizing and conceptual subitizing. However, due to the absence of a specific matching criterion in the design of the study, it is not possible to infer whether the performance of the DS group was in line with CA and / or MA.

In summary, past evidence indicates that while individuals with WS perform perceptual subitizing, results of the studies investigating subitizing skills in individuals with DS are conflicting. While the study by Sella et al. (2013) reported that participants with DS seemed to use counting processes by default even for low numerosities, the study by Zimpel and Rieckmann (2022) reported high level of accuracy for both perceptual subitizing and conceptual subitizing. However, it is important to consider that while one of the studies investigating enumeration skills in DS did not use a standard dot enumeration paradigm, the other did not report RTs. Hence, the interpretation of the results and comparison with other experiments should be done with caution. Additionally, it is worth considering that even when similar experimental designs are used (for example, in the case of the two studies investigating enumeration skills in WS), the instructions given to the participants differed greatly and this may have an impact on the strategies used to complete the enumeration tasks. For example, the use of the word “count” in the instructions might lead the participant to always choose to use the counting process when performing the task. Finally, none of these studies discussed took a developmental perspective, in that they did not explore whether in DS and WS enumeration skills change with CA.

2.1.2 Use of eye-tracking to study enumeration skills

2.1.2.1 Brief introduction to eye-tracking methodology

Eye-tracking (ET) provides insights into how individuals process visual information and it involves collecting a participant's overt visual attention (that is the act of physically directing the eyes to a stimulus) by recording eye gaze data through the measurements of well-defined ocular behaviours (Sharafi et al., 2020). Table 2.1 provides a brief description of the most common ocular behaviours collected by eye trackers, as reported by Sharafi et al. (2020).

Table 2.1. Ocular behaviours recorded by eye trackers.

Ocular behaviour	Description
Fixation	Stable position of the eye during a gaze. It usually lasts for 100 to 300 ms. However, it changes with task and participant's characteristics. During a fixation, the participant's visual attention is focused on a specific area of the stimulus, and it triggers cognitive processes. Information acquisition and processing mostly occurs during fixations.
Saccade	Common, continuous, and rapid movement of the eye between two fixations. It lasts for 40–50 ms. Vision is reduced during saccades.
Pupil dilation	The pupil is the aperture through which light enters the eye, whose dilation is controlled by the iris muscle.
Scan path	Series of areas of the stimulus visited by the participant and sorted in chronological order.

Eye movements have long been used "to draw inferences about perception, cognition, and brain function in many areas of psychology, cognitive science and applied research fields" (Hessels et al., 2018, p. 2). For example, the literature review by Rayner (1998) presented a summary of the main findings of ET studies in reading, music reading, typing, visual search, scene perception, auditory language processing, numeral reading, problem solving, face perception and studies investigating eye movements of patients with brain damage. However, ET provides only a proxy to cognitive processes (Sharafi et al., 2020), and this limitation needs to be taken into consideration when interpreting findings. In fact, one of the most crucial challenges in ET research is to properly link eye movements to the assumed underlying cognitive processes (Strohmaier et al., 2020). As reported by Sharafi et al. (2020), the relation between eye gaze and cognitive processing is based on two assumptions from the theory of reading: 1) the immediacy assumption and 2) the eye-mind assumption (Just & Carpenter, 1980). The immediacy assumption proposes that interpretation of the stimulus begins immediately as a participant sees it. The eye-mind assumption states that participants fixate their attention only on the part of the stimulus that is being processed. Hence, it is assumed that the position of the fixation reflects the position to which attention is drawn and from which relevant information is extracted. This suggests that eye movements provide a trace of where attention is being directed.

According to Sharafi et al. (2020), ET data can be classified in three categories. The first order of data includes raw data, that is unfiltered ET outputs. The second order of data includes fixations and saccades derived from the first order, through categorization of raw data into events, that can be done either manually by the researcher or automatically through an automated algorithm. The third order of data includes data obtained through the analysis of fixations and saccades. The current chapter runs analyses on the third order of data, in particular analyses of fixations, which are the most common measures used in ET studies (Sharafi et al., 2020). The metrics used in the current study are fixation count, also known as fixation frequency (that is the number of fixations within an area of interest; AoI) and fixation duration. The interpretation of these measures is context-dependent, and various studies have offered differing interpretations of the findings. For example, Kennedy (2016) suggests that a low fixation count may indicate that the task goal has been reached, that the participant is experienced, or that the search task is too simple. Some studies report saccadic frequency instead of fixation count. However, the relationship between these metrics is straightforward, at least for the designs including still images, in that the number of saccades should be equal or +/- 1 to the number of fixations (Kennedy, 2016). Metrics such as fixation duration and average fixation duration have been used as measures of the time needed to process the information at the fixated position and in some studies have been interpreted as the amount of visual effort (or difficulty) to perform a task (Mock et al., 2016). According to Just and Carpenter (1976) long fixation durations may be interpreted as the individual having a hard time extracting and processing visual information as well as the individual being “more engaged” by the stimulus.

The application of ET methodology to the investigation of cognitive processes can offer insights into the characteristics of eye movements during specific tasks – such as the enumeration task used in the current study – and their impact on performance. In the context of syndrome comparison, the use of eye movement data can also support the understanding of the subtle differences observed among different conditions, as suggested by Van Herwegen and Karmiloff-Smith (2015). In turn, this understanding can be used by researchers and educators to design effective interventions for different populations.

2.1.2.2 Findings from studies investigating enumeration skills using eye-tracking

There is a limited number of studies that have investigated enumeration skills using eye movement data, two studies conducted on TD populations (Schleifer & Landerl, 2011; Watson et al., 2007) and two studies assessing individuals with developmental dyscalculia (DD) or mathematical difficulties (MD) (Moeller et al., 2009; Schindler et al., 2020). These studies employed ET measures to examine eye movements such as fixation count (or saccadic frequency) and fixation duration during the enumeration task, and to validate the hypothesis that counting and subitizing should be considered as two distinct processes.

In the study conducted by Watson et al. (2007), eight TD participants (mean CA = 21.90 years) were presented for an unlimited amount of time with 1 to 9 red circles on a grey background and were asked to identify the number of circles on the screen (45 trials). Their findings showed a tight coupling between RTs and fixation count. Both measures exhibited a sharp increase when participants were presented with more than 4 items, indicating participants' engagement in counting, whilst they were subitizing for the lower numbers.

Schleifer and Landerl (2011) presented 60 TD participants with 1 to 10 black blocks randomly arranged on a white background (40 trials). Participants (age groups: 8 – 11 years, 14-years, and adults with mean CA = 26 years) were asked to determine the number of the blocks shown on the screen as quickly as possible. The typical performance described in Figure 1.6, characterised by a sudden increase of the slope as participants start to count, was found in RTs as well as in saccadic frequency for all the groups. These results indicated a tight coupling between RTs and eye movements, consistent with the results reported by Watson et al. (2007). In particular, the authors reported that TD children and adults used few or no saccades for enumerating 1–4 objects and reported a monotonic increase in the number of saccades when enumerating 5 or more objects. This finding was interpreted by the authors as indicating that participants were using two different processes to perform the task, that is subitizing for up to 4 objects and counting for higher numerosities. Fixation duration was not found to vary systematically with the number of the items displayed. This result was interpreted by the authors as fixation duration not being a good indicator of the enumeration processes employed, but it might also suggest that the same cognitive effort (or the same level of visual search difficulty) is needed to perform the enumeration task when using subitizing and counting processes. Moreover, the analysis of the number of saccades during the counting task showed a developmental change in the counting strategy used by the participants. In fact, 8-year-olds reported on average a saccadic count higher than the number of dots to be counted, suggesting that they scanned the displays rather unsystematically. The 11-year-olds reported a number of saccades that corresponded closely with the number of dots to be counted. Finally, the older age groups showed saccadic frequency that tended to be lower than the number of dots to be counted, probably reflecting some strategy of parallel processing of dot clusters within the array (Schleifer & Landerl, 2011).

The study by Moeller et al. (2009) investigated enumeration abilities of two children with DD (CA_1 = 10.58 years and CA_2 = 10.83 years) and compared them to a TD control group of 8 children matched for CA. Participants were presented with 1 to 8 black dots on a white background for an unlimited amount of time and were asked to press a button and say the number of dots shown on the screen (64 trials). Their findings showed that, when enumerating 1–3 dots, RTs were larger for participants with DD than those of the TD control group. However, while fixation count of both children with DD were higher than the control group, DD participants did not differ from the control group in terms of average fixation duration. Therefore, the authors interpreted these findings as children with DD showing a deficit when subitizing that resulted from using counting at least on some proportion of the 1-3 dots trials rather than from a general difficulty with extraction and processing the numerical information. When enumerating more than 3 dots the DD participants showed RTs in line with the control group. However, the looking behaviour differed between the two individuals with DD and reflected the use of different strategies in that one participant with DD exhibited not more but longer fixations than controls as the number of dots increased, while the other participant with DD showed an increase of fixation count, rather than longer fixations.

More recently, Schindler et al. (2020) used ET glasses to investigate enumeration skills of 10 individuals with MD (mean CA = 10.10 years) and to compare them to a TD control group matched for CA. Participants were presented with 1 to 9 red dots on a white background arranged either in a random or in a dice pattern and were instructed to say the number of dots that was presented as fast as possible (36 trials). The study reported significant group differences only for RTs for the conceptual subitizing condition (i.e., 5 – 9 dots arranged in a dice pattern), with the MD group being slower than the TD group. No group differences were reported for fixation count in any experimental condition. Fixation durations were not collected or included in the statistical analyses. The results reported by Schindler et al. (2020) are in contrast with the findings reported by Moeller et al. (2009). These differences are not surprising and can be explained by the different criteria used to define the DD and ML samples, by the different ET apparatus used, and by the different design of the two experiments. Moreover, the limited number of participants included in both studies prompts caution when interpreting these results due to the restricted statistical power.

While there are no studies that have investigated enumeration skills in individuals with DS or WS by means of ET, there are a few ET studies that investigated the overall eye gaze relating to visual stimuli in individuals with DS and WS. These studies highlighted that these populations show different patterns of eye movements. On one hand, Van Herwegen (2015) reported that most of the studies investigating eye movements in WS found evidence for impaired abilities to plan and execute eye movements, also described as “sticky fixations”, that is characterised by “infrequent fixations”. On the other hand, the research by Viñuela-Navarro et al. (2017) reported a fixation “stability deficit” in 21 children with DS (mean CA = 8.36 years) compared to a TD control group, which was characterised by significantly shorter fixation durations but no significant differences on the fixation count. It has been argued that such differences at the level of domain-general processes may affect domain-specific outcomes (Van Herwegen & Karmiloff-Smith, 2015). In the specific case of enumeration task, the reported sticky fixation and the stability deficit may impact the visual processing of numerical information and consequently not only the accuracy of the performance and the RTs needed to complete the task, but also the processes used by the participants to perform the task.

2.1.3 Current study

The current study employed ET to investigate enumeration skills in children and adults with WS and DS, compared to a control TD group matched for performance on RCPM.

The first aim of this study was to investigate the performance on an enumeration task of individuals with DS and WS in terms of accuracy and RTs. The study aimed to answer the following RQs:

- What is the accuracy of individuals with DS and WS on the enumeration task? How do they compare to the MA-matched TD group?
- What are the RTs of individuals with DS and WS on the enumeration task? How do they compare to the MA-matched TD group?
- Do individuals with DS and WS use the same processes (i.e. perceptual subitizing, conceptual subitizing, counting processes) to perform the enumeration task? If so, do they use these processes in the same experimental conditions?
- Is the overall performance of DS and WS groups related to non-verbal cognitive abilities or calculation skills?

Based on the existing literature in WS (Ansari et al., 2007; O'Hearn et al., 2011) and in DS (Sella et al., 2013), it was predicted that participants with DS and WS would report high levels of accuracy on the enumeration task.

Regarding the processes employed for executing the enumeration task, it was predicted that, based on existing literature mentioned above, the two groups would exhibit distinct outcomes. Participants with WS would use different enumeration processes to perform the task depending on the number of dots presented – i.e., perceptual subitizing for numerosity 1–3 and counting for numerosity 4–6. On the other hand, participants with DS would not use perceptual subitizing but would rather use counting to perform the enumeration task, regardless of the number of dots presented. Due to the lack of previous literature on conceptual subitizing, no predictions were made on the outcomes regarding this ability for the WS and DS groups. These predictions do not take into consideration the results from the study by Zimpel and Rieckmann (2022), as that study was published two years after the current study. For completeness, findings from Zimpel and Rieckmann (2022) are considered in the final discussion.

In line with previous studies investigating enumeration skills in DS (Sella et al., 2013), it was predicted that accuracy of the DS group would be positively correlated to measures of non-verbal cognitive ability and calculations skills. No predictions were made for the WS group, given the lack of previous studies investigating these relationships.

The second aim of this study was to analyse eye movements (i.e., fixation count and fixation duration) during the enumeration task and to answer the following RQs:

- What eye movements do individuals with DS and WS exhibit during the enumeration task? How do they compare to the MA-matched TD group?
- Do participants' eye movements impact the processes used by participants to perform the enumeration task and / or their performance?

Based on the existing literature in DS (Viñuela-Navarro et al., 2017) and in WS (Van Herwegen, 2015), it was predicted that group differences would be observed for both clinical groups, with participants with DS presenting shorter fixation durations than the control group and participants with WS presenting fewer fixations than the control group. Moreover, it was assumed that differences in fixation count and fixation duration would impact the performance on the enumeration task, in terms of accuracy and processes used. However, given the lack of literature in this field no predictions were made on the nature of this impact.

Finally, in line with the neuroconstructivist approach, the third was to provide a developmental account of both enumeration abilities and eye movements in children and adults with DS and WS and to answer the following RQs:

- Does the enumeration performance change across development?
- Do the eye movements change across development?

Due to the scarcity of studies investigating enumeration skills in DS and in WS and of studies employing ET in these populations, no predictions were made.

In summary, the current study investigated and compared enumeration skills in children and adults with DS and WS, with a TD control group matched for MA (RCPM). Specifically, it examined perceptual subitizing, conceptual subitizing, and counting skills. ET methodology was employed to record participants' eye movements. This provided insights into the presence of syndrome-specific differences in basic-level processes and into the impact of eye movements on the task performance. This will ultimately help clarify whether educational interventions supporting mathematical development in these neurodivergent populations should target these basic processes of mathematics and whether such interventions should be adapted for different populations.

2.2 Participants and Methods

2.2.1 Participants

Seventy-eight participants took part in the study. Twenty-four participants diagnosed with DS aged 8;08 to 49;02 years (12 female) were recruited via Down syndrome support groups across England and Wales, in the UK. Twenty-six participants diagnosed with WS by fluorescent in situ hybridisation test and aged 8;00 to 51;08 years (16 female) were recruited through the Williams Syndrome Foundation across the UK. Twenty-eight TD children (17 female) were recruited from local schools and social media. One TD participant was excluded because they did not complete the entire assessment. Of the 27 TD children that completed the whole assessment, only those who had similar MA as measured by the raw scores of the RCPM, were included in the control group (*n* = 25, 16 female) (see Van Herwegen et al., 2019 for a similar approach). Therefore, the TD participants in the control group (aged between 3;11 and 6;07 years) were much younger than the participants diagnosed with DS and WS but had similar non-verbal cognitive ability scores as measured by the RCPM. All participants had English as a first language and none of the TD participants had a diagnosis for a learning difficulty. All participants came from white middle class family background.

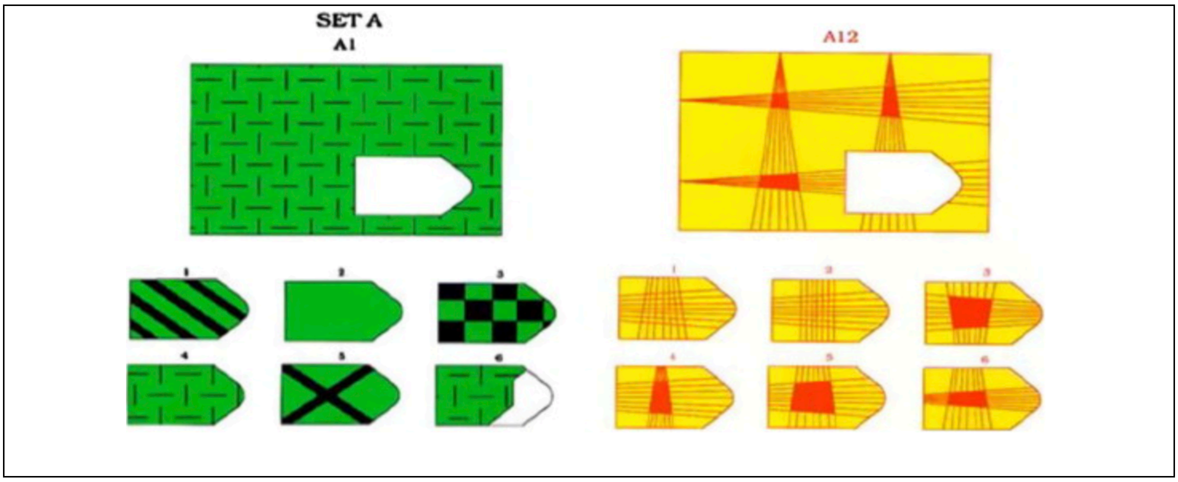
2.2.2 Measures

2.2.2.1 Raven's coloured progressive matrices

RCPM (Raven, 1993) is a standardised aptitude assessment with excellent psychometric properties that measures non-verbal cognitive ability (Ashworth et al., 2021). It is commonly used in neurodevelopmental research, including studies with individuals with DS (Facon & Nuchadee, 2010; Purser, 2015) and studies with individuals with WS (Ashworth et al., 2021). Moreover, Van Herwegen et al. (2011) compared the errors made by 53 individuals with WS (mean CA = 18;03 years) when completing the RCPM assessment with the errors made by TD children (mean CA = 5;08 years) matched by their raw RCPM score. Their findings reported that participants with WS made the same proportion of error types as the TD group and that the proportion of error types changed similarly over development. These findings further support the use of this assessment with the WS population.

The test includes 36 items which increase in difficulty and complexity within and across three sets. Each item contains on the top of the page an abstract figure with a missing part and six different options on the bottom of the page (Figure 2.1). The participant is asked to choose the option that completes the abstract figure. Each item was presented for an unlimited amount of time until the participant reported their response. If the participant did not know the answer or struggled with a particular item, they were allowed to guess. All participants started from item 1 and were presented with the 36 items. A score of 1 was given for every trial performed correctly. The minimum raw score is 0 and the maximum raw score is 36.

Figure 2.1. Item A1 and item A12 of the Raven’s Coloured Progressive Matrices test.



Note: From “Investigation of basic cognitive predictors of reading and spelling abilities in Tunisian third-grade primary school children” by Batnini S. & Uno A., 2015, Brain and Development, 37(6), p. 582 (https://doi.org/10.1016/j.braindev.2014.09.010). Licence number: 5751861408947.

2.2.2.2 Numerical Operations subscale of WIAT 2

The Numerical Operations subscale from the Wechsler individual achievement test second edition (WIAT-OP; Wechsler, 2005) was administered to assess basic numerical knowledge and calculation skills. This standardised assessment includes 54 items and measures the ability to identify and write numbers, to count using the 1:1 correspondence, to solve simple calculations involving the four basic operations, to use fractions, decimal numbers, and percentages, and to solve algebraic and geometric tasks. All participants started from item 1 and the assessment was terminated after 6 consecutive failed or skipped items. A score of 1 point was given for each correct response. The minimum raw score is 0 and the maximum raw score is 54.

2.2.2.3 Verbal counting

Participants were asked to verbally count from 1 to 20. In particular, the researcher asked the participant: “Can you count for me to 20? Start with 1”. If the participant was reluctant to count, the researcher would encourage them to count by saying “I will start. 1 and then comes…?” If the participant stopped counting, the researcher would ask the participant to continue by using probes such as: “What comes next?” or “And then…?”. The task was stopped either when the participant reached 20, when they made a mistake, or when they stopped because they did not know how to continue counting. A score of 1 was given for each item performed correctly. The minimum raw score was 0 and the maximum raw score was 20. If the participant made a mistake or stopped counting, a score equal to the last correct number reported by the participant was given and the type of error made was recorded by the experimenter. The same task has been used in previous studies with individuals diagnosed with NDCs, for example in Paterson et al. (2006).

2.2.2.4 Enumeration task

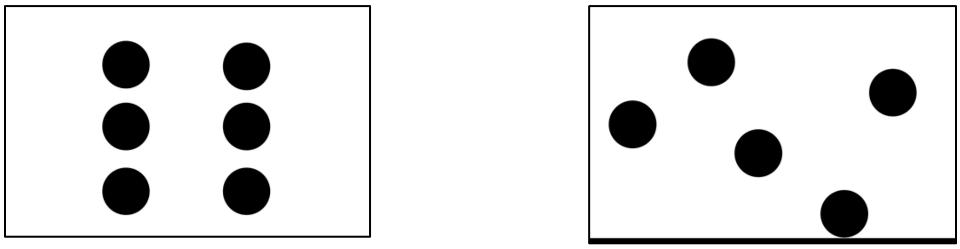
Enumeration skills were assessed with a researcher-developed, computer-based task and by using an eye tracker. The task started with a 5-point calibration task[12] initiated and controlled by the researcher, which was followed by the experimental task. The enumeration task usually took less than 5 mins, depending on the quality of the calibration task. In the instances in which the data for one or more calibration points was missing or was showing low accuracy or precision, the calibration was repeated for such specific points. To support the completion of the calibration, the researcher provided verbal instructions to the participant and used their finger to guide their gaze. One of the most common causes of low quality of the calibration was the thick lenses of the eyeglasses worn by participants, as they can produce noise due to reflections caused by the gaze tracker's lights. To overcome these difficulties, lenses were accurately cleaned by the researcher before starting the testing session. The quality of the calibration task was evaluated by the researcher in real time during the task. In cases of doubt about the overall accuracy of the calibration, the researcher took note of the difficulties encountered and the inclusion of the participant was discussed in laboratory meetings with more experienced members.

After completion of the calibration task, the enumeration task started. The participant was presented with static stimuli that contained between 1 and 6 black dots of the same size on a white background (Figure 2.2). The dots were arranged

[12] Calibration is the process through which the eye tracker firmware adapts the algorithms to the person sitting in front of the eye tracker. During the calibration the participant is asked to look at specific points on the screen located at known coordinates. At each location, the distance between the pupil and the corneal reflection is measured by the eye tracker. For this experiment, a 5-points calibration process was selected because 5-point calibration is reported to yield good results and is usually not experienced as intrusive by the users (Tobii AB, 2023).

either in a canonical pattern (dice condition, with the dots centred in the middle of the screen) or in a non-canonical pattern (random condition, with the dots positioned in random locations of the screen). There was one trial for each numerosity and for each condition, yielding a total of 12 trials. The 12 trials were arranged by the researcher in two predefined orders alternating between dice condition and random condition, which were used in turn. Participants were told they would see some black dots on the screen and that they had to say how many dots they saw as quickly and accurately as possible. A score of 1 was given for every trial performed correctly. The minimum score is 0 and the maximum score is 12. Accuracy, RTs, and eye movements were recorded during the enumeration task.

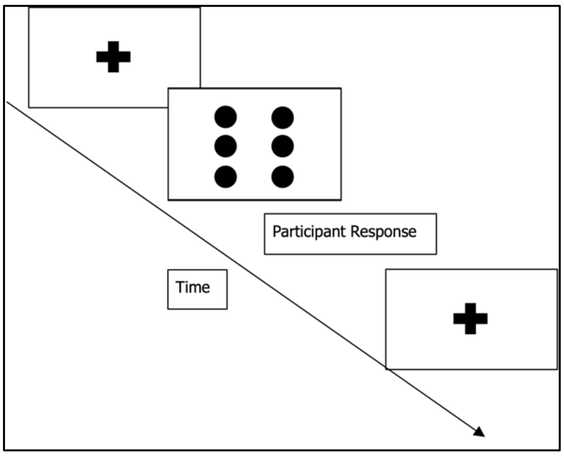
Figure 2.2. Stimulus D6 and stimulus R5.



Note: Stimulus showing 6 dots in dice pattern (D6) and stimulus showing 5 dots in random pattern (R5).

Figure 2.3 shows the enumeration task. Before each trial, a black fixation cross was displayed in the centre of the screen to capture the participant’s attention and to ensure a clear transition from one stimulus to the next. The experimenter initiated each trial only when the participant appeared to be attentive and looked at the fixation cross. Each stimulus was presented until the participant reported how many dots there were. The participant could report their response verbally or by using signing and did not receive any feedback for their response. The response was recorded by the experimenter and a fixation cross appeared on the screen before the following trial.

Figure 2.3. Enumeration task.



The experimental conditions were defined by the number and the arrangement of the dots in the stimulus. As shown in Table 2.2, four experimental conditions were identified: dice pattern in subitizing range (1 – 3 dots displayed in dice pattern; D1-D3), dice pattern in counting range (4 – 6 dots displayed in dice pattern; D4-D6), random pattern in subitizing range (1 – 3 dots displayed in random pattern; R1-R3), and random pattern in counting range (4 – 6 dots displayed in random pattern; R4-R6).

Table 2.2. Experimental conditions.

Range	Pattern	
	Dice	Random
Subitizing (1– 3)	D1-D3	R1-R3
Counting (4 – 6)	D4-D6	R4-R6

The subitizing range was defined for numerosities 1 to 3 because this study involved young children who might have a restricted subitizing range (Gilmore et al., 2018) and participants with WS who have been reported to have a reduced subitizing range (O'Hearn et al., 2011). Moreover, the decision to use a 2 (dice pattern vs random pattern) x 2 (subitizing range vs counting range) design rather than a 2 (pattern) x 6 (number of dots) design was taken to increase the power of the statistical analyses.

2.2.3 Procedure

The testing session started with the RCPM assessment, followed by the verbal counting task, the WIAT-NO test, and lastly the enumeration task. The entire session took approximately 50 minutes and breaks were taken between different tasks, when needed. Adult participants and the carers of young participants provided written informed consent prior to their participation. All participants provided verbal assent before starting the testing session. Before starting the testing session, the researcher showed them how the Tobii T120 eye tracker worked and why it was important that during the testing session the participant stayed as still as they could. The participant and their carer were then given the opportunity to ask questions before proceeding with the assessment.

The testing always took place face-to-face. Most of the testing sessions took place in the laboratory at the Kingston University campus while some participants were tested either at Grange Centre for People with Disabilities in Surrey (*n* = 3 participants with DS and *n* = 3 participants with WS) or at the University of Cardiff campus (*n* = 6 participants with DS). When testing took place at the Kingston University campus, participants completed the tasks in two rooms. Figure 2.4 shows the room where the enumeration task was assessed. This room was quiet, windowless, and with good and stable lighting to avoid interference with the infrared light of the eye tracker.

Figure 2.4. Testing room and eye-tracking equipment.

Note: The picture on the left shows the room of the university laboratory at Kingston University where the enumeration task was conducted. The desk on the left was occupied by the experimenter, while the desk on the right was occupied by the participant. The picture on the right shows the Tobii T120 screen-based eye tracker used for the experiment.

The participant was seated 60 cm away from the 17 inches monitor screen, where the static eye tracker was mounted. The participant was seated in an adjustable chair to accommodate for different participants' heights and was not subjected to movement restrictions through, for example, the use of head rest or head frames. The researcher sat next to the participants to reduce distractions and the carer sat behind the researcher, when present in the room. When needed, the curtains separating the participant and the researcher were used to reduce sources of distractibility for the participant. The participant was asked to stay as still as they could during the enumeration task. In case they found it difficult, the researcher asked their carer to stand behind them and to contain their movements by holding their head still and by trying to limit the participant leaning backward, forwards, or sideways for the duration of the enumeration task, without looking at the screen. The same testing conditions and testing environment were replicated as close as possible in the other two testing locations.

The study received ethical approval from the Kingston University Ethics Committee (ref. 161706); and from IOE, UCL's Faculty of Education and Society (data protection registration number: Z6364106/2022/04/75). No incentives were offered to participants for their participation in the study.

2.2.4 Data analysis process

2.2.4.1 Eye-tracking data

In the current study eye movements were recorded using a Tobii T120 screen-based eye tracker, which used infrared technology to measure corneal reflection in the participant. The Tobii T120 eye tracker was connected to a Windows pc used by the experimenter to perform calibration, to control stimuli presentation, and to monitor the quality of ET data collected during the experiment. The recording of the eye movements was controlled with Tobii's Studio software (version 2.06) at 120 Hz. Data was locally stored on the Tobii T120 eye tracker and back-up copies were created at the end of every testing session.

To extract and calculate ET data Tobii's Studio software requires the researcher to set up an area of interest (AoI), which is a predefined area of the stimulus. To ensure that the AoI would be large enough to capture all the relevant fixations, one AoI was set for each stimulus to cover the entire screen (Figure 2.5). This approach was chosen over the option of creating separate AoIs for each dot included in the stimulus because it avoided potential issues with overlapping AoIs and ensured a fair comparison between the stimuli, as they all had the same size.

Figure 2.5. Areas of interest set for the stimuli D6 and R5.

Note: The coloured area shows the AoI for the stimulus showing 6 dots in the dice pattern (left panel) and for the stimulus showing 5 dots in the random pattern (right panel).

Fixations were defined using the Tobii Velocity-Threshold Identification fixation classification algorithm averaged from both eyes (Anneli, 2012). The current study recorded (a) fixation count, which was measured as the number of times the participant fixated the AoI from the appearance of the stimulus on the screen until their response; (b) median fixation duration, which was defined as the median of the duration of individual fixations within the AoI identified from the appearance of the stimulus on the screen until participant’s response. Median fixation duration scores were used instead of mean fixation duration scores, as they are less strongly influenced by outliers (Schleifer & Landerl, 2011).

Participants for whom the ET data were not recorded for more than 50% of the total duration of the trial were excluded (n = 1 WS participant), as they were considered not reliable. This reduced the sample size of the WS group to 25 participants. Moreover, single trials were discarded from the analyses if a response was made in less than 0.2 s or more than 10 s from the trial onset (DS: n = 2 trials; WS: n = 2 trials; TD: n = 2 trials), as it was deemed that such RTs were too short or beyond a fixed threshold and were denoting poor attention (see Paul et al., 2017 for a similar approach).

2.2.4.2 Analysis plan

Accuracy rates, RTs, and eye movements were recorded during the enumeration task for each participant. RTs were measured as the duration of the visit within the AoI, that is from the appearance of the stimulus until the participant’s response.

Statistical analyses were conducted only on those trials for which participants produced a correct response. Hence accuracy analyses were conducted to define the dataset. Error analysis was not planned to be conducted, since a significant number of errors was not expected based on previous studies in TD and WS populations (Ansari et al., 2007; O'Hearn et al., 2011).

When the normality assumption was violated, non-parametric analyses were conducted instead of parametric analyses. When the assumption of equality of variance across groups was violated, Welch Analyses of variances (ANOVAs) were run.

To answer RQ1, ANOVAs or the related non-parametric tests were conducted to identify differences between groups for accuracy rates and RTs. Additionally, a series of Wilcoxon signed-rank tests was run to investigate whether participants' accuracy rates were influenced by the experimental conditions.

Then, an analysis focused on the data slopes for RT based on the study by Schleifer and Landerl (2011) was conducted. To perform the analysis on RTs, first regression lines for each experimental condition were individually computed. The average slope for each group was computed and submitted to four separate one-way Welch ANOVAs, one for each experimental condition, to determine if the three groups were using different enumeration processes in the same experimental conditions. Moreover, for each group a series of Wilcoxon signed-rank tests was conducted on the average RT slopes to investigate whether the processes used by participants were influenced by the experimental conditions in the same way.

Finally, for each experimental condition correlations were run to determine the relationship between mean accuracy scores and MA measured through RCPM, and between mean accuracy scores and calculation abilities.

To answer RQ2, fixation count and median fixation durations were examined. To examine differences in the eye gaze behaviour between groups, three-way mixed ANOVAs were conducted on fixation count and on median fixation duration for each experimental stimulus. In case of significant differences between groups were found, post hoc analyses were run.

For each group, the mean fixation count was compared with the number of dots to be enumerated to evaluate the efficiency of the scanning strategies used by participants.

To investigate the relationship between RTs and eye movements, correlations between RTs and fixation count and between RTs and median fixation duration were run for each group and for each experimental display. In case an association was found, it was planned to conduct the same analyses conducted for the RT slopes.

To answer RQ3, the cross-sectional developmental trajectories of accuracy rates, RTs, fixation count, and median fixation duration against CA were investigated through curve-fitting analyses selecting from linear, logarithmic, and power models. In case of lack of significant models, the relationship between these variables was investigated through visual inspection of the corresponding scatterplots.

2.3 Results

2.3.1 Description of the sample

Table 2.3 provides descriptive statistics for each group for CA, non-verbal intelligence (RCPM), calculation skills (WIAT-NO) and verbal counting performance.

Table 2.3. Descriptive statistics of participants.

Group		CA		RCPM		WIAT-NO		Verbal counting	
	N	*M* (*SD*)	Range	*M* (*SD*)	Range	*M* (*SD*)	Range	*M* (*SD*)	Range
DS	24	259 (132)	104 – 590	15.42 (6.32)	4 – 25	7.67 (2.82)	4 – 16	16.87 (4.88)	2 – 20
WS	25	219 (144)	96 – 620	15.44 (4.34)	7 – 25	8.88 (3.05)	3 – 17	19.58 (1.50)	14 – 20
TD	25	62 (8)	47 – 79	15.48 (4.87)	4 – 25	7.60 (2.02)	3 – 11	18.96 (3.17)	6 – 20

Note: CA reported in months. CA: Chronological age. RCPM: Raven's coloured progressive matrices. WIAT-NO: Wechsler Individual Achievement Test – Numerical Operation scale. DS: Down syndrome. WS: Williams syndrome. TD: Typically developing control group.

There were significant differences between the three groups for CA, Welch’s $F(2, 73) = 21.01$, $p < .001$. Games-Howell post hoc analyses revealed that the TD group (M = 62 months, SD = 8) was significantly younger than the WS group (M = 219 months, SD = 144, 156.76, 95% CI [84.55, 228.97], $p < .001$) as well as the DS group (M = 259 months, SD = 132, 197.04, 95% CI [129.30, 264.77], $p < .001$). CA of the WS group and of the DS group was not significantly different (40.28, 95% CI [-55.42, 135.98], $p = .569$).

Because of the matching criteria used for this study, there were no significant differences between the DS, WS, and TD groups for non-verbal cognitive ability measured with RCPM, Welch's $F(2,46.00) = .001$, $p = .999$.

There were no significant differences between DS, WS, and TD groups for calculation skills, measured by the WIAT-NO test, $F(2,45.43) = 1.62$, $p = .208$. All participants were able to correctly discriminate and recognise digits, to count using the 1:1 correspondence and cardinality principles, and most of them were able to solve one-digits additions and subtractions. Fewer participants were able to solve two-digits additions and subtractions, multiplications, and divisions.

There was a significant difference between the three groups for the verbal counting task, $F(2,36.52) = 3.67$, $p = .035$. Post hoc analyses showed that DS group ($M = 16.87$, $SD = 4.88$) had a statistically significant lower score than the WS group ($M = 19.58$, $SD = 1.50$, -2.685, 95% CI [-5.27, -.10], $p = .040$), but no other differences were statistically significant.

2.3.2 Enumeration processes

2.3.2.1 Accuracy rates

Table 2.4 shows the percent accuracy rates for each experimental display, by groups. Overall accuracy in the enumeration task was high. Analysis of the accuracy rates based on the number of dots presented indicates that accuracy rates decreased for all groups as the numerosity of the dots increased. Analysis of the accuracy rates based on the arrangement of dots shows higher accuracy rates for the dice pattern condition where accuracy rates ranged between 83% and 100%, while accuracy rates for the random pattern conditions ranged between 58% and 100%.

Statistical analysis of errors was not considered due to the low error rates.

Table 2.4. Percent accuracy rates.

Dots presented	Dice pattern			Random pattern		
	DS	WS	TD	DS	WS	TD
1	100%	100%	100%	100%	100%	100%
2	100%	96%	100%	100%	100%	100%
3	96%	100%	96%	92%	100%	96%
4	83%	96%	88%	88%	92%	96%

5	88%	88%	92%	79%	80%	80%
6	88%	88%	84%	58%	76%	71%

A Kruskal-Wallis *H* test was conducted for each experimental condition to determine if there were differences in accuracy rates between groups. Distributions of accuracy rates were similar for all groups in all conditions, as assessed by visual inspection of boxplots. For all experimental conditions the median accuracy rates were not statistically significantly different between the three groups (D1-D3: $\chi^2(2) = .59$, $p = .745$; D4-D6: $\chi^2(2) = 2.16$, $p = .340$; R1-R3: $\chi^2(2) = .58$, $p = .784$; R4-R6: $\chi^2(2) = .85$, $p = .655$).

For each group a series of Wilcoxon signed-rank tests was used to determine whether there was a median difference in accuracy between paired experimental conditions (Table 2.5). The pattern of significance in the accuracy rates was similar for the three groups with a few exceptions. First, only the DS group reported significantly different accuracy rates when enumerating D4-D6 compared to the accuracy rates in R1-R3 and in R4-R6. Moreover, the WS group reported no significant differences in accuracy when enumerating D1-D3 dots and D4-D6 dots, contrary to the other two groups that reported lower accuracy rates for the D4-D6 condition.

Table 2.5. Wilcoxon signed-rank test on accuracy rates between paired experimental conditions by group.

DS	D1-D3	D4-D6	R1-R3	R4-R6
D1-D3	—	$z = -2.26$, $p = .024$*	$z = -.58$, $p = .564$	$z = -2.92$, $p = .004$*
D4-D6		—	$z = -1.99$, $p = .046$*	$z = -2.33$, $p = .020$*
R1-R3			—	$z = -2.80$, $p = .005$*
R4-R6				—

WS	D1-D3	D4-D6	R1-R3	R4-R6
D1-D3	—	$z = -1.86$, $p = .063$	$z = 1.00$, $p = .317$	$z = -2.36$, $p = .018$*
D4-D6		—	$z = -1.89$, $p = .059$	$z = -1.04$, $p = .301$
R1-R3			—	$z = -2.57$, $p = .010$*
R4-R6				—

TD	D1-D3	D4-D6	R1-R3	R4-R6
D1-D3	—	$z = -2.27$, $p = .023$*	$z = .00$, $p = 1.000$	$z = -3.13$, $p < .001$*
D4-D6		—	$z = -1.93$, $p = .054$	$z = -.98$, $p = .329$

R1-R3	—	z = -3.13, p = .002*
R4-R6		—

Note: *p < .05.

For each group Spearman's correlation analyses were run to investigate the relationship between mean accuracy rates and CA, mean accuracy rates and MA measured by RCPM, and mean accuracy rates and calculation skills measured by the WIAT-NO test (Table 2.6). The analyses were only conducted for the experimental conditions D4-D6 and R4-R6, as the accuracy for the 1-3 numerosity range for both the dice and the random condition was close to ceiling for all the groups.

A strong positive correlation was found between CA and D4-D6 accuracy only for the WS group, $rs(25) = .61$, $p = .001$. Moreover, a moderate positive correlation between CA and R4-R6 accuracy was found for both the TD group, $rs(25) = .46$, $p = .019$ and the DS group, $rs(24) = .47$, $p = .021$.

RCPM scores correlated only with R4-R6 accuracy rates for all groups, TD: $rs(25) = .44$, $p = .028$; WS: $rs(25) = .45$, $p = .023$; DS $rs(24) = .48$, $p = .017$.

Finally, for all groups a moderate to strong positive correlation between WIAT-NO and mean accuracy was found in D4-D6 conditions as well as in R4-R6 conditions, D4-D6: TD: $rs(25) = .61$, $p = .001$; WS: $rs(25) = .56$, $p = .004$; DS $rs(24) = .42$, $p = .042$; R4-R6: TD: $rs(25) = .41$, $p = .040$; WS: $rs(25) = .40$, $p = .046$; DS $rs(24) = .55$, $p = .005$.

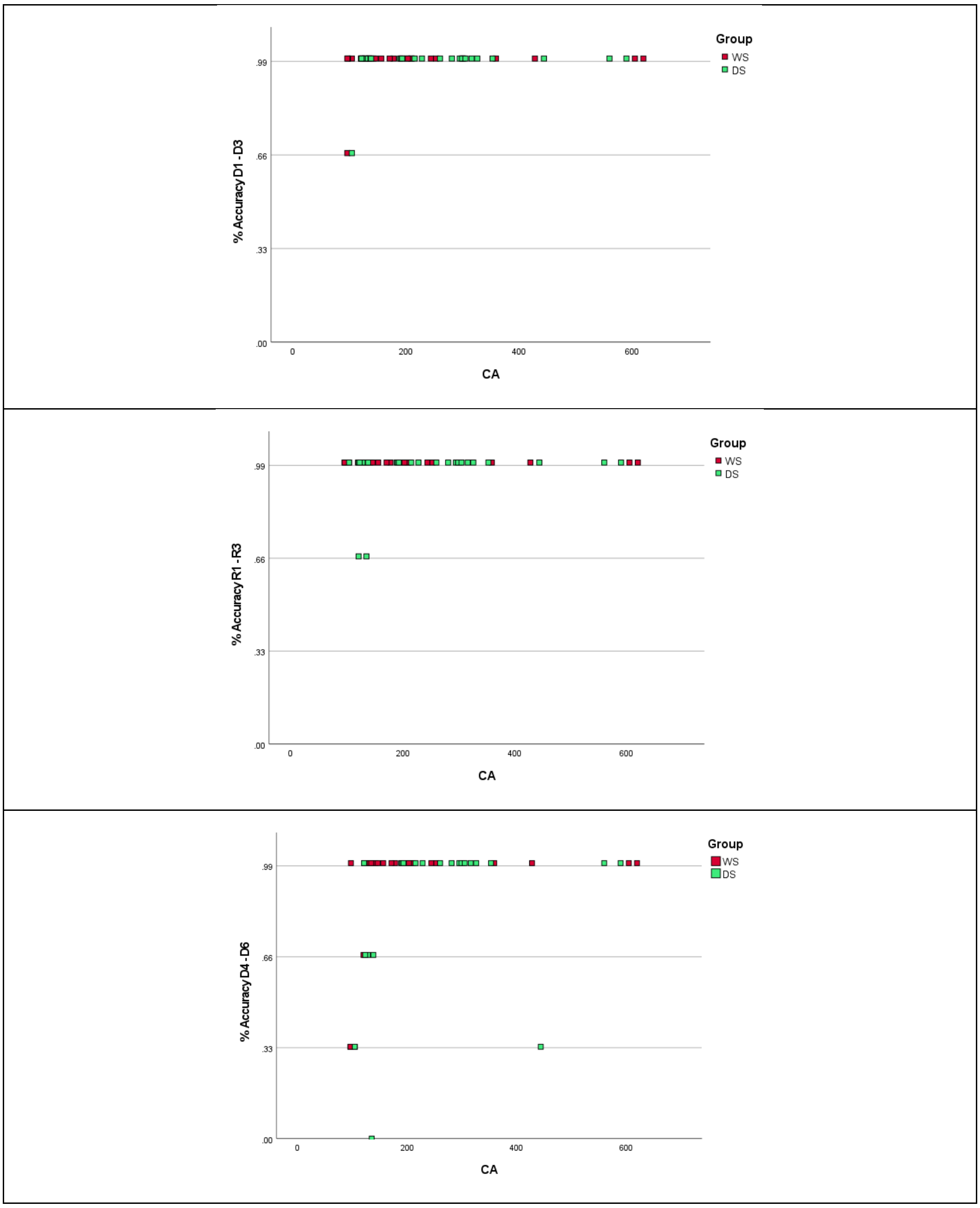
Table 2.6. Spearman's correlations for study variables.

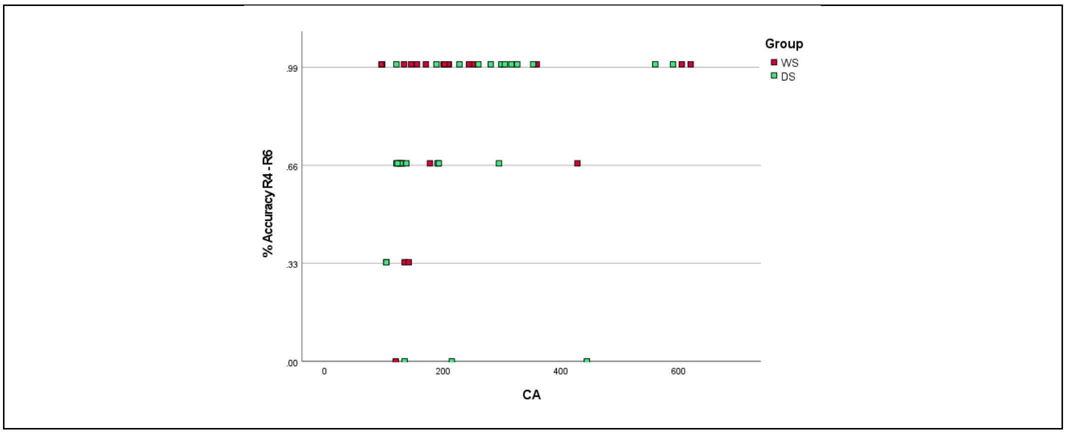
	Mean D4 – D6 Accuracy			Mean R4 – R6 Accuracy		
	DS	WS	TD	DS	WS	TD
CA	.37	.61**	.24	.47*	.34	.46*
RCPM	.38	.38	.14	.48*	.45*	.44*
WIAT- NO	.42*	.56**	.61**	.55**	.40*	.41*

Note: *p < .05. **p < .001. CA = Chronological age. RCPM = Raven's coloured progressive matrices test. WIAT-NO = Numerical operation scale of the Wechsler individual achievement test.

Finally, the scatterplots in Figure 2.6 show, for each clinical group, the relationship between percentage accuracy rates averaged between the four experimental conditions for each participant and CA. Visual inspection of the charts shows the above-mentioned high accuracy rates reported by both clinical groups, in particular for the D1-D3 and R1-R3 conditions. Furthermore, the scatterplots show that in both groups younger individuals made more mistakes than older individuals when subitizing. Moreover, there are instances of individuals with DS and with WS making mistakes when counting regardless of their age, with individuals with DS reporting more mistakes than individuals with WS. Statistical analyses were not conducted, and further inferences on the relationship between accuracy rates and CA could not be made due to the low variability in the percentage accuracy scores of the two clinical groups.

Figure 2.6. Association between percentage accuracy rates and CA by group split by condition.





Notes: accuracy rates are averaged for participants.

2.3.2.2 Reaction Times

Table 2.7 shows *M*s and *SD*s of RTs organised by experimental display and by group.

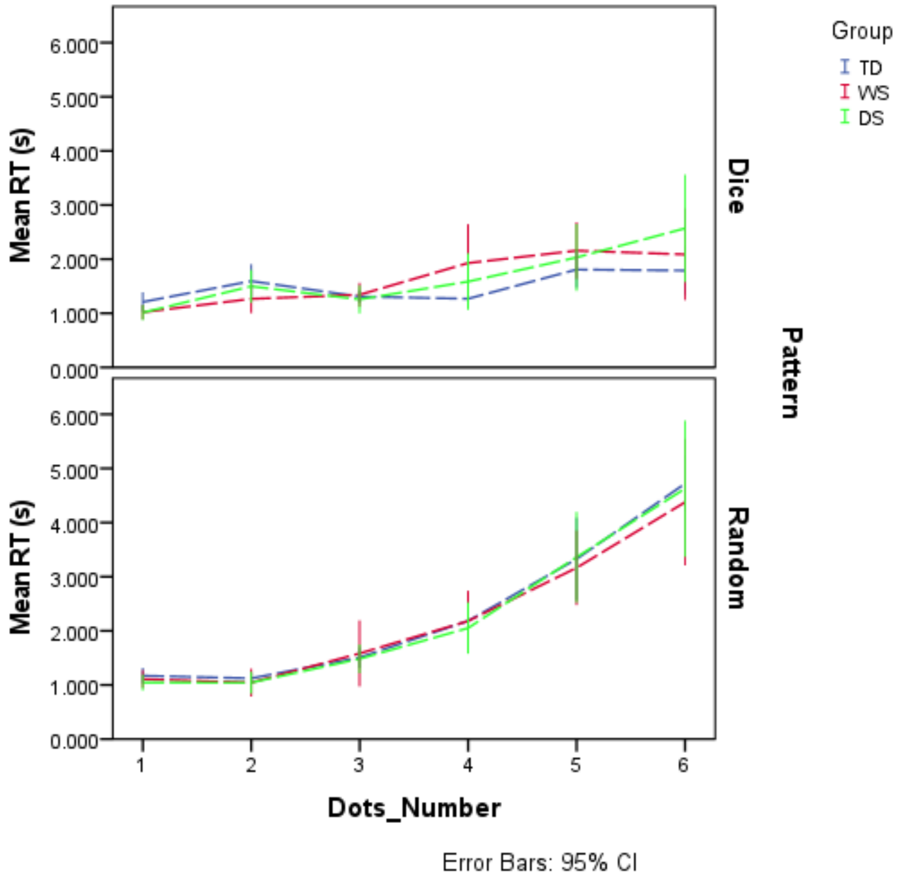
Table 2.7. RT Mean and Standard Deviation.

Dots presented	Dice pattern			Random pattern		
	DS	WS	TD	DS	WS	TD
1	1.01 (0.33)	1.02 (0.29)	1.21 (0.42)	1.05 (0.35)	1.08 (0.42)	1.17 (0.34)
2	1.50 (0.66)	1.22 (0.63)	1.60 (0.72)	1.04 (0.46)	1.05 (0.60)	1.12 (0.23)
3	1.26 (0.58)	1.34 (0.51)	1.31 (0.28)	1.48 (0.60)	1.58 (1.41)	1.51 (0.43)
4	1.59 (1.11)	1.93 (1.66)	1.27 (0.39)	2.05 (1.04)	2.18 (1.30)	2.18 (1.14)
5	2.03 (1.32)	2.16 (1.19)	1.81 (0.78)	3.36 (1.73)	3.03 (1.52)	3.33 (1.63)
6	2.57 (2.18)	2.00 (1.86)	1.79 (0.99)	4.63 (2.08)	4.38 (2.27)	4.72 (1.00)

Note. *SD* are presented in parenthesis.

Figure 2.7 shows mean RTs for each group for all experimental displays. Visual analysis of the graphs shows a somewhat flat function for both the dice and random conditions for the subitizing range (1 - 3) for all groups. In contrast, there is a steeper increase for the counting range (4 - 6) but only in the random condition, again in all groups. The increase observed in RTs for the higher numerosity but only with the random pattern suggests that participants were perceptually subitizing up to 3 dots in both conditions and conceptually subitizing 4 to 6 dots in the dice condition. Instead, participants were engaged in counting when enumerating 4 to 6 dots in the random condition. Moreover, Table 2.7 and Figure 2.7 show higher *SD* for both WS and DS groups for the experimental display R6, which were more than twice as large as the ones reported for the TD group. The same pattern is observed for the DS group for the D6 display.

Figure 2.7. Mean RT by group for each experimental display.



A three-way mixed ANOVA with group as between subject-factor and range and pattern as within-subject factors was conducted to understand the effects of pattern, number of dots, and group on RTs to perform the enumeration task. There was a significant main effect of pattern $F(1, 27) = 51.84$, $p < .001$, with higher RTs for the random condition. There was a significant main effect of number of dots $F(5, 135) = 47.71$, $p < .001$, with higher RTs for higher number of dots. There was not a statistically significant main effect of group, $F(1, 27) = .42$, $p = .659$. There was no statistically significant three-way interaction between pattern, range and group, $F(2, 65) = 2.14$, $p = .126$. There was a significant two-way interaction observed between pattern and number of dots for all the groups, $F(5, 135) = 18.33$, $p < .001$. Visual inspection of the charts showed higher RTs for higher number of dots for the random level, while RTs did not increase for the dice level of the within-subject factor pattern.

Statistical analysis of the slopes for RTs confirmed the results of the visual analysis of the graphs. Table 2.8 shows the averaged slopes and intercepts of individual regressions lines for RTs for each group. All groups showed average RT slopes for the D1-D3, D4-D6, and R1-R3 close to 0, indicating an almost-flat line. On the other hand, the average slopes for RTs for the R4-R6 condition were between 1.06 and 1.56, indicating a positive relationship between RTs and the number of dots to be counted.

Table 2.8. Average RT slopes, intercepts and R^2 of regression lines separately computed for each experimental condition by group.

Measure	D1 - D3			D4 - D6		
	DS	WS	TD	DS	WS	TD
Slope	0.18 (0.33)	0.12 (0.28)	0.02 (0.23)	0.37 (0.85)	0.37 (0.91)	0.24 (0.47)
Intercept	0.92 (0.49)	0.95 (0.49)	1.33 (0.67)	0.07 (3.66)	0.14 (3.84)	0.48 (2.15)
R^2	.54	.61	.45	.62	.62	.53
N	24	23	25	21	22	23

Measure	R1 - R3			R4 - R6		
	DS	WS	TD	DS	WS	TD
Slope	0.19 (0.31)	0.25 (0.67)	0.14 (0.24)	1.56 (1.76)	1.18 (1.08)	1.06 (0.86)
Intercept	0.79 (0.54)	0.71 (0.99)	0.99 (.50)	-4.19 (7.53)	-2.68 (4.67)	-2.02 (4.22)
R^2	.60	.60	.55	.94	.80	.80
N	24	24	24	20	21	22

Note. *SDs* are presented in parenthesis.

RT slopes were submitted to separate one-way Welch ANOVAs which showed no statistically significant differences between the three groups for all experimental conditions (*F* values between 0.30 and 2.13, all *p*s > .130). This suggests that the same enumeration processes were employed by all the participants in each experimental condition, regardless of their group.

Moreover, Wilcoxon signed-rank tests on the group averaged slopes for RT were used to determine whether there were any median differences between paired experimental conditions (Table 2.9). Results showed that for all groups average RT slopes for the experimental condition R4-R6 were significantly different from all the other conditions, and they were reported to be significantly higher. This confirms that when participants were shown 4 to 6 randomly displayed dots, they were using a different enumeration process (counting) than in the other experimental conditions, when they were subitizing.

Table 2.9. Wilcoxon signed-rank test on mean RT slopes for all experimental conditions by group.

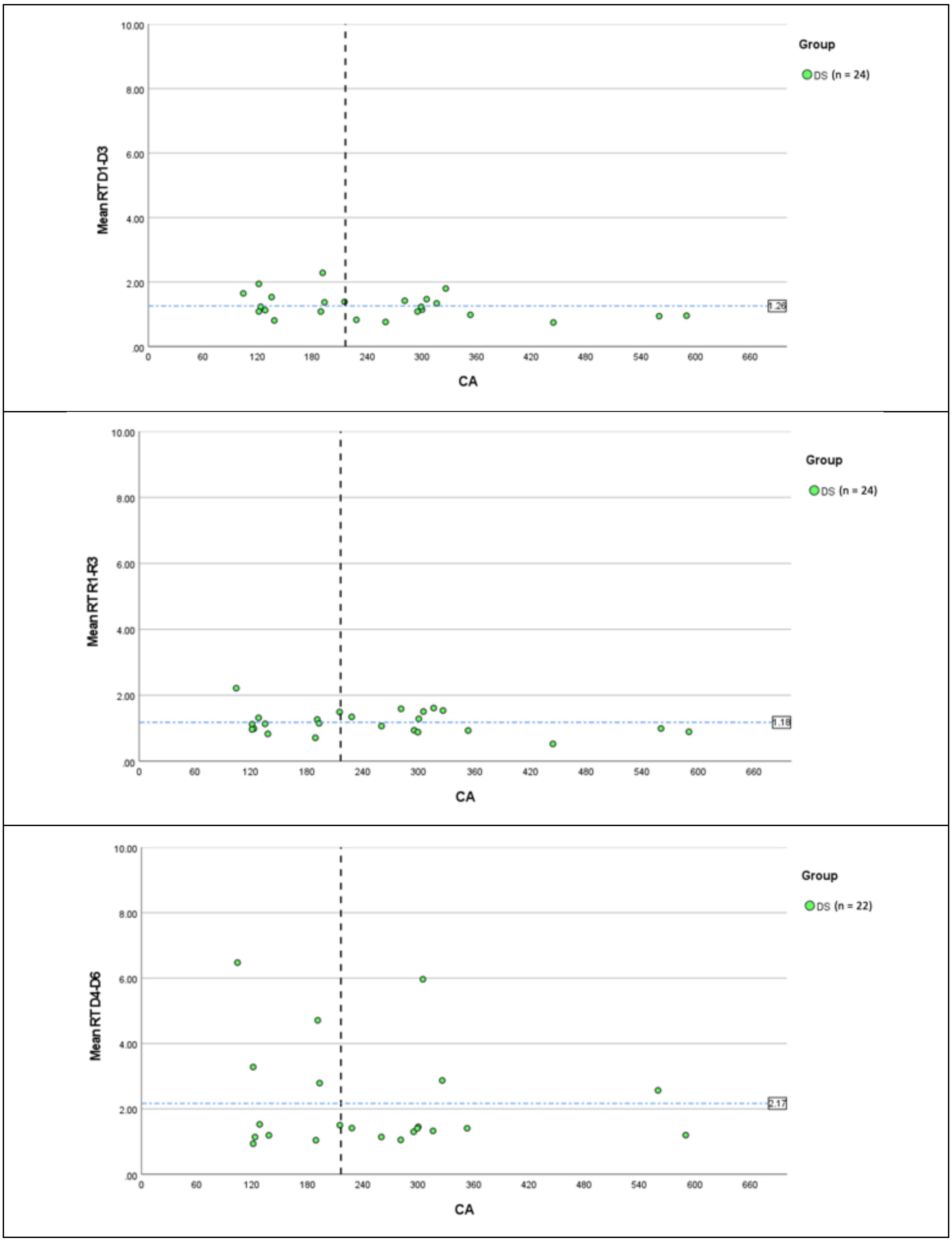
DS	D1-D3	D4-D6	R1-R3	R4-R6
D1-D3	—	$z = 0.33$, $p = .741$	$z = 0.71$, $p = .475$	$z = 3.02$, $p = .002$*
D4-D6		—	$z = -0.40$, $p = .689$	$z = 2.88$, $p = .004$*
R1-R3			—	$z = 3.14$, $p = .002$*
R4-R6				—
WS	D1-D3	D4-D6	R1-R3	R4-R6
D1-D3	—	$z = 0.43$, $p = .664$	$z = 0.43$, $p = .670$	$z = 3.21$, $p = .001$*
D4-D6		—	$z = -1.09$, $p = .277$	$z = 2.54$, $p = .011$*
R1-R3			—	$z = 2.62$, $p = .009$*
R4-R6				—
TD	D1-D3	D4-D6	R1-R3	R4-R6
D1-D3	—	$z = 1.31$, $p = .191$	$z = 1.91$, $p = .056$	$z = 3.46$, $p = .001$*
D4-D6		—	$z = -0.05$, $p = .961$	$z = 3.25$, $p = .001$*
R1-R3			—	$z = 3.49$, $p < .001$**
R4-R6				—

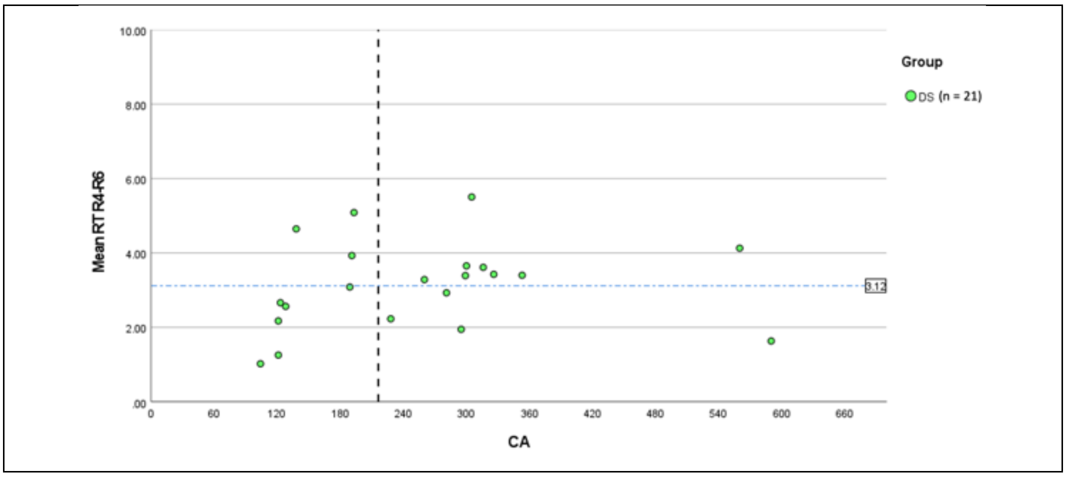
Note: *$p < .05$. ** $p < .001$.

The curve-fitting analyses of mean RTs against CA did not generate significant models for any of the experimental conditions and for any of the clinical groups (DS: F values between 0.25 and 3.25, all ps > .085, WS: F values between 0.02 and 1.73, all ps > .202), indicating that the data in the current study do not provide evidence for an effect of CA over mean RTs.

Figure 2.8 and in Figure 2.9 show the relationship between mean RTs and CA in each experimental condition separately for each clinical group. Visual inspection of the charts for the DS group shows similar levels of variability for children and adults in all experimental conditions (Figure 2.8). Moreover, the scatterplots highlight a higher variability for the counting range conditions than for the subitizing range conditions.

Figure 2.8. Association between mean RT and CA for the DS group split by condition.

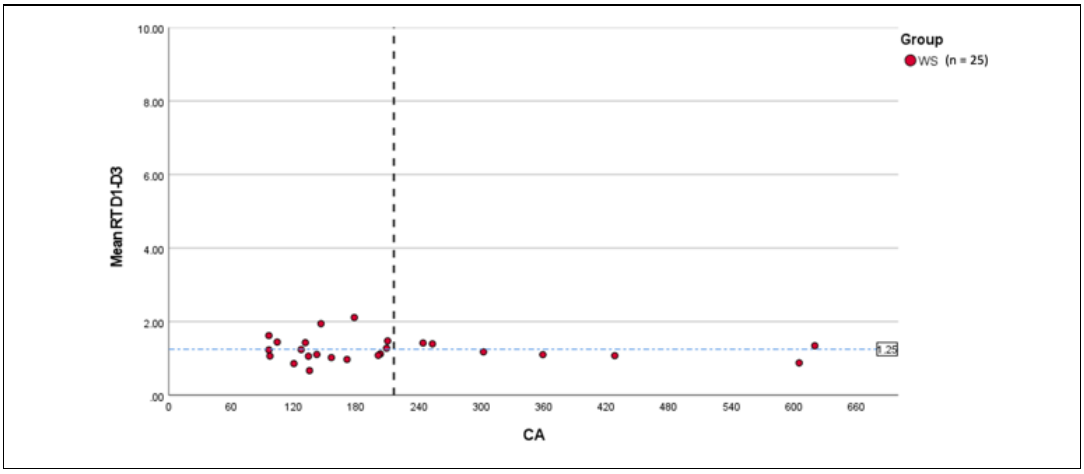


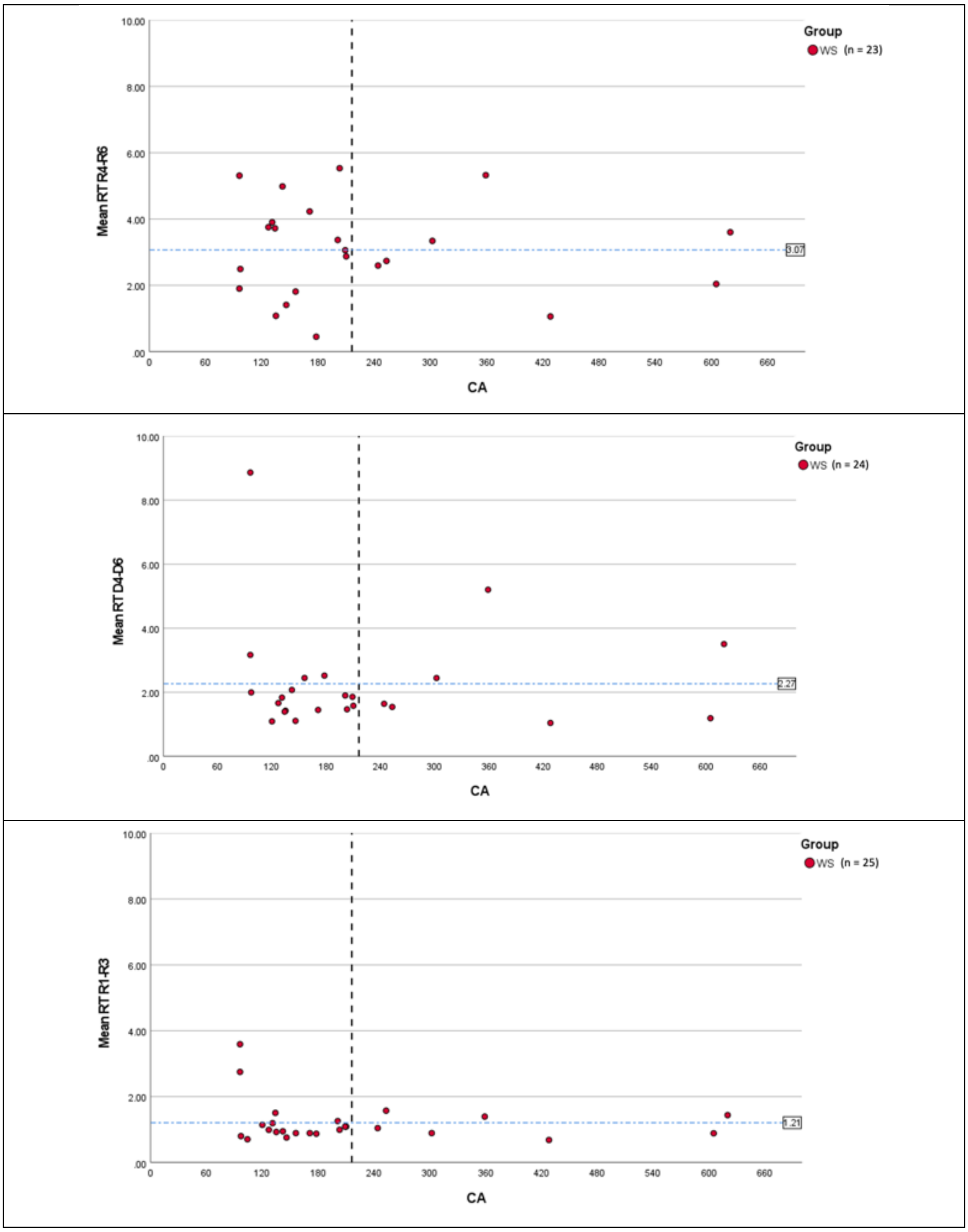


Note: The dashed black vertical line x = 216 represents the criterion used to distinguish between children and adults, that is CA = 18 years. The horizontal blue dotted line is set at the group mean RT for each experimental condition.

Visual inspection of the charts for the WS group shows similar levels of variability for children and adults for the D1-D3 and R1-R3 conditions, where the dots lie around the group mean (Figure 2.9). The D4-D6 condition shows higher levels of variability in adult participants. The R4-R6 condition reports high levels of variability for both young and adult participants.

Figure 2.9. Association between mean RT and CA for the WS group split by condition.





Note: The dashed black vertical line x = 216 represents the criterion used to distinguish between children and adults, that is CA = 18 years. The horizontal blue dotted line is set at the group mean RT for each experimental condition.

2.3.3 Eye gaze behaviour

2.3.3.1 Fixation count

Table 2.10 shows *M*s and *SD*s of fixation count organised by display and by group.

Table 2.10. Fixation count Means and Standard Deviations.

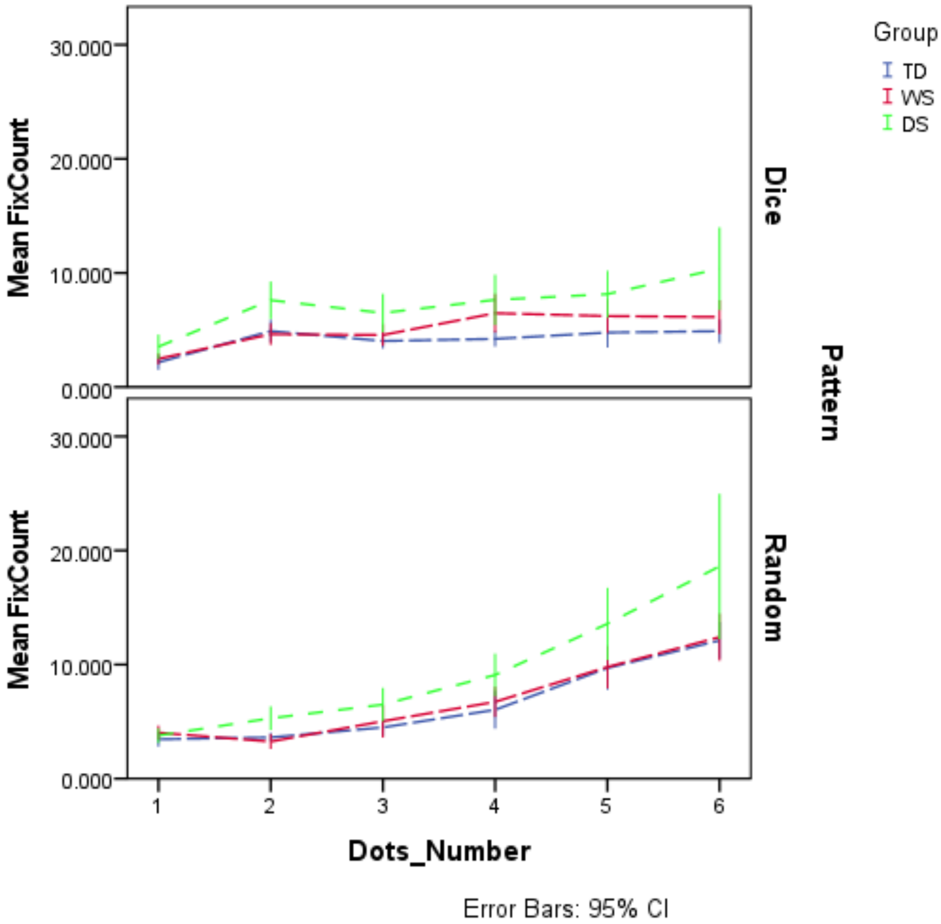
Dots presented	Dice pattern			Random pattern		
	DS	WS	TD	DS	WS	TD
1	3.54 (2.50)	2.48 (1.08)	2.17 (1.52)	3.79 (1.50)	4.04 (1.50)	3.46 (1.50)
2	7.62 (3.60)	4.64 (2.12)	4.19 (2.30)	5.30 (2.40)	3.26 (1.51)	3.63 (0.92)
3	6.50 (3.79)	4.57 (2.19)	4.04 (1.55)	6.50 (3.29)	5.04 (3.31)	4.50 (1.68)
4	7.65 (4.71)	6.48 (3.94)	4.23 (1.54)	9.10 (4.13)	6.74 (3.71)	6.04 (3.80)
5	8.15 (4.42)	6.23 (3.35)	4.78 (3.06)	13.58 (6.55)	9.79 (3.85)	9.70 (4.05)
6	10.38 (7.97)	6.14 (3.21)	4.91 (2.23)	18.62 (10.54)	12.41 (4.02)	12.12 (3.12)

Note. *SDs* are presented in parenthesis.

Inspection of mean fixation count on Table 2.10 shows that for both clinical groups, the average number of fixations is consistently higher than the number of dots being enumerated. The range of mean fixation count for the condition R4-R6 is wider than the others (DS: 9.10 to 18.62; WS: 6.74 to 12.41; TD: 6.04 to 12.12).

This data is visually represented in Figure 2.10, which shows mean fixation count for each group for all the experimental displays. Visual analysis of the graphs shows for all groups a similar pattern to the one that was observed for RTs, characterised by a flat function for both the dice and random conditions for the subitizing range (1 - 3) and by a steeper increase of fixation count for the counting range (4 - 6) but only in the random condition. Moreover, Table 2.11 and Figure 2.10 show higher *SD* values for the DS group for most of the experimental displays when compared to both WS and TD groups.

Figure 2.10. Mean fixation count by group.



A three-way mixed ANOVA with group as between subject-factor and range and pattern as within-subject factors was conducted to understand the effects of pattern, number of dots, and group on fixation count. There was a significant main effect of pattern $F(1, 27) = 77.69$, $p < .001$, with higher fixation count for the random condition. There was a significant main effect of number of dots $F(5, 135) = 52.05$, $p < .001$, with higher fixation count for higher number of dots. There was a significant main effect of group, $F(1, 27) = 7.09$, $p = .003$. Games-Howell post hoc analyses revealed that fixation count of the DS group was significantly higher than fixation count of both the WS group ($p = .012$) and the TD group ($p = .001$). There was no statistically significant three-way interaction between pattern, number of dots and group, $F(2, 27) = 1.17$, $p = .325$. There was a significant two-way interaction observed between pattern and number of dots for all the groups, $F(5, 135) = 23.92$, $p < .001$. Visual inspection of the charts showed higher fixation count for higher number of dots for the random level, while fixation count did not increase for the dice level of the within-subject factor pattern.

Table 2.11 shows that, as observed for the RTs, average fixation count slopes for the D1-D3, D4-D6, and R1-R3 (between 0.43 and 1.73) were smaller than the average fixation count slopes for R4-R6 (between 2.90 and 5.35), suggesting that participants were using different enumeration processes.

Table 2.11. Average fixation count slopes, intercepts and R^2 of regression lines separately computed for each experimental condition by group.

Measure	D1 – D3			D4 – D6		
	DS	WS	TD	DS	WS	TD
Slope	1.73 (2.03)	0.80 (1.59)	0.84 (1.37)	1.14 (4.07)	0.52 (1.70)	0.52 (1.52)
Intercept	2.51 (4.10)	2.36 (3.51)	2.13 (3.49)	2.90 (18.41)	3.57 (7.98)	2.00 (7.46)
R^2	.63	.63	.59	.50	.56	.44
N	24	22	25	21	22	23

Measure	R1 – R3			R4 – R6		
	DS	WS	TD	DS	WS	TD
Slope	1.27 (1.86)	0.43 (1.66)	0.55 (1.06)	5.35 (6.14)	3.21 (2.28)	2.90 (2.86)
Intercept	2.53 (2.94)	3.24 (2.67)	2.73 (2.30)	-12.14 (25.83)	-6.10 (10.21)	-5.84 (13.93)
R^2	.67	.51	.48	.91	.79	.80
N	22	23	21	20	21	20

Note. *SD* are presented in parenthesis.

Fixation count slopes were submitted to separate one-way Welch ANOVAs that showed no statistically significant differences between the three groups for all the experimental conditions (*F* values between .22 and 1.96, all *p*s > .160), in line with the results reported for the RT slopes.

The pattern of results of the Wilcoxon signed-rank tests run separately for the three groups to compare the averaged slopes for fixation count between paired conditions differed from the pattern of results reported for RTs. As shown in Table 2.12, significant differences in mean fixation count slope were found between two experimental conditions where the same enumeration process (subitizing) was observed for RTs: D1-3 vs D4-6 for DS group (z = -2.11, p = .035), and D1-D3 vs R1-R3 for WS group (z = -2.52, p = .012). In line with RTs data, a statistically significant difference in fixation count slopes was observed when participants were enumerating 4 to 6 dots in the random condition, compared to all the other experimental conditions.

Table 2.12. Wilcoxon signed-rank test on average fixation count slopes for all experimental conditions by group.

DS	D1-D3	D4-D6	R1-R3	R4-R6
D1-D3	—	*z* = -2.11, *p* = .035*	*z* = -1.65, *p* = .099	*z* = 1.96, *p* = .050*
D4-D6		—	*z* = 0.44, *p* = .662	*z* = 2.90, *p* = .004*
R1-R3			—	*z* = 2.62, *p* = .009*
R4-R6				—

WS	D1-D3	D4-D6	R1-R3	R4-R6
D1-D3	—	*z* = -1.95, *p* = .051	*z* = -2.52, *p* = .012*	*z* = 2.84, *p* = .004*
D4-D6		—	*z* = -0.73, *p* = .467	*z* = 3.40, *p* = .001*
R1-R3			—	*z* = 3.10, *p* = .002*
R4-R6				—

TD	D1-D3	D4-D6	R1-R3	R4-R6
D1-D3	—	*z* = -1.80, *p* = .072	*z* = -1.56, *p* = .119	*z* = 2.51, *p* = .012*
D4-D6		—	*z* = 1.24, *p* = .217	*z* = 3.18, *p* = .001*
R1-R3			—	*z* = 2.68, *p* = .007*
R4-R6				—

Note: **p* < .05.

As shown in Table 2.13, Spearman's correlations between fixation count and RTs showed a statistically significant, positive association for all groups, with a few exceptions. No significant correlations were found for the experimental display D1 for all groups, DS: *rs*(24) = .21, *p* = .331; WS: *rs*(21) = .31, *p* = .169; TD: *rs*(24) = .25, *p* = .247 for the experimental display R1 for the DS group, *rs*(24) = .26, *p* = .218, and for experimental display R2 for the TD and the WS groups, TD: *rs*(24) = -.08, *p* = .715; WS: *rs*(21) = .39, *p* = .070.

Table 2.13. Spearman's correlations between fixation count and RT.

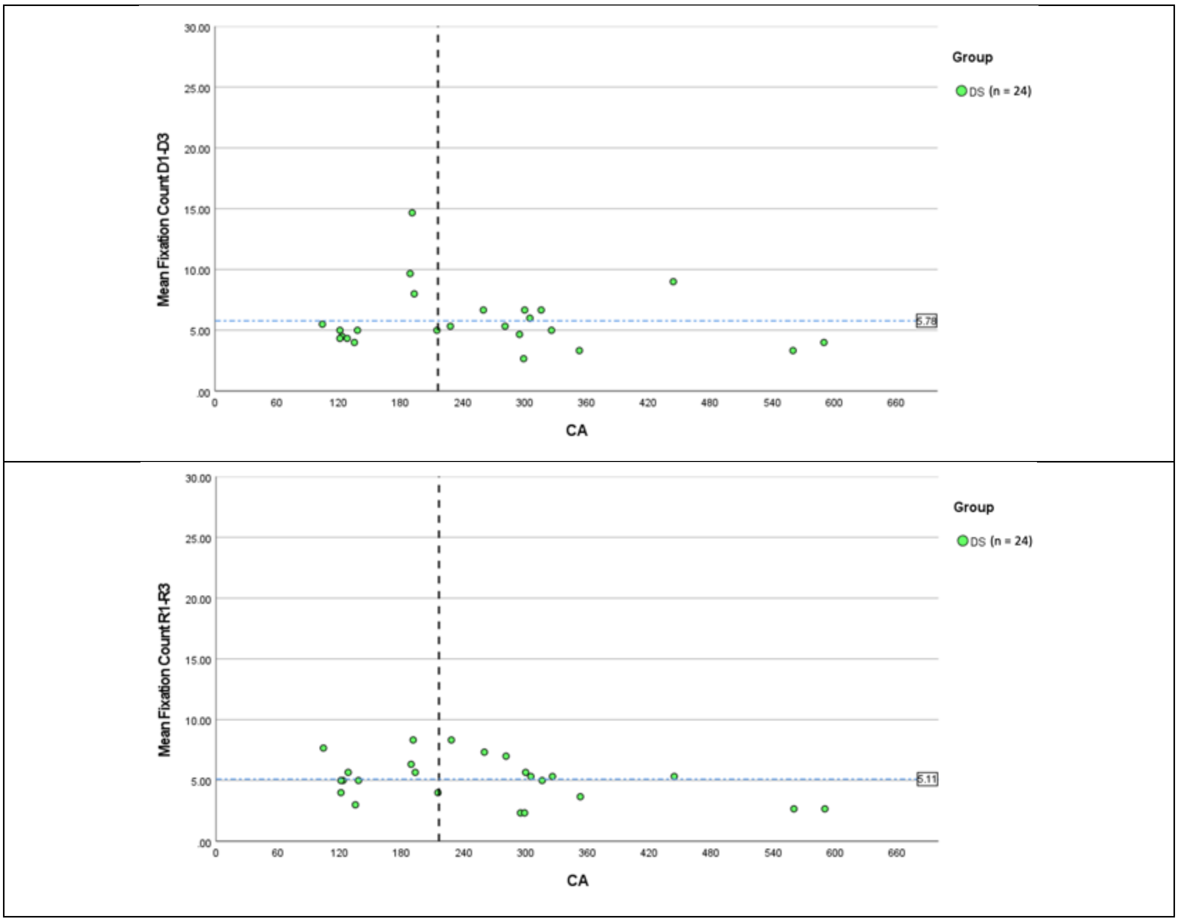
Dots presented	Dice pattern			Random pattern		
	DS	WS	TD	DS	WS	TD
1	.21	.31	.25	.26	.50*	.47*
2	.60**	.66**	.76**	.47*	.39	-.08
3	.52*	.56**	.44*	.45*	.43*	.59**
4	.73**	.59**	.71**	.87**	.78**	.85**
5	.54*	.55**	.57**	.78**	.50*	.72**
6	.67**	.77**	.46**	.67*	.63**	.80**

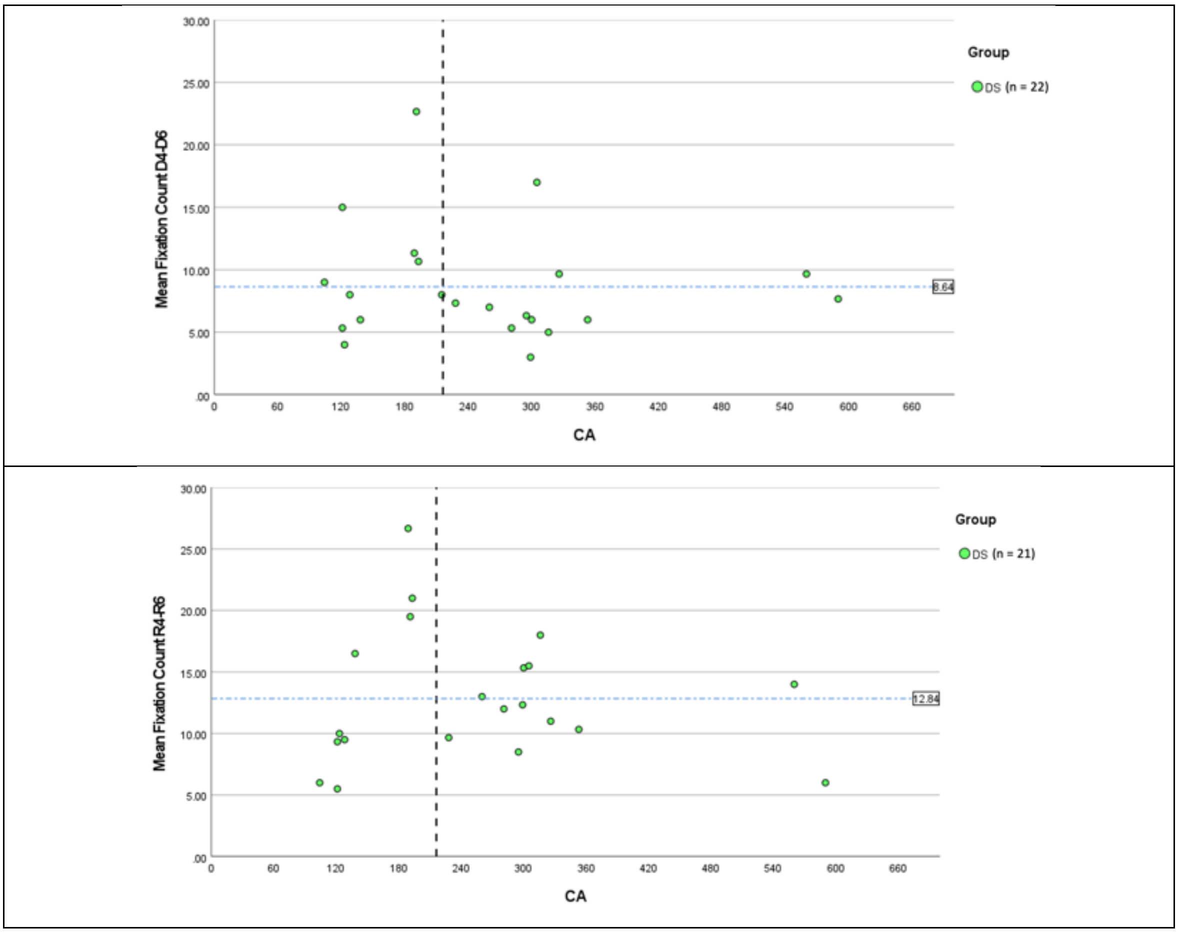
Note: **p* < .05. ***p* < .001.

The curve-fitting analyses of fixation count against CA did not generate any significant model for any of the experimental conditions and for both clinical groups (DS: *F* values between 0.07 and 4.07, all *p*s > .056; WS: *F* values between 0.01 and 0.26, all *p*s > .616), indicating that the data do not provide evidence for an effect of CA over fixation count.

Figure 2.11 and in Figure 2.12 show the relationship between mean fixation count and CA for each experimental condition, separately for the two clinical groups. Visual inspection of the charts shows, for both groups, similar ranges of variability for children and adults in all the experimental conditions. Moreover, a pattern characterised by increased variability in fixation count in the counting range compared to the subitizing range can be observed in both groups. However, this trend is more pronounced in the DS group. In fact, while for D1-D3 and R1-R3 the dots are closer to the group mean, dots are more spread around the group mean when participants with DS use conceptual subitizing and counting processes, independently from their age.

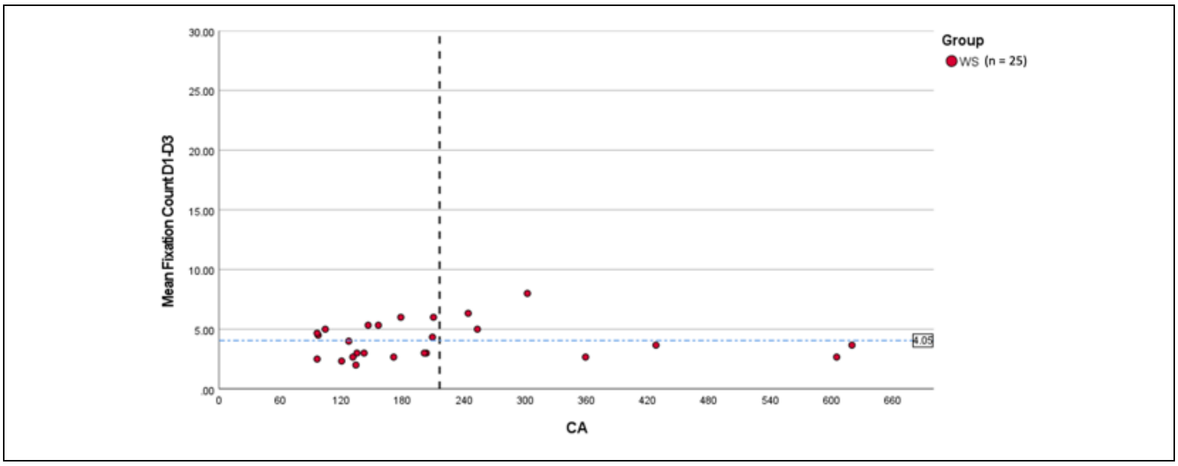
Figure 2.11. Association between mean fixation count and CA for the DS group split by condition.

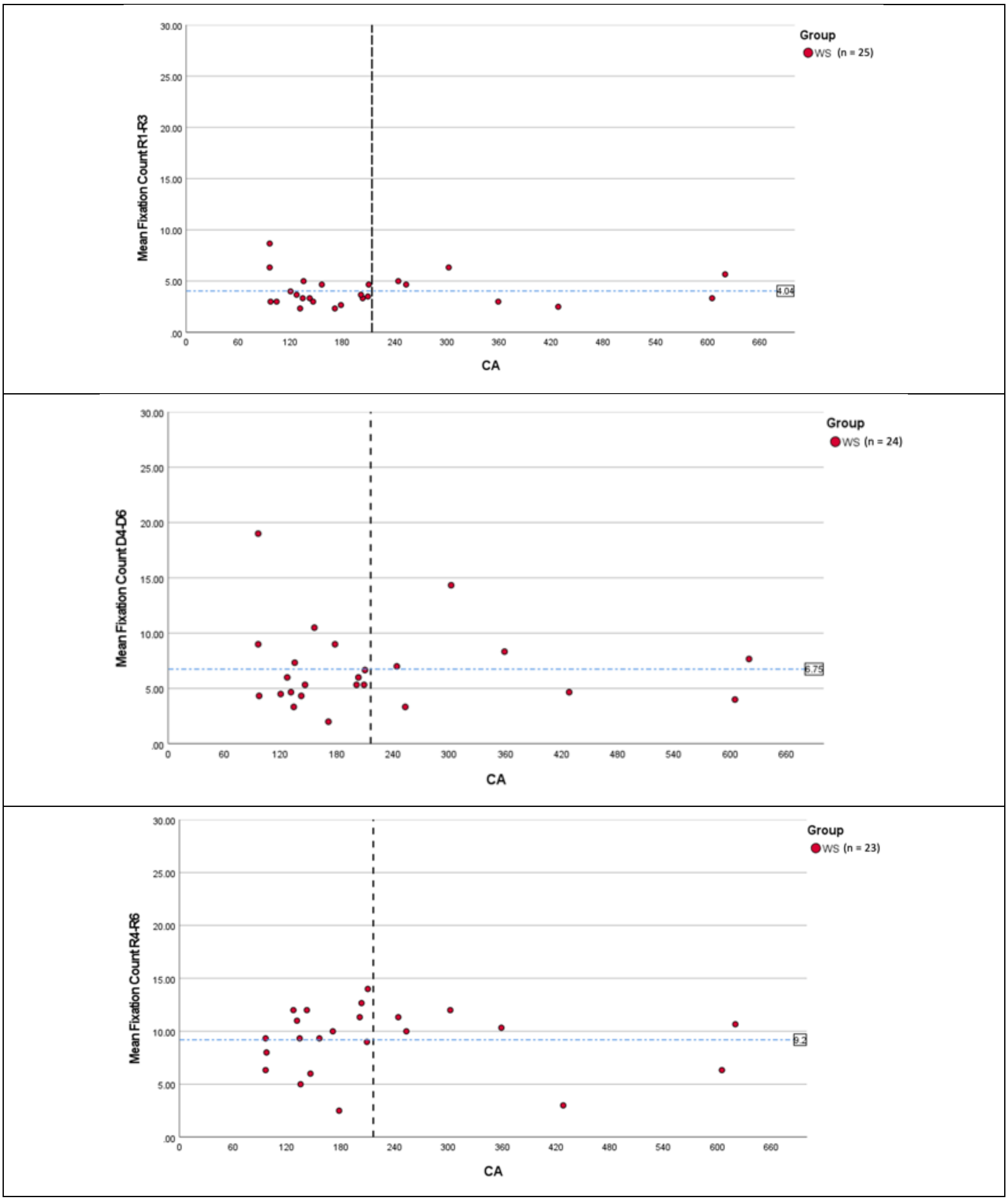




Note: The dashed black vertical line x = 216 represents the criterion used to distinguish between children and adults, that is CA = 18 years. The horizontal blue dotted line is set at the group mean fixation count for each experimental condition.

Figure 2.12. Association between mean fixation count and CA for the WS group split by condition.





Note: The dashed black vertical line x = 216 represents the criterion used to distinguish between children and adults, that is CA = 18 years. The horizontal blue dotted line is set at the group mean fixation count for each experimental condition.

2.3.3.2 Median fixation duration

Table 2.14 shows *M* and *SD* of median fixation duration per group for each experimental display.

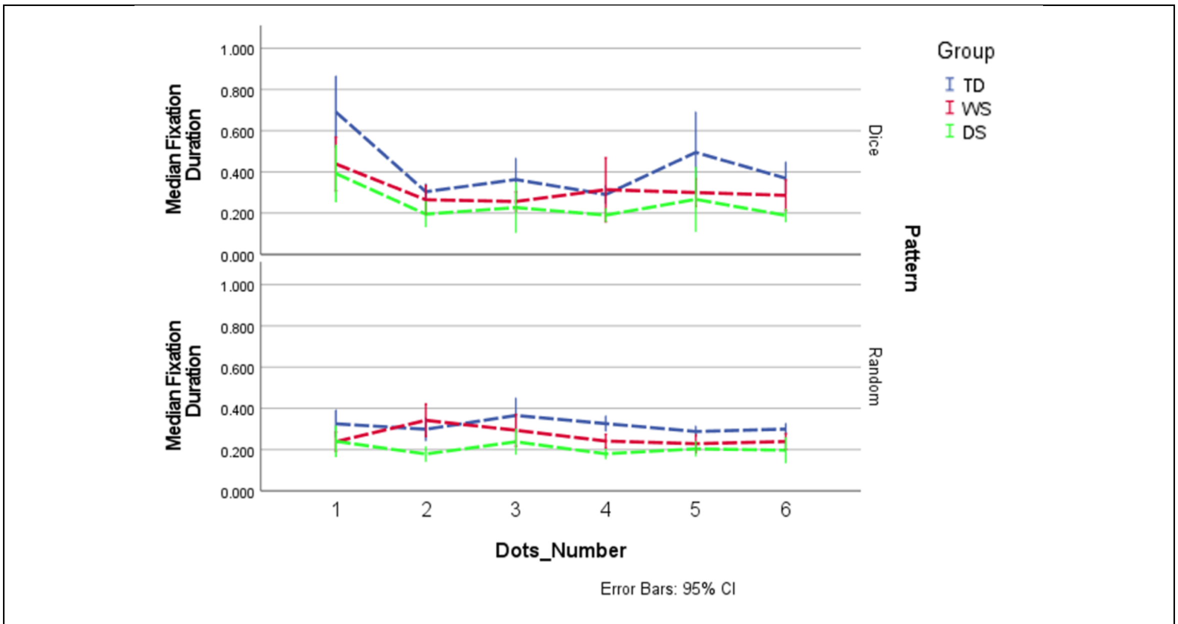
Table 2.14. Median fixation duration Means and Standard Deviations.

Dots presented	Dice Pattern			Random Pattern		
	DS	WS	TD	DS	WS	TD
1	0.39 (0.32)	0.44 (0.29)	0.69 (0.40)	0.24 (0.17)	0.24 (0.11)	0.33 (0.15)
2	0.20 (0.14)	0.27 (0.16)	0.30 (0.09)	0.18 (0.08)	0.34 (0.18)	0.30 (0.13)
3	0.23 (0.27)	0.26 (0.11)	0.36 (0.24)	0.24 (0.14)	0.29 (0.18)	0.37 (0.19)
4	0.19 (0.07)	0.31 (0.35)	0.29 (0.10)	0.18 (0.05)	0.24 (0.08)	0.33 (0.09)
5	0.27 (0.35)	0.30 (0.15)	0.50 (0.44)	0.20 (0.07)	0.23 (0.10)	0.29 (0.05)
6	0.19 (0.07)	0.29 (0.17)	0.37 (0.18)	0.19 (0.11)	0.24 (0.07)	0.30 (0.06)

Note. *SDs* are presented in parenthesis.

Figure 2.13 shows median fixation duration for each group for all experimental displays. Visual analysis of the chart shows that the median fixation duration for the TD group was higher than the one of the two clinical groups for most experimental displays. Moreover, the line representing the median fixation duration for both clinical groups is almost flat, suggesting that the median fixation duration was approximately the same across all the experimental displays.

Figure 2.13. Mean median fixation duration for each experimental display by group.



A three-way mixed ANOVA with group as between subject-factor and number of dots and pattern as within-subject factors was conducted to determine if there were differences in the median fixation duration between the three groups. Visual inspection of Figure 2.13 shows that the median fixation duration seems not to be affected by the different stimuli. Hence, the levels of the within-subject factor for number of dots were reduced to 2 (subitizing range and counting range) instead of the usual 6 levels used in previous analyses. This was done to reduce multiple comparisons which increase the chance of Type 1 error and to increase the power of the analysis. The analysis showed that median fixation durations were significantly different between the three groups, $F(2, 506.82) = 27.90$, $p < .001$. Games-Howell post hoc analyses revealed that the median fixation duration of the TD group was significantly higher than median fixation duration of the WS group ($p < .001$) as well as of the DS group ($p < .001$). Median fixation duration of the WS group was significantly higher than the one of the DS group ($p = .001$). Games-Howell post hoc analyses revealed that the DS group showed significantly shorter fixation duration than the TD group in all the experimental conditions (all *p*s < .001), while the WS showed significantly shorter median fixation duration than the TD group only in the D1-D3 condition ($p = .011$) and in the R4-R6 condition ($p < .001$).

As shown in Table 2.15, no statistically significant association was found between median fixation duration and RTs for TD and DS group for all the experimental stimuli. Instead, for the WS group were found a few moderate positive correlations between median fixation duration and RTs for the experimental conditions D5 and R2, D5: $r_s(22) = .44$, $p = .040$; R2: $r_s(21) = .45$, $p = .041$.

Table 2.15. Spearman's correlations between median fixation duration and RT.

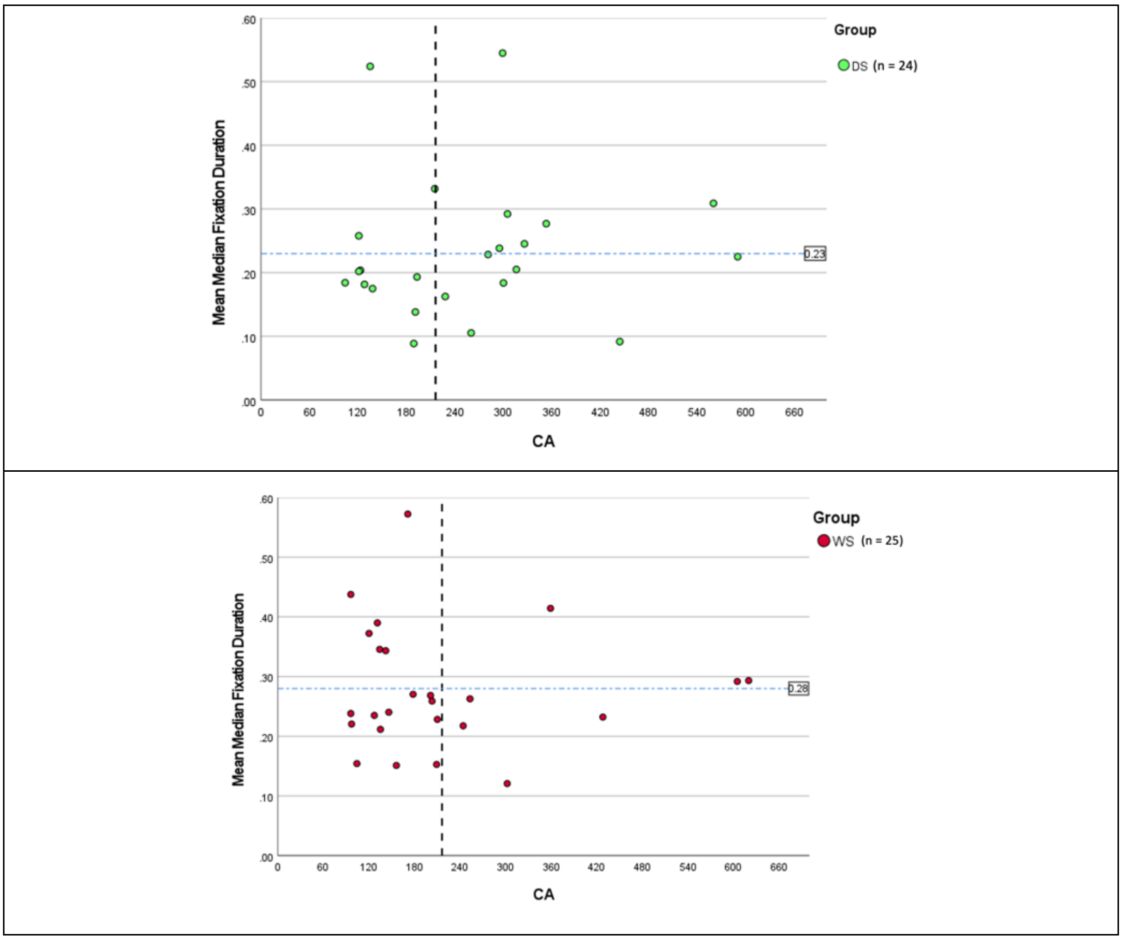
Dots presented	Dice pattern			Random pattern		
	DS	WS	TD	DS	WS	TD
1	.20	-.04	.12	.26	.30	.11
2	.20	-.17	.06	.30	.45*	-.07
3	.40	.10	.08	-.01	.40	-.28
4	.02	.35	-.09	.06	.23	-.05
5	.37	.44*	.30	.25	.22	-.10
6	.26	-.12	.25	.07	.43	.05

Note: *$p < .05$.

The curve-fitting analyses of median fixation duration against CA did not generate any significant model for both clinical groups (DS: F values between 0.05 and 0.06, all ps > .811; WS: F values between 0.01 and 0.08, all ps > .776). Hence, the data do not provide evidence for an effect of CA over fixation duration.

Figure 2.14 shows for each clinical group the relationship between CA and the median fixation duration averaged across the 12 experimental trials. Visual inspection of the charts shows a high level of variability of fixation duration across development for both groups, and smaller variations of median fixation duration in older participants.

Figure 2.14. Association between mean median fixation duration and CA by group.



Note: The vertical black line at x = 216 represents the criterion used to discriminate children from adults, that is CA = 18 years. The horizontal blue line is positioned at the mean value of median fixation duration for each group.

2.4 Discussion

In this study ET methodology was employed to investigate enumeration processes in children and adults with WS and DS, compared to a TD group matched for performance on RCPM. Participants were asked to enumerate visual sets with 1 to 6 dots arranged either in a dice pattern or in a random pattern. This allowed the investigation of the overall performance on the enumeration task, the processes used in different experimental conditions, and the participants' eye movements while performing the task.

2.4.1 Performance on the enumeration task

Overall, high accuracy rates were reported, in line with previous studies investigating enumeration skills in DS (Zimpel & Rieckmann, 2022) and in WS (Ansari et al., 2007; O'Hearn et al., 2011). Performance on the enumeration task of both clinical groups was in line with MA measured through RCPM for all the experimental conditions.

Analyses of RTs showed the typical discontinuity between subitizing and counting for all groups, in line with previous studies in TD population (Schleifer & Landerl, 2011; Watson et al., 2007). Although there were no statistically significant group differences for RTs in all experimental conditions, both clinical groups presented larger SDs than the control group, especially in the R6 display. This indicates a large variability in the clinical groups for RTs in this experimental display, with some individuals with DS and WS taking longer to correctly perform the enumeration task and some individuals showing RTs closer to the ones observed for the TD group.

Analyses of RT slopes showed that participants with DS and participants with WS used the same enumeration processes used by the TD control group, in all experimental conditions. This means that individuals with WS and individuals with DS in this study were able to perform both perceptual subitizing and conceptual subitizing, with subitizing ranging from 1 to 3. This result is in line with the findings from O'Hearn et al. (2011) that reported a limited subitizing range for children and adults with WS. Moreover, these findings are in line with the results reported by Zimpel and Rieckmann (2022), but contrast with the conclusions of the study by Sella et al. (2013), where participants with DS showed a pattern consistent with the use of a counting, even for low numerosities. This discrepancy could be caused by the different experimental design used by the authors in that study, where participants were asked to compare a target image with a sample image. In fact, the task set-up by Sella et al. (2013) did not only measure enumeration skills but also relied heavily on the participants' working memory abilities because they were asked to compare sequential stimuli.

Since these results suggest that both clinical groups performed conceptual subitizing, an implication of this study for practice would be to leverage this process in educational training programmes as an alternative to counting to support enumeration abilities and overall mathematical development. In fact, the use of conceptual subitizing can support the understanding of other key mathematical concepts and strategies, such as the understanding of number composition, set combination and the emergence of groupitizing (Starkey & McCandliss, 2014).

Finally, examination correlation analyses showed for all groups that accuracy in conceptual subitizing was correlated to developmental level measured by RCPM and calculation skills measured through he WIAT-NO, while accuracy on counting was only correlated to calculation skills. These results not fully in line with findings by Sella et al. (2013) which investigated enumeration skills in DS and found a moderate positive correlation between accuracy in counting and RCPM scores, but did not found a significant correlation between accuracy in counting and mathematical abilities measured through the Numerical Intelligence Scale. This contrasting finding might be explained by the use of the different tools to assess enumeration skills. Analyses of patterns of correlations showed that while individuals with DS showed a pattern of correlation similar to the TD group this was not the case for the WS group, which, contrary to the other two groups, showed significant correlations of CA for accuracy in conceptual subitizing but not in counting. These findings should be further investigated through the use of regression analysis to evaluate the patterns of predictors for WS and DS, and specially the role of CA.

2.4.2 Eye gaze behaviour

The analyses of eye movements indicated that fixation count and median fixation duration were significantly different between the three groups. In fact, this study found that individuals with DS presented overall gazing behaviour characterised by significantly shorter median fixation duration than the TD and the WS groups in all the experimental conditions. This finding is in line with the study predictions and with previous research by Viñuela-Navarro et al. (2017). Moreover, the DS group showed significantly higher fixation count than both the WS and the TD group for specific experimental stimuli. The fixation instability observed in the DS group did not have an impact on their performance in the enumeration task.

Moreover, the results showed that individuals with WS presented shorter median fixations than the TD group in the conditions D1-D3 and R4-R6. This is contrary to the study predictions, which expected to observe differences between the TD control and the WS group for the fixation count, but not for fixation duration. This finding can be explained by the age range of the sample included in this study. In fact, "sticky fixations" have been reported in studies with toddlers and children younger than 5 years old (see Van Herwegen, 2105 for a review). Given that the youngest participant with WS in this sample was 8 years old, this may explain the absence of a "sticky fixation" pattern in the current study. As observed for the DS group, the fixation instability observed in the WS group did not have an impact on their overall performance on the enumeration task. According to Just and Carpenter (1976), the shorter fixation durations might be interpreted as individuals with DS and WS having "less of a hard time" to process the visual information than the TD group or as individuals with DS and WS being "less engaged" than the TD group.

The analysis of the eye movements suggested that fixation count, but not mean fixation duration, appeared to be a reliable indicator of the processes employed during the enumeration task for both clinical groups. These results are in line with previous studies in TD populations (Schleifer & Landerl, 2011; Watson et al., 2007) and in individuals with DD or MD (Moeller et al., 2009; Schleifer & Landerl, 2011). In fact, no significant correlations between median fixation duration and RTs were found for almost all the experimental displays, showing that, on average, for all groups neither the number nor the spatial arrangement of the dots presented influenced the visual processing time. This finding seems to indicate that the cognitive load required for subitizing and counting processes was similar for all groups. The only exceptions were found for the WS group, where a moderate positive correlation between median fixation duration and RTs were found for the R2 and D5 displays. The reasons behind these findings are not clearly understood at this point. On the other hand, for all groups a positive correlation between RTs and fixation count was found for almost all the experimental displays, meaning that enumeration processes were well described by fixation count. The only exceptions were observed when participants were presented with either 1 or 2 dots, where no significant correlations between RTs and fixation count was found. This can be explained by the low variability observed in RTs for these experimental displays. Finally, analyses of the patterns of significance of correlations between RTs and fixation duration for each experimental condition showed that individuals with WS reported some differences in comparison with the TD group. These findings are in line with the ones reported in previous studies which investigated eye movements in DS and in WS while performing a non-symbolic magnitude comparison task (Van Herwegen et al., 2019) and a number line task (Simms et al., 2020).

The findings showing the tight coupling between RTs and fixation count were supported by the observation for all groups of the typical performance of the enumeration task for both RTs and fixation count, which is characterised by a flat line when subitizing, and a sharp increase when individuals start engaging in counting. Moreover, these findings were supported by the analyses of fixation count slopes, which replicated the pattern observed for the RT slopes, with few discrepancies. When comparing fixation count slopes in D1-D3 condition with the ones in D4-D6 condition for the DS group, a statistically significant difference was observed. This was not reflected in the corresponding RT slopes. Also, when comparing the fixation count slopes in D1-D3 condition with the ones in R1-R3 condition for the WS group, a statistically significant difference was observed. Again, this was not reflected in the corresponding RT slopes. These findings might indicate that ET measures are more sensitive than RTs. Hence, these findings support the methodological choice of combining RTs and eye movements discussed by Schleifer and Landerl (2011) to obtain a deeper understanding of the processes underlying enumeration skills.

Finally, the analyses of the fixation count for different experimental displays showed results in line with previous research in TD and MD populations which demonstrated the occurrence of saccadic movements within the subitizing range (Schleifer & Landerl, 2011; Watson et al., 2007). Indeed, the findings from the current study indicate that, on average, all groups exhibited more than one fixation when they engaged in subitizing. Also, analyses of the fixation count for the displays where participants were counting showed that, on average, for all groups the average number of fixations was higher than the number of dots to be enumerated. For the TD group, this result is in line with findings from Schleifer and Landerl (2011) that showed that only older TD participants (those aged 11 years old and adults) used systematic scanning strategies characterised by lower saccadic frequency than the number of dots to be enumerated. However, a greater number of fixations than the number of dots to be enumerated was reported not only for younger participants with DS and WS, but also for the group of older participants with WS and with DS (aged 11 or older), thus suggesting that both adolescents and adults in the DS group and in the WS group were using inefficient scanning strategies.

2.4.3 Changes of performance and eye gaze behaviour across development

Findings from the investigation of cross-sectional developmental trajectories did not show any clear relationship between CA and RTs, CA and fixation count, and CA and median fixation duration. Hence, the data do not provide evidence of an effect of CA on the enumeration task performance and of CA on the eye gaze behaviour for both clinical groups. In other words, this study did not find any influence of age on how individuals with DS and WS performed the enumeration task, showing that they do not get quicker or slower at subitizing and counting with age. This might be due to the wide age range of the samples investigated. Moreover, the study did not find any influence of CA on the eye movements of individuals with DS and WS, suggesting that their eye gaze behaviour does not change across development when numerical information is processed. This is in line with previous studies in TD population showing that the distributions of fixation duration for children and adults are quite similar (Rayner, 1998).

Visual inspection of the scatterplots plotting RTs, fixation count, and fixation duration against CA, provided insightful information on the variability observed within the clinical groups, and whether variability changes with age. Visual inspection of the charts plotting RTs against CA showed lower levels of variability in RTs for the conditions where participants were using perceptual subitizing and higher ranges of variability for the conditions where participants were using conceptual subitizing and counting. Moreover, the scatterplots showed similar levels of variability in RTs between young and old participants in all the experiential conditions for both groups. Visual inspection of the charts plotting fixation count against CA indicated similar patterns of variability to the ones observed for RTs. In fact, for both groups, lower levels of variability in fixation count could be observed in the experimental conditions where participants used perceptual subitizing compared to the others. Moreover, similar levels of variability between young and old participants are observed in all conditions. Visual inspection of the charts plotting median fixation duration against CA showed high levels of variability for both clinical groups. Furthermore, the charts showed that the range of median fixation duration decreased with age. However, these results should be interpreted with caution, as the reduced variability in fixation duration for the older participants could be explained by the small number of adults included in the sample.

Finally, a series of correlations between CA and accuracy when performing conceptual subitizing and counting suggest that the level of experience influences performance differently for the two clinical groups. In fact, older individuals with WS were better than their younger counterparts at conceptual subitizing, while older individuals with DS were better than their younger counterpart at counting.

2.4.4 Limitations of the study

There are some limitations to consider when interpreting the results of this study. First, only one trial per experimental display was assessed. The number of trials of the enumeration task was kept at a minimum level to avoid increasing the length of the testing battery and to limit participants' fatigue. However, having more trials per experimental display would lead to a more robust estimation of the slopes for RT and fixation count and, in general, to more statistical power.

Second, the current study included participants from a wide age range. This was to reach a sample size in line with ET experiments in mathematics education research, that has been reported to be on average 28.56 (SD = 21.70) (Strohmaier et al., 2020). This target was particularly challenging to meet with the WS group, given the difficulties related to conducting ET studies with participants with such a rare disorder (Martens et al., 2008). The relatively small number of participants compared to the large age range may have hidden age-specific group differences. Enumeration skills should be further investigated in younger samples with narrower age ranges.

Finally, although the lack of a control group matched for CA was justified by the scope of the present study, it did limit the conclusions regarding enumeration processes in DS and in WS. In fact, based on the current study, it can only be stated that enumeration skills in DS and WS are in line with MA (RCPM). Further studies are needed to investigate whether enumeration skills are typical in these populations.

2.4.5 Implications and future directions

The implications of this study are relevant for both research and practice. On one hand, the use of ET in this study proved to be a valuable tool to investigate enumeration skills in DS and WS populations. Results showed that fixation count seems to be a more sensitive measure than RTs to describe the processes used by participants during enumeration and for delineating cross-syndrome differences. However, the interpretation of ET data is still debated and their implications still unclear (Strohmaier et al., 2020) and more knowledge around the interpretation of ET findings in the field of mathematical cognition is needed.

On the other hand, these findings might have implications for interventions and educational programmes with regards to how to support enumeration abilities in individuals with DS and WS. In fact, this study showed that children and adults with DS and with WS can perform perceptual subitizing and conceptual subitizing. Hence, learning objectives should target these basic components of mathematics. Moreover, evidence from TD populations showed that perceptual and conceptual subitizing are important factors for mathematical development, and in particular for the development of counting, calculation skills, and overall mathematics achievement (Butterworth, 2005; Clements, 1999; Kroesbergen et al., 2009; Özdem & Olkun, 2021; Reigosa-Crespo et al., 2013). Further studies should investigate if this is also the case for DS and WS populations. Furthermore, despite the differences observed between the clinical and the control groups for fixation count and median fixation duration, these do not seem to have an impact on the enumeration processes used and on the overall performance when enumerating 1 to 6 dots. Hence, these findings indicate the presence of syndrome-specific mechanisms which do not impact enumeration skills. Finally, the high variability observed in both groups when using conceptual subitizing and counting strategies highlights the importance of considering the design of individualised educational programmes to support mathematical development in these populations.

2.5 Conclusion

The performance of individuals with DS and WS in the enumeration task based on accuracy and RTs was in line with the TD control group matched for MA on RCPM. However, further studies are needed to investigate whether enumeration skills are typical and whether the predictors of subitizing in DS and WS are the same.

Findings from this study indicate that children and adults with DS and WS can perform both perceptual subitizing and conceptual subitizing and that they use inefficient scanning strategies when counting.

The analysis of the eye movements during the enumeration task showed that individuals DS and WS exhibit shorter fixations than the TD control group and that individuals with DS exhibit higher fixation count than the WS and the TD control group. However, this did not seem to affect their overall performance and the processes used during the enumeration task. The causes and implications of these shorter fixations and of higher fixation count need to be further explored. Moreover, this study confirmed that fixation count is a good indicator to identify the processes used during enumeration.

High levels of variability were observed in both clinical groups both in terms of RTs and eye movements, especially in the conditions where participants were using conceptual subitizing and counting.

Chapter 3: Maths at Home: Development and use of a parental questionnaire

This study was conducted with the aim to investigate and compare the home learning environment of primary school children diagnosed with Down syndrome (DS) and Williams syndrome (WS) through a web-based parental survey.

Aligned with the neuroconstructivist perspective which recognises a crucial role of the environment in shaping development, this study aimed to fill the gap in the current literature and to explore the home learning environment, and in particular the home mathematics environment (HME), in these two groups in terms of frequency and type of home-based activities, parental expectations, and child and parent's attitudes towards learning.

Moreover, as this study used the methodology of parental web-based survey, this has been an opportunity to collect parental perspectives on mathematical abilities and on academic performance of their child and to investigate whether such perspectives impact the type and frequency of learning activities which occur at home.

The final aim of this study was to analyse HME across syndromes and across development. The cross-syndrome comparison can help answer the question about syndrome specificity. Finally, in line with the neuroconstructivist approach, this study takes a developmental perspective and provides a developmental account of HME by exploring changes across development of frequency and type of maths-based activities used at home, parental expectations, and child's attitude towards mathematics.

The findings from this study have been published in The home learning environment of primary school children with Down syndrome and those with Williams syndrome. *Brain Sciences*, 11(6). 733. https://doi.org/10.3390/brainsci11060733; available online on 20 April 2021. This chapter presents a more detailed and expanded version of this work.

3.1 Background and rationale of the study

3.1.1 Definition of home learning environment

As discussed in Section 1.2.7, the home learning environment comprises structural indicators and functional indicators. Structural indicators refer to the physical characteristics of the resources used to facilitate learning, such as type and frequency of the learning activities that occur at home. Functional indicators relate to the quality of the implicit and explicit learning support that the child receives from their caregivers and may include parents' expectations regarding their child's academic outcomes, parents' and child's attitudes towards learning, or quality of parent-child interactions.

The home learning environment can be differentiated into home literacy environment (HLE) and home mathematics environment (HME). HLE is often defined as the frequency of literacy-related activities in the home, such as shared parent–child book reading. Additionally, measures such as the age of onset of parent–child book reading, the number of books in the home, the frequency of trips to the library, and parental attitudes such as the enjoyment of reading and beliefs about reading are also considered to be important aspects of the HLE (Payne et al., 1994). Senechal et al. (1998) suggested that children at home can be exposed to both formal literacy instruction activities, i.e., those activities where the attention is on the print itself, and informal literacy experiences, i.e., those activities where the print is present but is not the focus of the parent–child interaction. HME is defined in the literature as a multifaceted construct consisting of various components. However, there is still no consensus on the specific components that should be included, as well as how they should be measured. Maths-related activities have also been classified into formal and informal (LeFevre et al., 2009). As for the literacy-based activities, informal maths activities encompass real-world tasks in which parental teaching occurs without a predefined mathematical target. Because the informal acquisition of skills is often incidental, mathematical learning may happen as part of playful and extra-curriculum activities that involve the use of numbers or shapes. Formal maths activities are characterised by their explicit emphasis on mathematical abilities and are employed by caregivers with the specific intention of supporting mathematical skills. This may involve activities such as repeating the number sequence or making calculations. Some work has also further categorised formal maths activities on the basis of the level of difficulty (for example, basic and advanced), and of the type of mathematical skills they target (Daucourt et al., 2021).

3.1.2 Findings from studies investigating the home learning environment in typically developing populations

Findings of studies investigating HLE in typically developing (TD) populations suggest that children's exposure to books, in both formal and informal activities, is related to the development of vocabulary and listening comprehension skills and that parental involvement in teaching children about reading and writing words is related to the development of early literacy skills (Senechal & LeFevre, 2014).

Findings from studies investigating HME in TD populations show that most parents report counting objects to be the most frequent mathematics-based activity (Blevins-Knabe et al., 2000; Mutaf-Yildiz et al., 2020; Zippert & Rittle-Johnson, 2020) and that at home occur very few activities focusing on other aspects of mathematics, as parents often fail to grasp opportunities to incorporate mathematical components other than numbers into their daily activities (Cheung & McBride, 2017). Furthermore, del Rio et al. (2017) found that parents who had high expectations for their child's mathematical outcomes also reported engaging more frequently in advanced number activities at home.

Regarding research comparing HLE and HME in TD populations, findings indicate that parents' engagement is lower in mathematics compared with literacy activities. For example, studies by Blevins-Knabe et al. (2000) and by LeFevre et al. (2009) showed that home maths-based activities occur less frequently than literacy-based activities. The study by Blevins-Knabe et al. (2000) reported that the frequency with which parents involved their 4- to 6-year-old child in mathematics-based activities was positively related to the frequency of their engagement in literacy-based activities. Moreover, this study demonstrated that the frequency of parental engagement in mathematical activities with their child was related with the parents' personal attitudes towards mathematics. Specifically, parents who enjoyed mathematics offered mathematical activities more frequently (Blevins-

Knabe et al., 2000). Finally, findings in a qualitative study by Cahoon et al. (2017), it was found that reading experiences were integrated into daily routines, with parents dedicating specific time to engage with their child aged between 3.1 and 4.9 years. In contrast, experiences related to numbers were less structured and did not occur at a prescribed time.

3.1.3 Findings from studies investigating the home learning environment in Down syndrome and Williams syndrome

The results from studies investigating the HLE of individuals with DS reported that most parents place a high value on supporting their child's literacy development as they are involved in regular literacy interactions with their child (Al Otaiba et al., 2009; Lusby & Heinz, 2020), although more so for reading than for writing (van Bysterveldt et al., 2010). The HLE of children with DS described in the literature is rich (Fitzgerald et al., 1995) and positive, with the majority of children introduced to books when they are 1 year old and the majority of the families with mixed socioeconomic status (SES) reporting having more than 50 children's books at home (Al Otaiba et al., 2009; van Bysterveldt et al., 2010) as well as a wide range of writing materials (Trenholm & Mirenda, 2006; van Bysterveldt et al., 2010). Parents of children with DS aged from 3 months to 6 years reported reading to or with their child on a daily basis and using reading instructional materials, such as flash cards or magnetic letters, on a daily basis (Al Otaiba et al., 2009). Parents also reported additional ways in which they facilitated the development of literacy skills of their child with DS, including active teaching, language games, exposure through TV programmes and other electronic media, and library visits (Ricci & Osipova, 2012; van Bysterveldt et al., 2010). These studies also reported that parents of primary school children with DS reported their child to have a positive interest in reading, regardless of their age (Lusby & Heinz, 2020). Furthermore, Al Otaiba et al. (2009) investigated the lifelong literacy goals that parents had for their preschool child with DS and found that developing their child's literacy skills was a high priority target for parents who reported a wide range of reading goals such as recognising the alphabet, reading for meaning, reading for pleasure, reading chapter books, and reading for job purposes. These findings were confirmed by the study by Ricci and Osipova (2012), where parents of children with DS aged between 3 and 13 years described reading as a key goal for their child. Finally, Trenholm and Mirenda (2006) investigated the home and community literacy experiences of individuals with DS and collected data from 224 parents and guardians across Canada. Their findings showed that more than 50% of the respondents reported that their family member used a computer for reading activities and / or for writing activities and that adolescents and adults engaged with technology more often than the younger groups and in ways which were more functional than school-based activities.

The HLE of individuals with WS was investigated by a recent exploratory study by Lettington and Van Herwegen (2024). This study described and compared HLE of children, adolescents, and adults with DS and WS (CA range = 3.58 – 36.33 years) and reported similar features of the HLE for the two groups, with the only difference being that individuals with DS engaged less than individuals with WS in informal literacy activities. Moreover, the study found that the support that individuals with DS and WS received at home (measured as the frequency of literacy-based formal and informal activities) remained stable across the different age groups.

While the HLE of individuals with DS and WS has been investigated, to the best of my knowledge, there is no published peer-reviewed research investigating the HME of these populations. Although the studies that examined the HLE of these population suggest that parents offer a rich HLE to their child, it is unclear what resources and activities parents use at home to support their child's mathematical development, what is the frequency of home-based maths activities, what are the parental attitudes towards maths and the parental expectations towards their child's mathematical development. Moreover, it is unclear whether the structural and functional indicators of the HME differ between groups and whether they change across development. This study addresses this gap and uses a parental questionnaire to investigate and compare the HLE and HME of primary school children with DS and WS.

3.1.4 Current study

This study developed an online-based parental questionnaire to investigate and compare the home learning environment of primary school children with DS and with WS.

The first aim of this study was to explore and compare the HLE and the HME of children with DS and WS. Based on the existing literature on the HLE of individuals with DS and WS, and on the studies investigating HME in TD populations, it was predicted that:

- Similar to TD population, literacy-based activities would be more frequent than mathematics-based activities both in DS and in WS.
- Similar to TD population, within the mathematics-based activities, counting would be the most frequent activity.
- In line with parents' expectations towards reading in DS population, parents' expectations towards their child's academic outcomes in literacy and maths would be in line with the targets set by the English national curriculum (Department for Education, 2014)
- Similar to TD population, the frequency of mathematics-based activities would be positively correlated to the parent's attitude towards mathematics and to the frequency of their engagement with literacy-based activities.
- No predictions were made on the child's level of interest and motivation towards mathematics-based activities, due to lack of previous studies.
- When technology was used at home to support learning, it was predicted that the pattern observed in the type and frequency learning activities occurring at home would be the same as the one observed when no technology was involved.

The second aim of this study was to investigate the academic performance and the general level of functioning of children diagnosed with DS and WS, as reported by their parents. Moreover, the study aimed at investigating the association between the home learning environment and the child's general level of functioning. Based on the literature on DS and WS populations discussed in Chapter 1 and on studies investigating HLE in TD populations, it was predicted that:

- When rating their child's general level of functioning, carers would report different profiles, with carers of children with DS reporting lower expressive skills scores than the ones reported by carers of children with WS.

- Caregivers from both groups would acknowledge difficulties in their child's mathematical skills.
- Differences in the general functioning profile and in the mathematical academic profile reported by the carers would be reflected in differences in the HME in terms of type and frequency of learning activities occurring at home.

Finally, the third aim of this study was to investigate the effect of the chronological age (CA) of the child on the type and frequency of maths-based activities which occur at home of primary school children with DS and WS. Due to the scarcity of studies investigating home learning environment developmentally in these populations, no predictions were made. These predictions do not take into consideration the results of the study by Lettington and Van Herwegen (2024), as that study was published 2 years after the current study. For completeness, findings from Lettington and Van Herwegen (2024) are considered in the final discussion.

The ultimate objective of this study was to gather insightful information from parents that could inform the development of parental interventions employing learning resources already present in the home learning environment and the development of guidelines to improve current practices.

3.2 Participants and Methods

3.2.1 Respondents

Sixty-four carers of primary school children with DS (*n* = 39) and WS (*n* = 25) completed the web-based survey. Five participants were excluded from the final sample as they did not provide complete demographic data of their child (*n* = 2 respondents reported a wrong date of birth, *n* = 2 respondents did not provide date of birth and *n* = 1 respondent did not provide the sex of the child). Hence, data from 59 respondents (DS: *n* = 35, WS: *n* = 24) were included in the final sample.

Table 3.1 provides sociodemographic characteristics of the included respondents. Fifty-eight respondents (98%) were females, one respondent (2%) was male. Overall, most of the families were well educated (75% were educated at the university level), white (90%), and based in the UK (96%). The lack of variability in the sociodemographic characteristics of the sample could be due to the sampling techniques, to the recruiting method, and to the online format of the survey.

Table 3.1. Sociodemographic characteristics of respondents.

Characteristic	Carer of child with DS		Carer of child with WS		Tot	
	n	%	*n*	%	*n*	%
Sex						
Female	35	100	23	96	58	98
Male	-	-	1	4	1	2
Ethnicity						
White	30	86	23	96	53	90
Asiatic	2	6	1	4	3	5
Other	3	8	-	-	3	5
Highest level of education						
Post-graduate degree or equivalent	15	43	8	33	23	39
University degree or equivalent	10	29	11	46	21	36
School leaving certificate (e.g., GCSE or A levels or equivalent)	7	20	5	21	12	20
Vocational training	2	6	-	-	2	3
No formal qualification	-	-	-	-	-	
Missing	1	2	-	-	1	2
Highest level of mathematical education						
Post-graduate or university level	1	3	-	-	1	2
School leaving certificate level (e.g., GCSE or A levels or equivalent)	27	77	17	71	44	75
Missing	7	20	7	29	14	23
Country						
UK	35	100	22	92	57	96
Ireland	-	-	1	4	1	2
New Zealand	-	-	1	4	1	2

3.2.2 Drop-out rates

As reported in Table 3.2, a total of 177 respondents used the anonymous link and accessed the questionnaire, read the information sheet, and completed the consent form. However, only 60% of them completed Section 1 about their child's demographic and academic abilities, showing a drop-out rate of 40%. The drop-out rate further increased as respondents progressed through the questionnaire, with 64% of the respondents that accessed the questionnaire not completing the questionnaire in its entirety (drop-out rate A). Drop-out rate B is calculated over the number of participants that completed the first section of the survey (n = 107), rather than over the number of participants who gave initial consent to be part of the study. Drop-out rate B remains high but does not exceed 40%. Consideration of the causes of the drop-out rates and possible improvements to the web-survey are addressed in the discussion.

Table 3.2. Drop-out rates.

Section	Completed section (*N*)	Drop-out rate A	Drop-out rate B
Section 0: Information sheet and consent form	177	-	
Section 1: Children's demographic data and academic profile	107	40%	-
Section 2: Frequency of home learning activities	88	50%	17%
Section 3: Use of technology	84	53%	21%
Section 4: Parents' expectations	67	62%	37%
Section 5: Importance of competences	67	62%	37%
Section 6: Children's and parents' attitudes towards numeracy and literacy	66	63%	38%
Section 7: Vineland Adaptive Behaviour Scale II	64	64%	40%
Section 8: Respondents' demographic data	64	64%	40%

Note: Drop-out rate A is calculated as the number of uncompleted surveys divided by the total sample size (n = 177). Drop-out rate B is calculated as the number of uncompleted surveys divided by the number of respondents that completed Section 1 (n = 107).

3.2.3 Measures

3.2.3.1 The construction of the "Maths at home" questionnaire

A new questionnaire was developed for this study. The survey was created using Qualtrics. The new questionnaire was developed starting from existing surveys (Table 3.3). The "Maths at home" questionnaire was based on two parental surveys used in previous studies with TD children. One survey was used in the study by LeFevre et al. (2009) which investigated the home learning environment of preschool and primary school TD Canadian children. The other survey was used in the study by De Keyser et al. (2020), which examined the HME of pre-school TD Belgian children. Finally, some items were drawn from the study by Costa et al. (2023) that surveyed 83 early years practitioners based in the UK and explored their beliefs and practices concerning mathematics. The academic targets included in the questionnaire and the level of competences reported were based on the English national curriculum (Department for Education, 2014).

The survey's design and the formatting were crafted to prevent survey fatigue, to alleviate respondent burden and frustration, and to ensure clarity.

Branch logic and display logic features were built into the survey design to shorten it where possible. A progress bar that gradually filled as the respondent completed the survey was included at the top of each page of the survey. Questions were phrased in a simple, clear, and concise language, and spelling was checked throughout the development of the tool.

Once the draft of the survey was completed, pre-test and pilot assessments were conducted to improve the quality and the efficiency of the survey. The questionnaire was pre-tested with 5 individuals, including experts of the field, colleagues, and friends. Revisions to the questionnaire included changes aimed at standardising the Likert scales – e.g., the use of the same starting point, and the display of both verbal and numerical labels on the scale. Some instructions were rephrased to avoid the use of negative wording and key terms were highlighted either in bold or in italic font to improve clarity. Then the survey was piloted with 2 parents, one mother of a 5-year-old child with DS and one mother of a 10-year-old child with WS. Parents were asked to provide their feedback about their experience with completing the survey, especially in terms of clarity of instructions, length of the questionnaire and relevance of the questions asked. The revisions included the rephrasing of 4 questions. One of the amendments made was based on the acknowledgment that children with special educational needs may use assistive technology to write. Hence, respondents were asked to consider their child's typing skills when commenting on their child's writing skills. The pilot data was not included in the final data sample.

The final questionnaire included 66 questions, both close-ended questions (CEQ) and open-ended questions (OEQ). CEQs used frequency-based scales to measure the frequency of home-based activities, rating scales to rate socioemotional factors, and checklists to ask respondents to indicate a list of educational apps their child was using. OEQs were used to collect participants' views of their child academic abilities and one OEQ was included at the end of each section to collect additional comments from participants. As shown in Table 3.3, the questionnaire was organised in 8 sections, each one clustering questions according to a specific topic. Each section was named with a self-explanatory title and was introduced by a short paragraph that described the main topic and which included section-specific instructions, when applicable. A copy of the questionnaire can be found in Appendix A.

Table 3.3. Structure of the "Maths at Home" questionnaire.

#	Topic	Items	Source
0	Information sheet and consent form		
1	Children's demographic data and academic profile	5 CEQ; 10 OEQ	
2	Frequency of home learning activities	6 CEQ; 1 OEQ	De Keyser et al. (2020) LeFevre et al. (2009) Costa et al. (2023)
3	Use of technology	5 CEQ; 3 OEQ	
4	Parents' expectations	7 CEQ; 1 OEQ	Costa et al. (2023)
5	Importance of competences	7 CEQ; 1 OEQ	De Keyser et al. (2020) Costa et al. (2023)
6	Children's and parents' attitudes towards numeracy and literacy	2 CEQ; 1 OEQ	De Keyser et al. (2020)
7	Vineland adaptive behaviour scale 2	9 CEQ; 1 OEQ	Sparrow et al. (2005)
8	Respondents' demographic data	6 CEQ; 1 OEQ	

Note: CEQ: close-ended question. OEQ: open-ended question.

Section 0. Information sheet and consent form. The survey started with an information sheet that introduced the researcher and the research project to participants. This was followed by the consent form, which was compulsory.

Section 1. Children's demographic data and academic performance. The first section of the survey asked about children's age, sex, clinical diagnosis, language spoken at home, school setting, and type of additional support they had received in the last year. Additionally, the questions included in this section collected qualitative data to investigate the child's academic performance. Respondents were asked to report at what level of the English national curriculum their child was working at in mathematics, reading and writing. As the design of this study did not include a control group, a decision was made to measure children's mathematical, reading and writing skills towards the benchmarks reported in the English national curriculum (Department for Education, 2014). If the child was working at the level of the English national curriculum, respondents were asked to specify whether their child was working towards the expected standard, at the expected standard or at greater depth within the expected standard. If the child was working below the standards of the English national curriculum, the respondents were asked to specify whether their child was following the P scales[13]. Finally, for each academic area, parents were asked to report whether their child's academic abilities were better, in line, or worse than their overall abilities.

Section 2. Frequency and type of home learning activities. In the second section, parents were presented with a list of 36 home activities (Table 3.4) and were asked to rate on a 5-point Likert scale, ranging from 0 (never) to 4 (every day), how often they engaged in these activities with their child. Parents could opt out and select the option "not age appropriate" if they considered their child was too old or too young for the item presented. On the basis of previous studies, a list of home-based activities and resources were identified and classified into six areas: 1) literacy-based activities (*n* = 6), 2) everyday life activities (*n* = 6), 3) activities related to domain-general abilities which support mathematical development (e.g., visuospatial abilities) (*n* = 6), and mathematics-based activities broken down into 4) number skills activities (*n* = 6), 5) calculation skills activities (*n* = 6), and 6) broader mathematical skills activities (*n* = 6). Literacy-based activities were included in the list to compare their frequency with mathematics-based activities and to determine whether the same pattern of occurrence observed in the TD population was observed in DS and in WS. Everyday activities were included in the list as control items. Number skills activities included learning activities targeting number recognition and counting. Calculation skills included activities which involved the use of arithmetic operations. Broader mathematical skills included activities targeting functional mathematics and geometry. To measure the frequency of home-based learning activities, the Activity Frequency (AF) score was computed for each participant as the average score of the 6 items presented in each category. In the case of items reported as "not appropriate", these were excluded from the computation. Where more than 3 items were reported as "not appropriate" in the same category, the category was coded as "not appropriate". The AF score ranged from 0 to 4 for each category, with higher scores indicating more frequent home-based activities.

[13] In the UK, P scales were used to assess the progress of children aged 5-14 who had special educational needs and whose abilities did not reach Level 1 of the English national curriculum. P Scales have been replaced by Pre-key stage standards in 2021.

Table 3.4. List of home learning activities.

Activity	Formal	Informal
Number skills	1.Using Numicon resources 2.Using number flashcards 3.Reading number story books that include numbers or counting	4.Using sticker reward charts 5.Singing number songs together (e.g., five little monkeys) 6.Counting during daily activities (e.g., counting the number of apples when cooking)
Calculation skills	1.Worksheets on addition and subtraction 2.Using number activity books 3.Doing maths homework	4.Elementary calculations during daily activities (e.g., "There are five apples in the fruit bowl. If I take one, how many apples are left?") 5.Playing board games that require elementary computations (e.g., with two dice) 6.Recognising and finding half of a quantity, length, set of objects or shape (e.g., can I have half of your sweets?)
Broader mathematical skills	1.Handling and naming common 2-D or 3-D shapes 2.Playing dominoes 3.Using measuring tools such as a ruler when drawing or a scale when cooking 4.Playing estimation games (e.g., guess which one is more?) 5.Talking about money when shopping 6.Telling the time	
Literacy skills	1.Writing letters and/or words (e.g., writing birthday cards) 2.Writing/typing your child's name 3.Reading books	4.Paying attention to letters and/or words during daily activities (e.g., cooking) 5.Playing games that include writing and/or reading (e.g., Fishbowl game) 6.Learning new words during daily activities
Domain-general abilities that support maths	1.Drawing 2.Playing memory games (e.g., Shopping List) 3.Playing jigsaw puzzles 4.Doing connect the dots activities 5.Creating patterns with concrete materials (e.g., creating a necklace alternating red and blue beads) 6.Playing with building blocks such as Lego	
Everyday life	1.Playing sports 2.Doing shopping 3.Watching TV 4.Listening to music 5.Playing with toys/video games together 6.Cooking together	

Section 3. Use of technology. The third section of the survey asked parents how often their child used tablets, computers, and TV. Respondents were also asked to report if they had any concern about their child's use of technology. Finally, respondents were asked to indicate a list of 3 educational apps which their child was using and a list of 3 educational apps which they liked. While many studies which explore the use of technology in NDCs discuss the challenges

associated with its usage or its application in therapy settings (Feng et al., 2010), very little is known regarding its use at home to facilitate educational activities. This section was included in the survey to fill this gap.

Section 4. Parents' expectations. Parents were asked to indicate on a 11-point Likert scale from 0 (not at all) to 10 (very well) how well they expected that their child would master specific competencies at the end of primary school. The 56 competencies listed in this section replicated the classification used for the home-based activities in the Section 2. In addition, to validate participants' responses, eight control items were included (control category). These included mathematical competences derived by the English national curriculum for KS2 (e.g., "converting between miles and kilometres") and everyday life skills characterised by high levels of independence (e.g., using public transports independently). To measure parents' expectations, the average score of the eight items presented in each category was computed for each participant, i.e., the expectation score (ES). The ES ranged from 0 to 10, with higher scores indicating higher parental expectations.

Section 5. Importance of competences. The fifth section of the survey asked respondents to indicate on a 11-point Likert scale from 0 (not important at all) to 10 (very important) how important it was for them that their child mastered specific competencies at the end of the primary school. The competences presented reproduced the 56-item list presented in Section 4. To measure levels of perceived importance for the different competences, the average score of the eight items presented in each category was computed for each participant, i.e., the Importance Score (IS). The IS ranged from 0 to 10, with higher scores indicating higher levels of perceived importance from the respondents.

Section 6. Children's and parents' attitudes towards numeracy and literacy. In this section, respondents were asked to rate on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) the extent to which they agreed with six statements about their own attitudes towards mathematics (e.g., "I like mathematics") and literacy (e.g., "Reading is important"), as well as five statements about their child's attitudes towards mathematics (e.g., "My child enjoys mathematics"). When rating their child's attitudes, respondents could select the option "not appropriate" if they considered that the statement was not appropriate for the stage of development of their child. For each participant, the parental attitude (PA) score and the child's attitude (ChA) score were calculated as the average score of the items presented. Both PA and ChA ranged from 1 to 5, with a higher score indicating more positive attitudes.

Section 7. Vineland adaptive behaviour scale. This section of the survey included items from five subdomains of the Vineland adaptive behaviour scale, second edition (VABS 2; Sparrow et al., 2005): receptive (13 items), expressive (15 items), written (13 items), living in the community (22 items), and fine motor skills (26 items). VABS 2 is a standardised assessment which measures the general level of functioning and adaptive behaviour in individuals from birth to 18 years. It has been used in previous studies with children with DS (Ricci, 2011) and with children with WS (Mervis & John, 2010). Respondents were asked to score their child's behaviour on a 3-points Likert scale based on the frequency of the behaviour and on the level of support provided to the child. The scale ranged from 0 (child never performs the behaviour or never performs it independently) to 2 (child usually performs the behaviours independently). Participants could also select the option "I don't know" in case they had no knowledge about their child's performance on the behaviour described. For each participant, the total raw score separately for each subdomain was computed, with higher scores indicating more adaptive behaviours.

Section 8. Participants' demographic data. The final section of the survey asked participants to provide their demographic information. The questions included participants' ethnicity, country, and highest level of education and of mathematical education completed. Parental education was chosen as a proxy measure for SES because it is predictive of variables such as income and occupation (Calvo & Bialystok, 2014).

3.2.4 Procedure

Participants completed the online survey between December 2018 and May 2019 through an anonymous link that was made available through Qualtrics, an online survey platform.

The final sample was recruited through non-random sampling methods, which included convenience and snowball sampling (Ruel et al., 2016). The online survey was open to all parents of children diagnosed with DS and WS aged between 4 and 11 years. The survey was advertised via several channels, including social media platforms, charities and foundations, support groups, as well as families within the researcher's network. Charities and support groups were used for participants' recruitment to increase trust in the study and participants' cooperation.

Participants could stop completing the survey at any point and they could resume from where they left off within a 2-week window. In case respondents failed to resume the survey within the 2 weeks deadline, their contribution was regarded as incomplete. Respondents could omit answering any question except for the questions included in Section 0 and the question asking about the child's CA. On average it took respondents 25.6 hours to complete the survey including breaks. To prevent participants from taking the survey more than once, the security option "prevent multiple submission" was selected on Qualtrics[14].

All participants were informed about the content and scope of the study and gave written informed consent before starting the online survey. This project was reviewed according to procedures specified by Kingston University London and was allowed to proceed (approval n. 1718CHA12). No incentives were offered to respondents for their participation in the study.

3.2.5 Data analysis

3.2.5.1 Quantitative analyses

Participants' responses were collected and stored through the Qualtrics survey platform. Once data collection was completed, the dataset was exported as

[14] This feature operates by placing a cookie on the respondent's browser upon submission of their response.

.csv file. After data cleaning processes, the dataset was imported into SPSS where exploratory and statistical analyses were conducted.

Data cleaning processes included re-coding of items, analyses of drop-out rates, and analyses of missing data.

Exploratory analyses were conducted to check the normality assumption of the data distribution of the study variables for the two groups. Parametric tests were preferred to their non-parametric version. The decision of conducting non-parametric analyses was taken in the instances of violation of the assumption of normality and considering the small sample size of the WS group (n = 24). Exploratory analyses also included the checks of the three assumptions needed to conduct Chi-square tests for association[15]. In the instances where the cells of the crosstabulation table showed expected counts lower than 5 and the crosstabulation was bigger than 2x2 contingency table, the categories of one of the two variables were collapsed gradually and until the crosstabulation table showed values greater than 5 for all expected counts. If the gradual process of collapsing categories resulted in a 2x2 contingency table, Fisher's exact test was used instead of the Chi-square tests to assess association between the variables investigated.

Preliminary analysis highlighted two main issues. The first issue was related to the distributions of the variable measuring the importance perceived by the respondents (IS), which were highly skewed towards the higher end of the scale for both groups and for all the categories, meaning that parents rated as very important all the competences listed in the survey. Since the same pattern was observed for the distribution of the control category, it was deemed that the quality of the data collected with this scale was not satisfactory. Consequently, this variable was removed from the dataset. The second issue was related to the lack of variability observed for the ordinal variable SES. This resulted in the lack of monotonic relationship between SES and any of the other variables, that is an assumption of Spearman's correlation. Hence correlational analyses involving the SES variable were not conducted.

---

[15] Two assumptions are related to the study design and include checks on the level of measurement of the variables (which must be categorical) and the independence of the observations. The third assumption is related to whether the data fitted the chi-square test model, in that all cells of the crosstabulation table must show expected counts greater than 5.

To answer RQ1, a series of Mann–Whitney *U* tests were conducted to examine differences in structural and functional indicators of the home learning environment between DS and WS. Moreover, a series of Friedman tests was run to determine if there were differences between the frequency of home-based activities and between different categories of parental expectations. Pairwise comparisons were performed running separate Wilcoxon signed-rank tests with a Bonferroni correction for multiple comparisons.

In relation to RQ2, the academic performance of the children was investigated through single-item analyses of the VABS-II subscale "community" and through a series of Chi-square tests to examine the significance of the association between child diagnosis and the level of challenges reported by parents and between child diagnosis and the target level their child was working at. Demographic variables of the two groups were compared using both t-tests and Mann–Whitney U tests. Fisher's exact test was run to examine the significance of the association between child diagnosis and the type of school setting. Finally, Spearman's correlations were run to determine whether there was an association between frequency of home-based activities, parental expectations, parental attitude towards learning, child's attitudes towards mathematics, and CA and general functioning profile of the child.

To answer RQ3, either Kruskal-Wallis *H* or one-way ANOVA tests were conducted to understand whether frequency of activities, parental expectations, and child's attitudes towards mathematics changed based on the year group that the child was attending at school.

In instances where statistically significant differences between DS and WS groups were not reported, the above analyses that were planned at group level were collapsed by clinical group to increase the power of the statistical analyses conducted. When this approach was taken, this is reported in the relative result section.

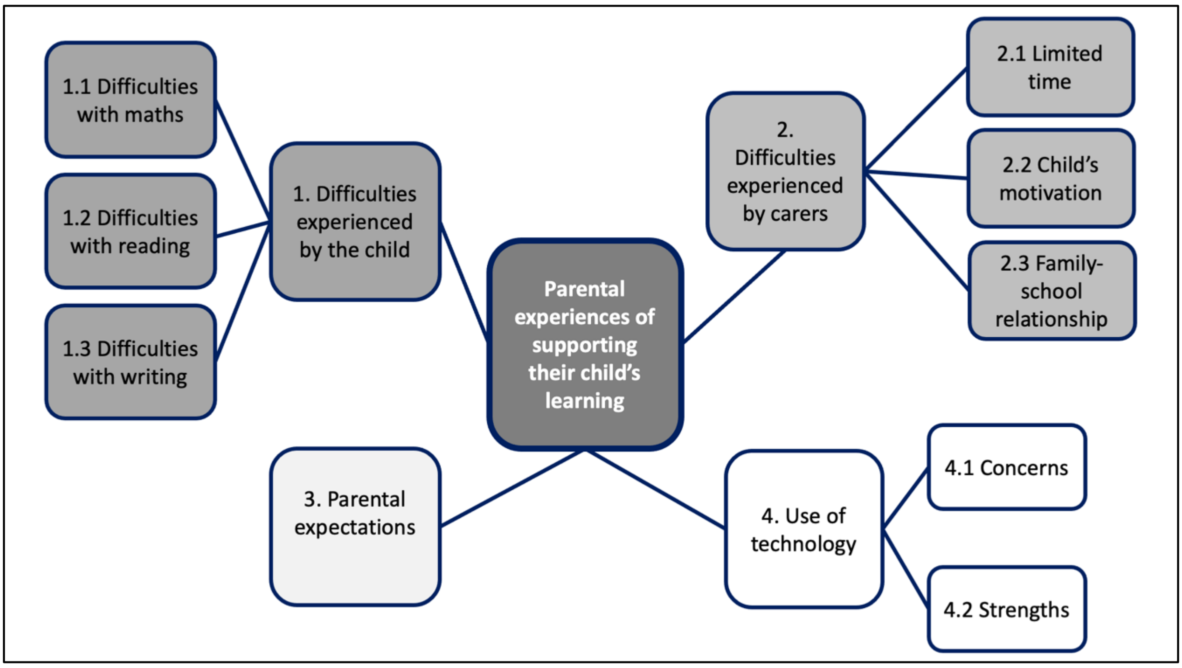
3.2.5.2 Qualitative analysis

A reflexive thematic analysis approach was employed to examine the qualitative data obtained from the respondents' comments. This methodology was chosen to identify, analyse, and report patterns (themes) within the qualitative data collected. According to Braun and Clarke (2019)'s guidelines, a six-step process was executed iteratively. First, I immersed myself in and became familiar with the collected data. Then, I began generating codes, which were subsequently grouped into initial themes. These initial themes underwent a review, leading to the refinement and naming of the final themes before the writing up of the analysis. The reflexive thematic analysis was conducted using NVivo.

3.3 Results

Figure 3.1 shows the main themes and the sub-themes identified through reflexive thematic analysis. The four main themes include: 1) difficulties experienced by the child, 2) difficulties experienced by the carer, 3) parental expectations, and 4) use of technology. The findings from the reflexive thematic analysis are discussed in conjunction with the results from the quantitative analyses.

Figure 3.1. Thematic map of participants' experience in supporting their child's learning at home.

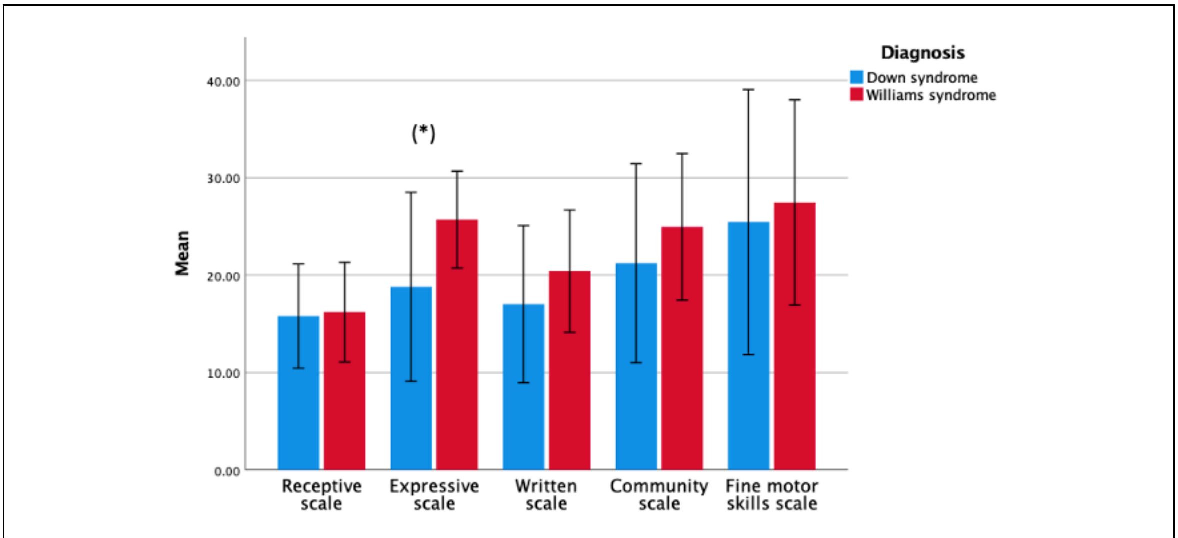


3.3.1 Description of the sample

The final sample included 35 children diagnosed with DS (16 female) and 24 children diagnosed with WS (9 female). One child had a double diagnosis of WS and autism, one child with DS had an additional diagnosis of developmental dyscalculia, and one child with DS had an additional diagnosis of cortical visual impairment. English was the primary language for 96% of the children, with one child speaking Romanian at home and one family not providing this information.

There were no significant differences in age between the DS group and the WS group, *U* = 451.5; *p* = .627, with children across the groups aged between 4;01 and 11;07 years. Figure 3.2 shows that the general level of functioning and adaptive behaviour of children with DS and with WS, was similar and characterised by high levels of variability for both groups. There was a significant difference between the two groups only for one scale of the VABS 2, the expressive scale, *U* = 585.0; *p* = .011, with the DS group (*Mdn* = 21.00, *IQR* = 1.00 – 30.00) reporting lower scores than the WS group (*M* = 28.00, *IQR* = 9.00 – 30.00).

Figure 3.2. Scores on the VABS scales by clinical group.



Note: *n* DS = 35. *n* WS = 24. Error bars show standard deviation.
*p < .05.

Table 3.5 shows the number of children with DS and with WS in different year groups, categorised according to the UK education system classification described in Section 1.3.4. Children were distributed unequally within the different year groups, with most of the children attending Key Stage 2 (KS2) (DS = 16; WS = 10), and a lower number of children attending early years (EY) (DS = 9; WS = 4).

Table 3.5. Distribution of children by year group and by clinical group.

Year group	DS	WS	Total
EY			
Early years (3-4 years)	2	2	4
Reception (4-5 years)	7	2	9
KS1			
Y1 (5-6 years)	7	4	11
Y2 (6-7 years)	3	6	9
KS2			
Y3 (7-8 years)	2	1	3
Y4 (8-9 years)	4	2	6
Y5 (9-10 years)	4	4	8
Y6 (10-11 years)	6	3	9

Note: EY: Early years. KS1: Key Stage 1. KS2: Key Stage 2.

As shown in Table 3.6, for both groups most children were attending a mainstream school (DS: 77% and WS: 71%). There was no statistically significant association between child diagnosis and the type of school attended (mainstream vs non-mainstream[16]), as assessed by Fisher's exact test, p = .536.

Most of the children received some form of additional support in their last academic year (Table 3.6). It was not possible to run a Chi-square test to assess the association between child diagnosis and additional support, because of the lack of independence of the observations. However, most of the children in both groups, but not all of them, received speech and language therapy (DS: 94%; WS: 75%) and special educational needs support (DS: 77%; WS: 96%). Lower percentages of children received occupational therapy support (DS: 54%; WS: 38%). A higher percentage of children in the DS group (60%) received visual support than children in the WS group (25%). Some respondents reported in their comments that their child was receiving Lego therapy, play therapy, and art therapy as further additional support. Only one respondent reported receiving additional support for mathematics, and it was not the family whose child was diagnosed with developmental dyscalculia.

[16] Because the third assumption of the Chi-square test was not met, the analysis was run on a 2x2 table, where the categories "SEN school", "Mainstream school with SEN unit on site", "Dual placement" and "Home educated" were collapsed in one category labelled "Non-mainstream".

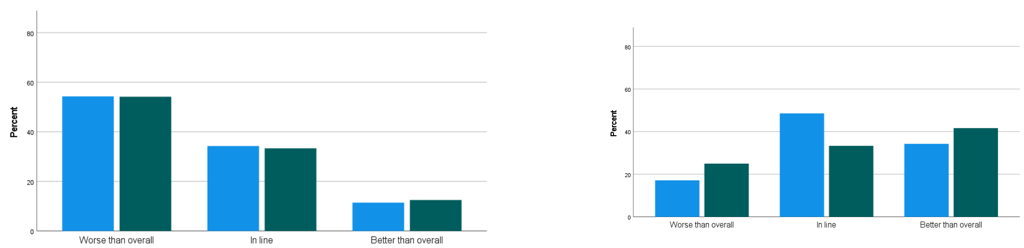
Table 3.6. Type of schooling and additional support received by clinical group.

Measure	DS		WS	
	n	%	*n*	%
Type of schooling				
Mainstream school	28	80	17	71
SEN school	3	8	4	17
Mainstream school with SEN unit on site	2	6	3	12
Dual placement	1	3	0	0
Home educated	1	3	0	0
Additional support received in the previous year [a]				
Speech and language therapy	33	94	18	75
Special educational needs support	27	77	23	96
Occupational therapy (sensory)	19	54	9	38
Visual supports	21	60	6	25
Extra reading help/phonics	21	60	15	63
Life skills teaching	7	20	8	33
Physiotherapy	4	11	9	38
Music therapy/music lessons	6	17	7	29

Note: [a] Total exceeds 100% because respondents were asked to check all options that applied. SEN: special educational needs.

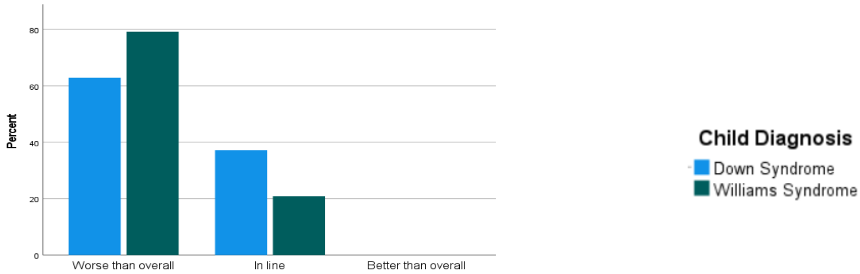
When asked to compare their child's overall abilities with their child's mathematical, reading and writing abilities, parents' ratings were similar for both groups in all the academic domains, as assessed by a series of Chi-square tests. In fact, for each academic skill, there was no statistically significant association between child diagnosis and level of challenges experienced in the three academic areas (mathematics: p = .089; reading: p = .492; writing: p = .181). Figure 3.3 shows that most of the parents reported that mathematics and writing skills were a challenge for their child, with more than half of the parents in each group reporting that their child's mathematical and writing skills were worse than their child's overall abilities. On the other hand, more than 70% of the parents in both groups reported that their child's reading abilities were either in line or better than their child's overall abilities.

Figure 3.3. Comparison of children's academic abilities with their overall abilities.



Mathematical abilities

Reading Abilities



Writing abilities

Note: *n* DS = 35. *n* WS = 24.

When explaining mathematical difficulties of their child, most of the parents mentioned difficulties with “memory” and challenges of their child “to retain maths knowledge”. For example, one parent reported that their child “used to know 3 + 3 = 6, etc., but she doesn’t recall those as much now” (P23; DS). Some parents mentioned that their child found it difficult to “apply maths knowledge” (P49; DS) and that they “struggle with abstract concepts” (P15; DS) and “to apply [mathematical learning] and make links with prior learning” (P32; DS). When commenting on the challenges associated with reading abilities of their child, most of the parents reported poor comprehension skills. Finally, for challenges related to writing, parents referred to “low muscle tone” (P4; DS) or “hypermobile joints” (P46; DS), and several parents commenting on their child writing abilities reported that their children were starting to learn to type instead of focusing on handwriting.

One of the sub-themes which was identified through thematic analysis involved the educational programmes used at home and in school. For example, one parent reported “[my child] can confidently join letters to make sounds th, ch, sh, etc., but this is not yet done with him at school. I am keeping up with my own reading support for him as I feel it is beneficial and not confusing him” (P7; WS), while another respondent reported that “the school are not using the RLI programme [Reading and Language Intervention for children with Down syndrome] that we previously used in pre-school with success, but they are having great success combining various other methods” (P36; DS).

As shown in Table 3.7, when asked to report whether their child was following the national curriculum and at what level their child was working, a similar pattern was observed for the two groups. For each academic skill, there was no statistically significant association between child diagnosis and the target level of academic performance, as assessed by a series of Chi-square tests[17] (mathematics: p = .723; reading: p = .226; writing: p = .649). For all the academic abilities investigated, most of the respondents indicated that their child was working either towards the expected standard (mathematics: 42%, reading: 51%, writing: 41%) or below the standard set by the English national curriculum (mathematics: 37%, reading: 29%, writing: 44%). A small number of respondents reported that their child was working at a greater depth within the expected standard (mathematics: 2%, reading: 3%, writing: 3%). The two children who were working at a greater depth within the expected standard in mathematics were one child diagnosed with WS aged 5 years old and one child diagnosed with WS aged 10 years old.

Table 3.7. Level of children's academic abilities.

Level of academic abilities	DS		WS		Total	
	n	%	*n*	%	*n*	%
Mathematical abilities						
Working at greater depth within the expected standard	-	-	2	8	2	3
Working at expectations	4	11	1	4	5	9
Working towards the expected standard	13	37	12	50	25	42
Working below the standard of the national curriculum	14	40	8	33	22	37
Missing / Don't know	4	11	1	4	5	9
Reading abilities						
Working at greater depth within the expected standard	-	-	3	13	3	5
Working at expectations	4	11	2	8	6	10
Working towards the expected standard	17	49	13	54	30	51
Working below the standard of the national curriculum	13	37	4	17	17	29
Missing / Don't know	1	3	2	8	3	5
Writing abilities						
Working at greater depth within the expected standard	1	3	2	8	3	5
Working at expectations	1	3	1	4	2	3
Working towards the expected standard	15	43	9	38	24	41
Working below the standard of the national curriculum	15	43	11	46	26	44
Missing / Don't know	3	9	1	4	4	7

Table 3.8 reports for each academic area the number of children that were reported to be working below the standard of the English national curriculum grouped by their year group and by their clinical group. Due to the small size of the sub-samples, it was not possible to run any statistical analyses. However, the data

[17] Because the third assumption of the Chi-square test was not met, the analysis was run for each academic skill in a 2x3 table, where the categories "Working at greater depth within the expected standard" and "Working at expectations" were collapsed in one category labelled "Working at expectations or at greater depth".

show that 44% of children with DS and 25% of children with WS in the EY group were working below the maths standard of the English national curriculum. This percentage increases to 50% and 40% for KS1, respectively for DS and WS. The academic area of reading reported the lowest number of children working below the standard of the English national curriculum, while writing reported percentages closer to 50% for all year groups.

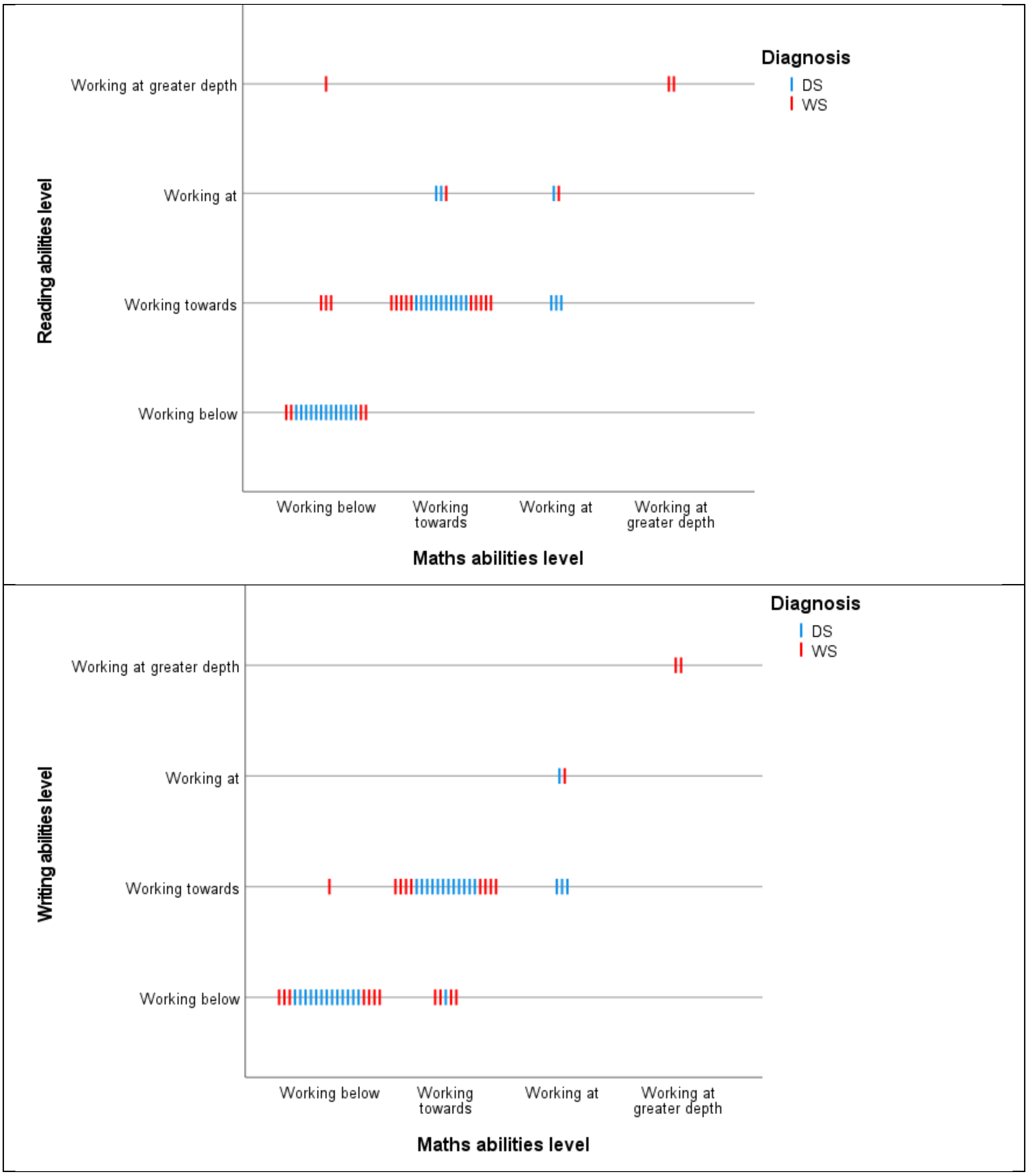
Table 3.8. Distribution of children working below the standard of the national curriculum for mathematics, reading, and writing by clinical group.

Year group	Maths		Reading		Writing	
	DS	WS	DS	WS	DS	WS
EY	4 (44%)	1 (25%)	4 (44%)	1 (25%)	5 (55%)	2 (50%)
KS1	5 (50%)	4 (40%)	4 (40%)	-	5 (50%)	5 (50%)
KS2	6 (38%)	4 (40%)	5 (31%)	3 (30%)	5 (31%)	4 (40%)

Note: Percentages are calculated on the total number of children in their Key Stage group. EY: Early years. KS1: Key Stage 1. KS2: Key Stage 2.

Figure 3.4 shows two charts where the level at which the child was working at for mathematics was plotted with the level the child was working at for reading and writing, respectively. For both groups it is evident that most of the dots are placed in the diagonal of the graphs, showing that children were working at the same level in maths and reading (DS = 85%; WS = 79%) and in maths and writing (DS = 89%; WS = 72%). Dots placed on the top of the diagonal show children who were reported to be working at the lower level for maths, than for reading or writing. Indeed, nine children (DS = 3; WS = 6) were reported to be working at a lower level in maths than reading targets, while three children (DS = 1; WS = 2) were reported to be working at a lower level in maths than writing. Dots placed under the diagonal show children who were reported to be working at a higher target level for maths than for reading (DS = 3; WS = 0), and at a higher target level for maths than for writing (DS = 4; WS = 4).

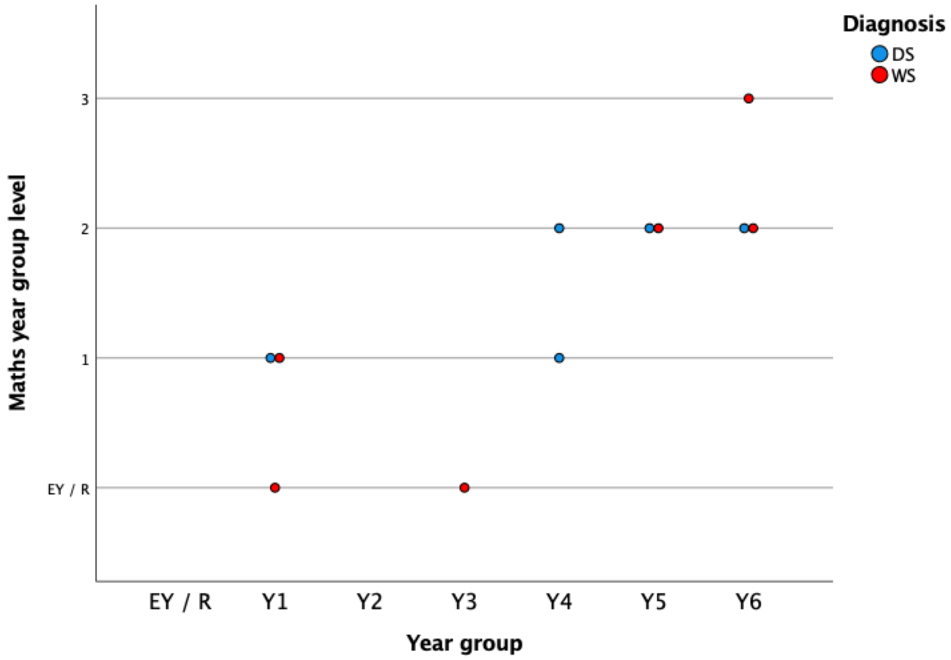
Figure 3.4. Association between child's mathematical level and reading level, and between child's mathematical level and writing level.

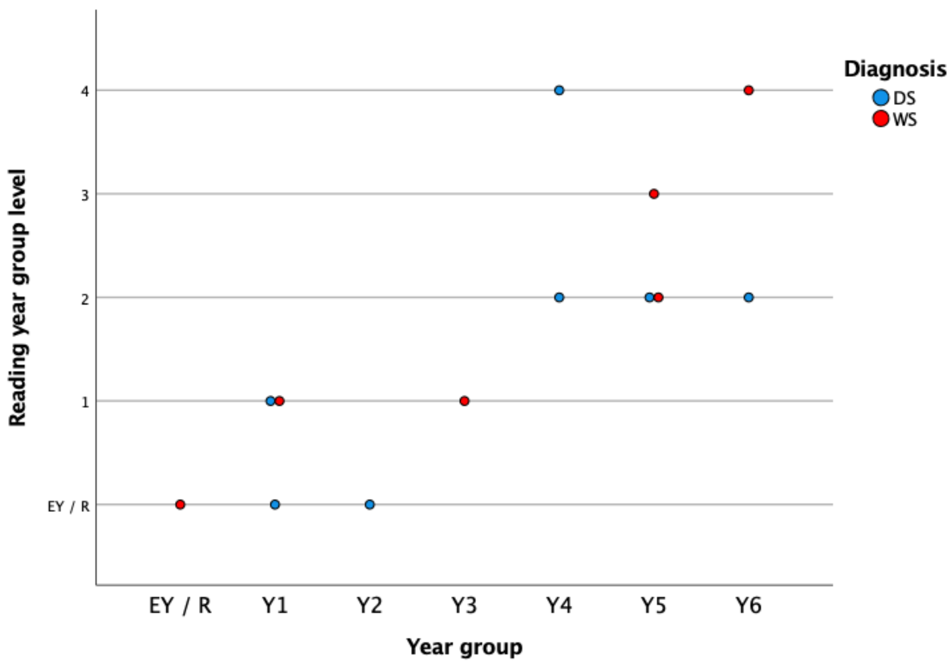


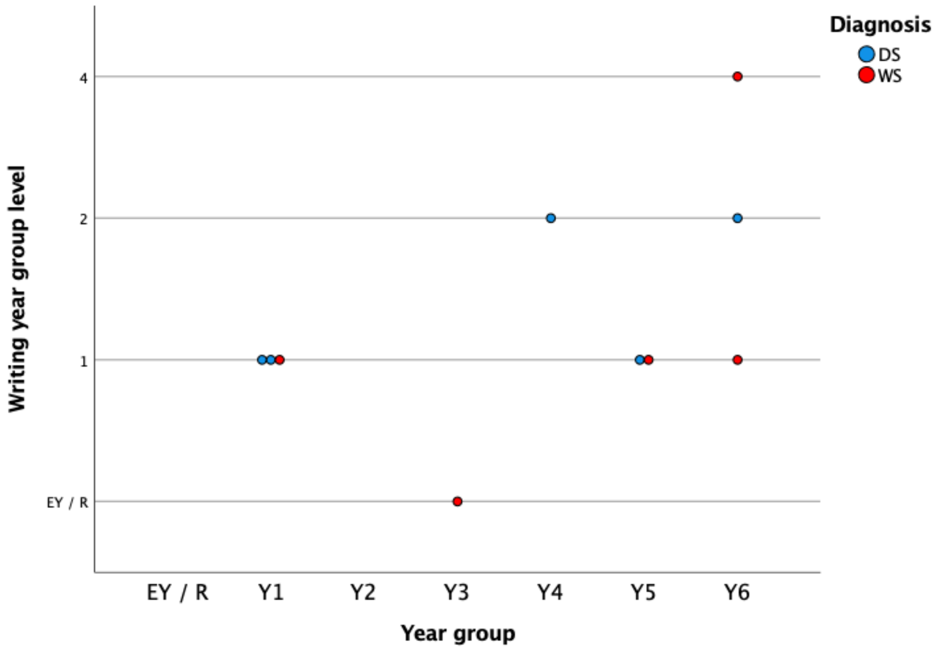
Note: Mathematical level versus reading level: *n* DS = 31. *n* WS = 22. Mathematical level versus writing level: *n* DS = 31. *n* WS = 23.

A minority of parents (20%, for a total of 36 instances) reported the year group level that their child was working at in their open-ended responses (mathematics *n* = 13; reading *n* = 13; writing *n* = 10). This data is plotted against the CA of the child in Figure 3.5. Due to the small size of the sub-samples, the data should be interpreted with caution. Figure 3.5 shows that of all the respondents who provided this information, only 2 children at Y1 (DS = 1; WS = 1) were working at their year group level in mathematics, while most of the children were working at a lower level than the school year they were attending. Moreover, the graph shows that the gap between the year group that children were attending and the maths year group they were working at increases with their age. This might be attributed to the fact that almost all the older children are reported to be working at Y2 group level in maths. The same pattern is replicated for reading. However, here the number of children working at their group level is higher (DS = 2; WS = 2) and the upper limit of the range of year group the children are working at is higher (Y4) than the one reported for mathematics (Y2). Regarding writing skills, most of the children were reported to be working at Y1 level, but some children were reported to be working at Y2 and Y4 levels. Only 3 of the participants were working at their group level (DS = 2; WS = 1).

Figure 3.5. Year group each child is working at in mathematics, reading, and writing, by group.







Note: Mathematical level: *n* DS = 5. *n* WS = 6. Reading level: *n* DS = 7. *n* WS = 6. Writing level: *n* DS = 5. *n* WS = 5.

Finally, Table 3.9 shows the analysis of the specific items of the VABS 2 which focused on mathematical skills. For both groups, most of the children could count at least 10 objects one by one, and that they were able to recognise numbers independently or with some sort of support. Three children with DS aged between 5 and 6 years could not count 10 objects, two of them could not recognise numbers and discriminate them from letters. Moreover, while 20% of parents of children with DS reported that their child performed these tasks independently only partially or only sometimes, only 8% of the parents of children with WS selected this option. As for the use of money and reading the time, the data showed a similar pattern for the two groups where at least half of the children in each group demonstrated an understanding of the function of money and of the clock. However, when it came to using money in terms of identifying the value of coins and using money and reading the time using a clock, parents reported that their child could not perform the task independently. The same pattern was observed for the use of the calendar. Respondents reported that most of the children in both groups were able to say the current day of the week either independently or with support, but they could not use a calendar.

Table 3.9. Responses to the VABS 2 items focused on mathematical skills by clinical group.

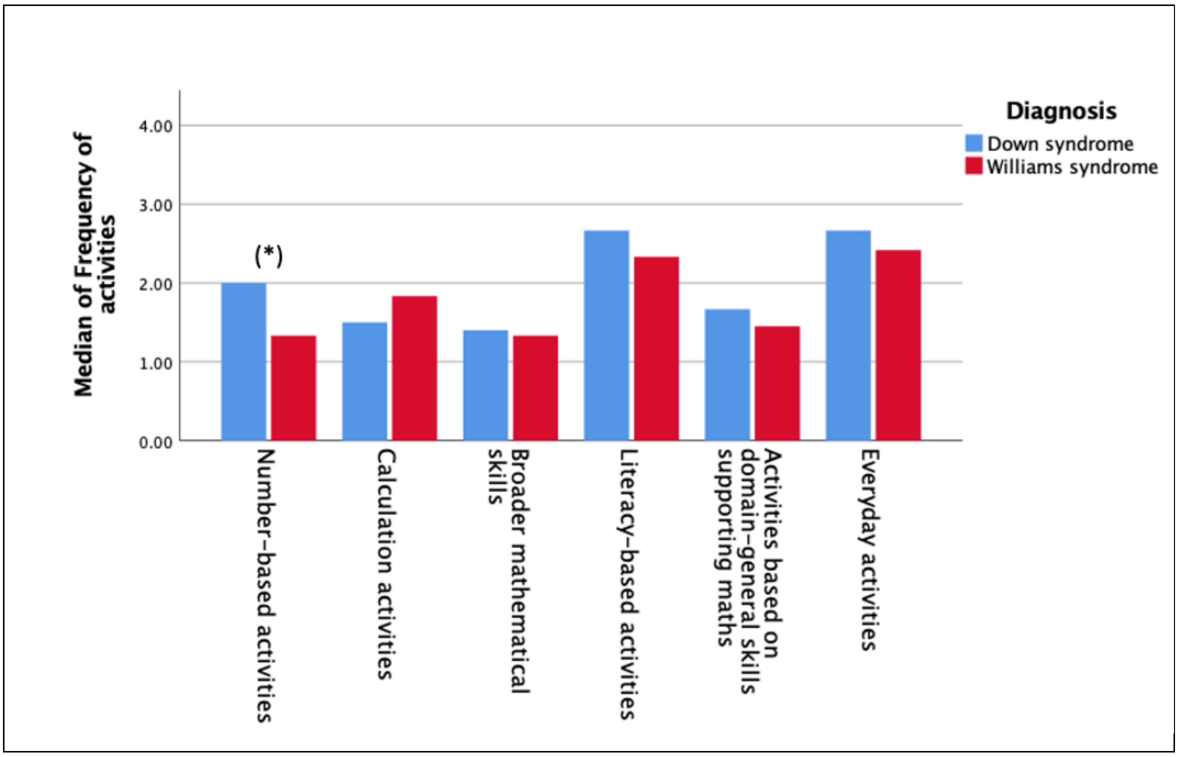
VABS 2 item	DS								WS							
	Usually		Sometimes		Never		I don't know		Usually		Sometimes		Never		I don't know	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Counting																
Counts at least 10 objects one by one	25	71	7	20	3	9	0	0	22	92	2	8	0	0	0	0
Number recognition																
Identifies one or more alphabet letters as letters and distinguishes them from numbers	31	89	2	6	2	6	0	0	22	92	2	8	0	0	0	0
Use of money																
Demonstrate understanding of the function of money (for example, says, "Money is what you need to buy things at a store", etc.)	20	57	7	20	8	23	0	0	17	71	4	17	2	8	1	4
Identifies penny, pence and pounds by name when asked; does not need to know the value of coins	4	11	9	26	18	51	4	11	2	8	13	54	7	29	2	8
States value of penny, pence and pounds	1	3	10	29	22	63	2	6	1	4	8	33	15	63	0	0
Demonstrate understanding that some items cost more than others (for example, says, "I have enough money to buy gum but not a candy bar"; "Which pencil costs less?" etc.)	1	3	7	20	24	69	3	9	1	4	9	38	13	54	1	4
Discriminates between bills of different denominations (for example, refers to £1 bills, £5 bills, etc., in conversation; etc.)	1	3	6	17	24	69	4	11	1	4	2	8	21	88	0	0
Use of the calendar																
Says current day of the week when asked	9	26	12	34	13	37	1	3	7	29	15	63	2	8	0	0
Points to current or other date on calendar when asked	6	17	6	17	21	60	2	6	4	17	6	25	14	58	0	0
Reading the time																
Demonstrates understanding of function of clock (for example, says,	18	51	7	20	10	29	0	0	20	83	3	13	1	4	0	0

“Clocks tell time”; “What time can we go?” etc.)																
Tells time using a digital clock or watch	2	6	8	23	22	63	3	9	1	4	11	46	11	46	1	4

3.3.2 Home-based learning activities and resources

Parents were asked to provide information about the frequency of home-based learning activities. Figure 3.6 shows for each activity the median of the frequency reported by respondents.

Figure 3.6. Frequency of home-based learning activities by clinical group.



Note: *n* DS = 35. *n* WS = 24. Scale: (0): Never, (1): Less than once a week (2): Once a week, (3): More than once a week (4): Every day.
 $*p < .05$

A significant difference in the frequency of home-based learning activities between the DS group and the WS group was found only for the number skills category U = 201.0, p = .001, with parents of children with DS (*MdN* = 2.00, *IQR* = 1.67 – 2.33) engaging more often in activities including counting and number recognition than parents of children with WS (*MdN* = 1.33, *IQR* = 1.00 – 1.92). Because a significant group difference in the frequency of home-based learning activities was found only for one category, statistical analyses included in this section were conducted both separately for the two groups and on the whole sample. Because the same pattern of results was found when analyses were run separately for the two groups and for the whole sample, the statistics and the analyses reported below are collapsed by clinical group to increase the power of the statistical analyses.

Table 3.10. Distribution of the frequency of home-based learning activities for the whole sample.

Activity Frequency	Formal			Informal			Total			
	Q1	*Q2*	*Q3*	*Q1*	*Q2*	*Q3*	*Q1*	*Q2*	*Q3*	*Range*
Number skills	1.00	1.33	2.00	1.67	2.00	2.33	1.33	1.66	2.17	0.17 – 3.00
Calculation skills	1.00	1.66	2.33	1.00	1.66	2.33	1.00	1.50	2.17	0.00 – 2.83
Broader mathematical skills	n/r	n/r		n/r	n/r	n/r	1.00	1.33	2.00	0.20 – 2.83
Literacy skills	2.33	2.67	3.00	2.00	2.67	3.00	2.17	2.50	1.67	1.17 – 3.00
Domain-general skills supporting maths	n/r	n/r		n/r	n/r	n/r	1.00	1.67	2.50	0.33 – 3.00
Everyday life	n/r	n/r		n/r	n/r	n/r	2.17	2.50	2.83	1.33 – 3.00

Note. N = 59. n/r: not reported. Scale: (0): Never, (1): Less than once a week (2): Once a week, (3): More than once a week (4): Every day.

Table 3.10 shows that three mathematics-based activities and the activities related to domain-general abilities supporting mathematical development were reported to occur less than once a week while literacy-based activities and everyday activities were reported to occur once a week.

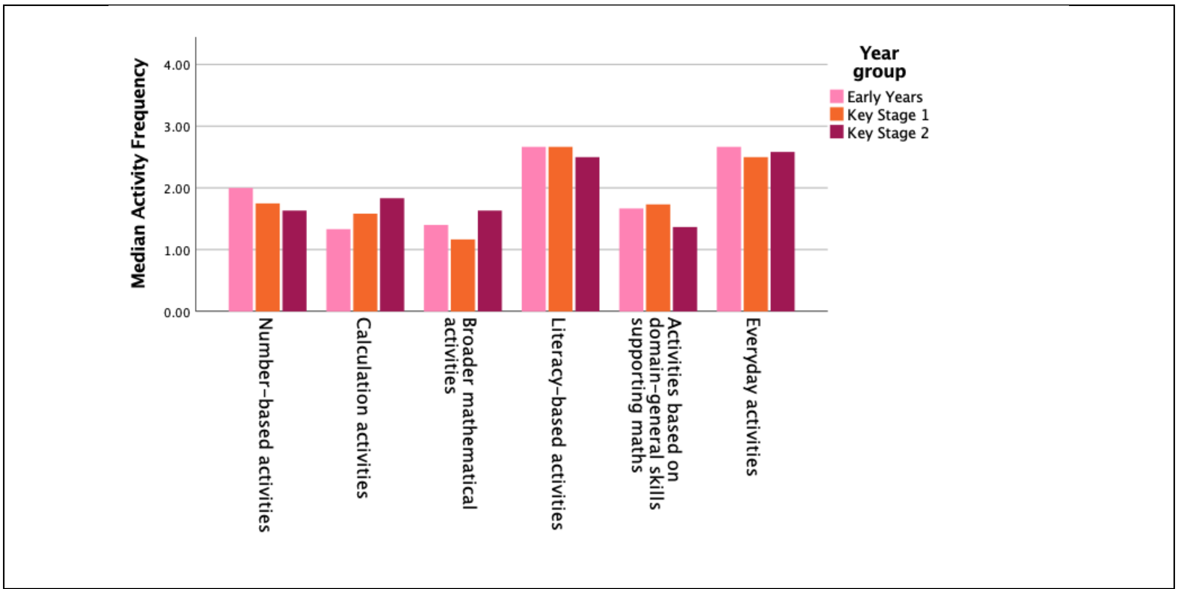
Overall, there was a statistically significant difference between the frequency of occurrence of home learning activities $\chi2(5) = 145.93$, $p < .001$. The post hoc analyses with Wilcoxon signed-rank test were conducted with a Bonferroni correction applied, resulting in a significance level set up at $p = .003$ (calculation: .05 / 15 = .003). Statistically significant differences were found between the frequency of occurrence of literacy activities and all the other categories, except for everyday life activities (z values between 5.93 and 6.49, all p-values < .001) and between the frequency of everyday life activities and the occurrence of activities included in all the other categories, except for literacy (z values between 5.76 and 6.56, all p-values < .001). No significant differences were found between the frequencies of the three mathematics-based activities.

When comparing the frequency of formal and informal activities within the mathematical-based activities and the literacy-based activities, a significant difference was found only with respect to the frequency of activities based on number skills ($z = -4.43$, $p < .001$) with participants performing informal number skill-based activities (*MdN* = 1.67, *IQR* = 2.00 - 2.33) more often than the formal ones (*MdN* = 1.00, *IQR* = 1.33 - 2.00).

In their comments, parents reported a wide range of different mathematics-based activities such as “count and sort toys”, “playing shops using pretend money”, “chant in 5s”, “read house numbers and bus stop adverts”, and having “discussions over dinner about maths problems”. Moreover, the analysis of the comments confirmed that there were several challenges, on both the child’s and the parent’s side, which affected the occurrence of mathematics-based activities at home. The challenges reported on the child’s side were lack of motivation, lack of interest in mathematics, and difficulties with getting the child’s attention after the day at school. One parent reported “I find that after school my child is tired and has limited tolerance for attending to further educational activities” (P2; WS) and another parent said, “I try to involve him, but he does not seem interested, and often prefers to read instead” (P44; WS). The challenge reported on the parent’s side was “lack of time”, with some parents reporting their wish to be more consistent in engaging in these activities with their child.

A series of Kruskal-Wallis *H* tests were conducted to determine if the frequency of the home-based learning activities was different for different age groups. As shown in Figure 3.7, no statistically significant differences between the mean rank scores of the frequency of the type of home-based activities were found between the three year groups (*H* values between 0.86 and 5.47, all *p*-values >.065).

Figure 3.7. Frequency of home-based learning activities by year group.

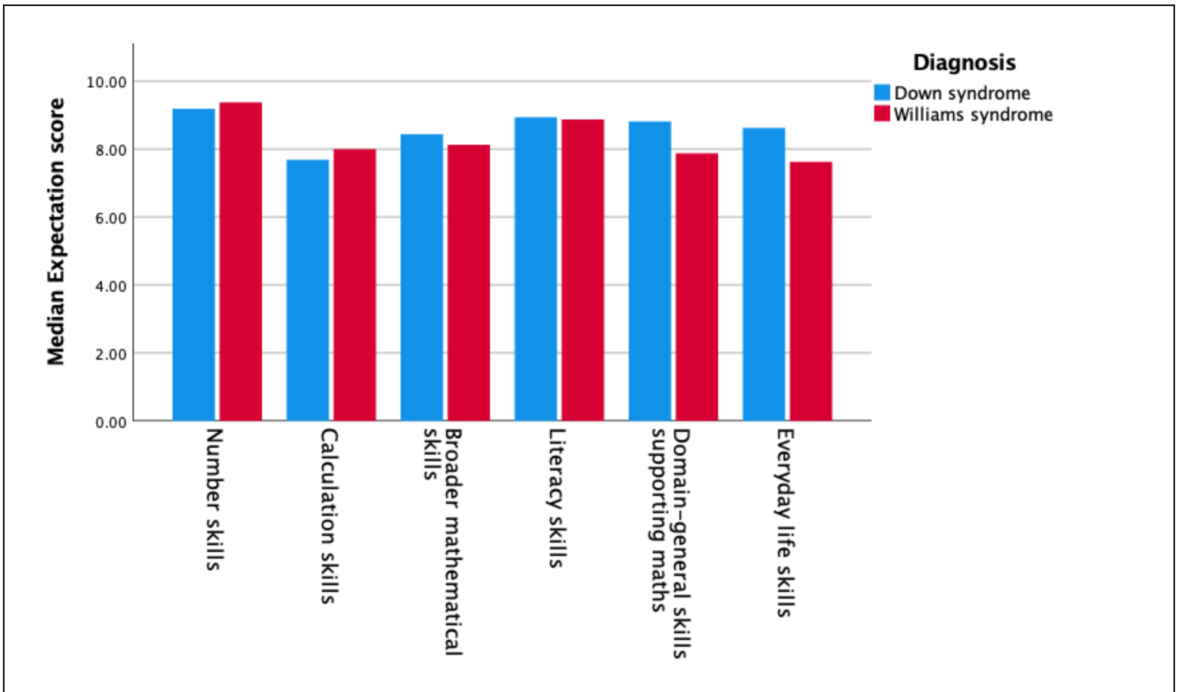


Note. *n* EY = 13. *n* KS1 = 20. *n* KS2 = 26. Scale: (0): Never, (1): Less than once a week (2): Once a week, (3): More than once a week (4): Every day.

3.3.3 Parents' expectations

Figure 3.8 shows the median parental expectation scores by groups. No significant differences between the DS and the WS groups were found with respect to parents' expectations towards their child's abilities at the end of primary school (*U* values between 316.5 and 454.5, all *p*-values >.110). As such, all the statistics reported in this section and further analyses are collapsed by clinical group.

Figure 3.8. Parental expectations by clinical group.



Note: *n DS* = 35. *n* WS = 24. Expectation score range: 0-10.

Table 3.11 shows that the median expectation score (ES) ranged from a minimum of 7.75 (calculation skills) to a maximum of 9.25 (number skills). A statistically significant difference between parents' expectations scores in the different categories was found: $\chi2(5) = 66.56$, $p < .001$. The post hoc analyses with Wilcoxon signed-rank test were conducted with a Bonferroni correction applied, resulting in a significance level set up at $p = .003$ (calculation: .05 / 15 = .003). Parents' expectations for the control items were significantly lower than other categories (*MdN* = 4.00, *IQR* = 1.78 – 5.38). This provides assurance for the quality of the data collected as this result was expected. In fact, the items included in the control category involved skills that are targeted in academic years later than KS2. Moreover, statistically significant differences were found between parents' expectations towards number skills and all the other categories, except for literacy (*z* values between −6.03 and −3.17, all *p*-values < .001), between parents' expectations towards calculation and all the other categories (*z* values between 3.28 and 6.03, all *p*-values < .001), and between parents' expectations towards literacy and broader mathematical skills (*z* = 4.28, *p* < .001).

Table 3.11. Median and dispersion of the distribution of parental expectations for the whole sample.

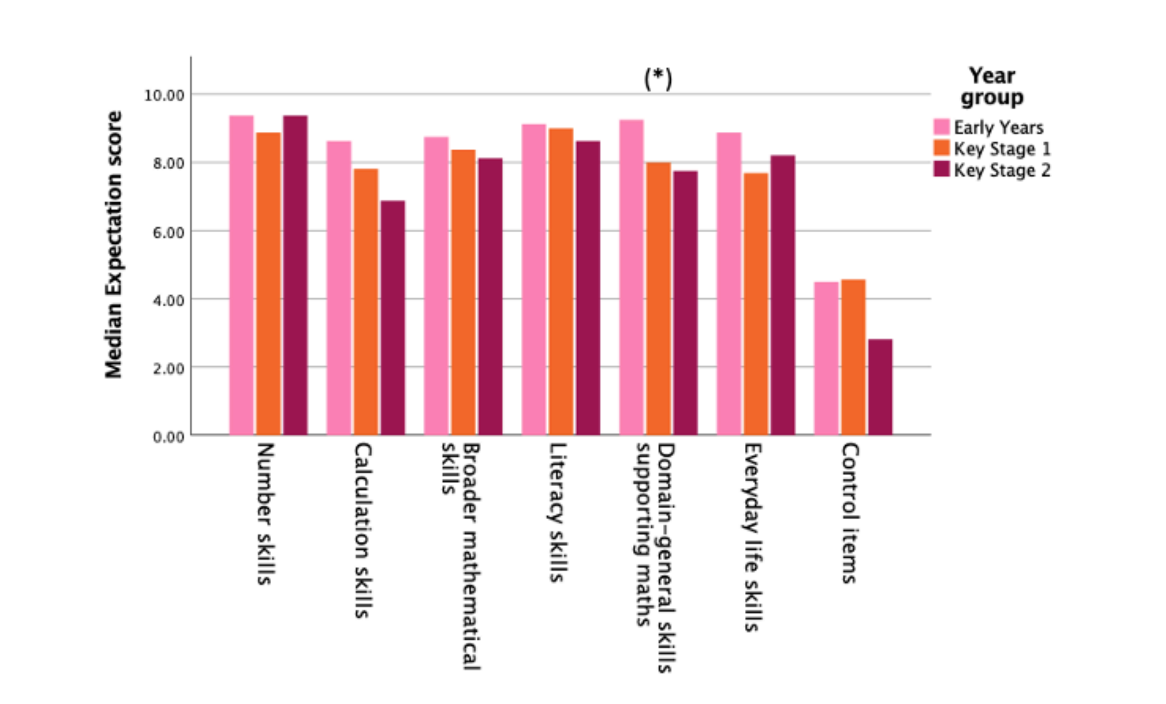
Expectation category	*Q1*	*Q2*	*Q3*	*Range*
Number skills	8.00	9.25	9.75	0.13 – 10.00
Calculation skills	5.75	7.75	9.13	0.38 – 10.00
Broader mathematical skills	6.75	8.38	9.13	0.75 – 10.00
Literacy skills	8.00	8.88	9.38	1.13 – 10.00
Domain-general skills	7.13	8.25	9.25	2.88 – 10.00
Everyday life	7.13	8.29	9.13	1.75 – 10.00
Control items	1.78	4.00	5.38	0.00 – 10.00

Note. *N* = 59. Expectation score range: 0-10.

Despite having overall high levels of expectations for their child's academic abilities, one parent reported that "[…] she's in her final year at primary [school]. I would have hoped she'd have achieved more of these as I would have had higher expectations if she was just starting school" (P51; DS). Also, when commenting on their expectations for their child's academic abilities, most of the parents stressed the importance of functional skills, with one parent reporting "I want her to have life skills, to be able to live and maybe work independently, to handle money day to day and to be able to budget or understand which product to buy" (P28; WS).

A series of Kruskal-Wallis *H* tests were conducted to determine if the parental expectations for each category were different for different year groups. As shown in Figure 3.9 significant difference in parental expectations between the year groups was found in one instance, indicating that for that category the distribution of at least one year group was different from the distribution of another year group. Pairwise comparisons were performed using Dunn's procedure with a Bonferroni correction for multiple comparisons. Statistical significance was accepted at the *p* < .017 level (calculation: 0.5 / 3 = .017). A significant difference in parental expectations was found for domain-general skills, *H* (2) = 8.52, *p* = .014. Post hoc analyses revealed statistically significant differences in parental expectations scores between EY (mean rank = 41.96) and KS2 (mean rank = 28.50) (*p* = .012).

Figure 3.9. Parental expectations by year group.

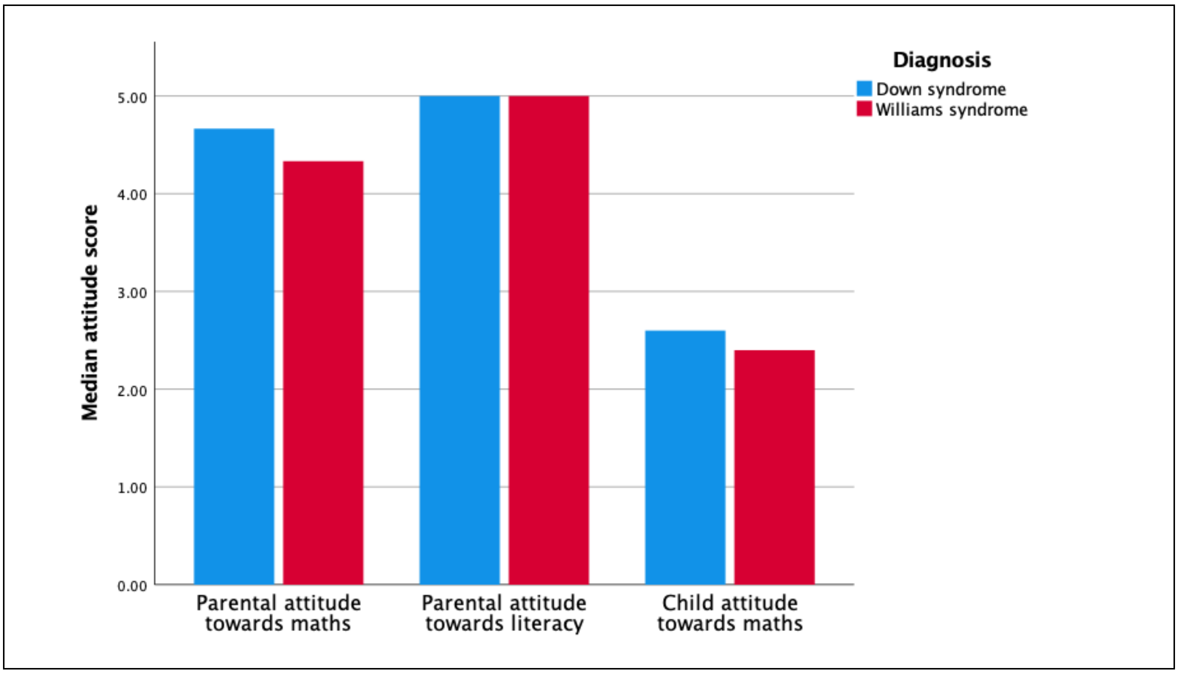


Note: n EY = 13. n KS1 = 20. n KS2 = 26. Expectation score range: 0-10. *$p < .017$.

3.3.4 Parents' and children's attitudes towards literacy and mathematics

Figure 3.10 shows the median scores of parental and child's attitudes towards literacy and mathematics by groups. No significant differences between the DS group and the WS group were found in parents' attitudes towards literacy: $U = 389.0$, $p = .509$, parents' attitudes towards mathematics: $U = 394.5$, $p = .682$, and in the children's interests towards mathematics: $U = 329.5$, $p = .404$. As such, the statistics reported in this section and all further analyses are collapsed by clinical group.

Figure 3.10. Parent and child attitudes towards literacy and mathematics by clinical group.



Note: *n* parental attitude DS = 35. *n* parental attitude WS = 24. *N* child attitude DS = 33. *N* child attitude WS = 23. Attitude score range: 1-5.

Table 3.12 shows the attitude scores for the whole sample. Parents reported significantly more positive attitude scores towards literacy ($M = 4.79$, $SD = 0.61$) as compared with mathematics ($M = 4.34$, $SD = 0.86$): $z = 4.65$, $p < .001$. One parent reported "I know that some of my reluctance to tackle maths with my daughter is my own poor experience of maths at school […] and people with Down's syndrome find maths difficult so perhaps I've been a bit defeatist" (P15; DS).

Table 3.12. Median and dispersion of the distribution of attitude scores for the whole sample.

Attitude score	*N*	*Q1*	*Q2*	Q3	Range
PA literacy	59	5.00	5.00	5.00	1.00 – 5.00
PA mathematics	59	4.00	4.66	5.00	1.00 – 5.00
ChA mathematics	56	2.00	2.50	3.20	1.00 – 5.00

Note: PA: Parents' attitude. ChA: Child's attitude. Attitude score range: 1-5.

Parents' comments on their child's attitudes towards mathematics were mixed. A few parents reported that their child "loves numbers and maths" (P3; DS) and that their child "always chooses to do her maths homework first" (P38; DS). However, most of the parents reported that their child found mathematics "very hard" and reported different levels of engagement with the subject. One parent reported that their child "actively resists it" (P24; DS), one respondent said that their child “wouldn’t choose to do maths homework but is okay about it with encouragement” (P31; WS), while another parent commented that their child often “enjoys maths more than she thinks she is going to” (P20; DS). Three respondents reported all the statements presented in the list as not appropriate, but they did not provide additional comments to explain their rate.

A one-way ANOVA was conducted to determine if the child’s attitude towards maths was different for different year groups. Mean scores of child’s attitudes towards mathematics were low for all groups (mean between 2.59 and 2.85) and no significant differences between the year groups were found, $F(2, 55) = .26$, $p = .774$.

3.3.5 Use of technology

Fifty-seven parents out of the 59 participants in the study sample reported that their child used technology at home. The results reported in this section only included data from these 57 participants and present consolidated data for the overall sample.

Table 3.13 shows that 52 parents (91%) reported that their child had access to technology (e.g., tablets, computers, and smartphones) daily, and 72% of respondents reported that their children owned their own iPad or tablet. In their open-ended comments, most of the parents recognised the benefits of the use of technology and described it as an “essential tool of daily life” (P44; WS), with one parent reporting that their child “has a strength using technology and this should be maximised” (P5; WS). Another respondent observed that their child “seem[s] to learn a lot by what she watched” (P45; DS). Technology was mostly used to access videos and television programmes (86% of respondents daily), while it was not used as often for playing video games, making video calls, and reading e-books. When comparing the frequency of watching literacy- and mathematics-based educational programmes, it was found that the average frequency of watching literacy-based TV programme ($Mdn = 3.00$, $IQR = 0 - 3$) was higher than the average frequency of watching mathematics-based TV programme ($Mdn = 1.00$, $IQR = 0 - 3$) and that the difference in the scores was statistically significant, $z = 4.05$, $p < .001$.

Table 3.13. Frequency of technology use at home.

Item	Daily		Weekly		Monthly		Never	
	n	%	*n*	%	*n*	%	*n*	%
Has access to technology	52	91	3	5	2	4	0	0
Watches videos on YouTube	49	86	5	9	1	2	2	3
Watches literacy educational programmes	29	51	2	4	15	26	11	19
Uses drawing apps	14	25	9	16	10	17	24	42
Watches mathematical educational programmes	11	19	9	16	20	35	17	30
Plays video games	8	14	11	19	6	11	32	56
Makes video calls	8	14	12	21	9	16	28	49
Reads e-books	5	9	2	4	7	12	43	75

Note: N = 57.

Furthermore, the analysis of the data related to the use of mathematics-related apps during the month prior completing the survey (Table 3.14) showed that most of the children were using apps targeting counting (63%), matching (53%), and number recognition (49%). At least half of the parents reported that their children were not using apps targeting calculation, work with number lines, and digital puzzles.

Table 3.14. Use of mathematics-related apps at home.

Item	Yes		No		Not Appropriate	
	n	%	*n*	%	*n*	%
Counting apps	36	63	19	33	2	4
Size/matching apps	30	53	25	44	2	3
Number recognition apps	28	49	23	40	6	11
Addition and Subtraction games	23	40	28	49	6	11
Mathematics-related websites	22	39	33	58	2	3
Digital puzzle games	21	37	34	60	2	3
“Filling the gap” number games	18	32	35	61	4	7
Racing games	16	28	39	68	2	4

Note: N = 57.

When commenting on the use of educational apps, parents reported that their child’s level of engagement was high, with one parent reporting that their child “loves the educational apps and will practice and practice things in a way he won’t do with me as he doesn’t like to make public mistakes” (P58; DS). Another parent observed that “through the use of iPad apps/websites her numeracy has made massive strides forward in the last 6 months” (P45; DS).

Finally, Table 3.15 shows parents' responses around their concerns with the use of technology. More than one third of the parents (37%) reported that they were not concerned about their child's use of technology. The remaining parents reported that their main concerns were related to the time spent in front of the screen and to the content of the applications not being appropriate. These results were confirmed by the comments of the parents which reported the use of screen time rules and measures to limit the access to child-friendly applications in order to control the appropriateness of the applications' content. When voicing their concerns, some parents mentioned the lack of knowledge around which applications were useful, and one parent reported that "there aren't enough apps for kids with learning disabilities. There are some great apps out there […] but they aren't adaptable, for example, in terms of speed of response" (P32; DS).

Table 3.15. Parents' responses to the question "Do you have any concern about your child's use of technology?"

Response	*n*	%
I am concerned about the time my child spends on the screen	27	47
I am concerned about undesirable and/or not age-appropriate content	19	33
I am concerned about accidental in-app purchases	14	25
I am concerned about the effectiveness of these apps	7	12
I don't have any concern	21	37

Note: N = 57. Total exceeds 100% because respondents were asked to check all options that applied.

3.3.6 Correlations

To increase the power of statistical analyses and in view of the homogeneity of the two groups, correlations were run on the whole sample rather than separately for the DS and WS groups.

Spearman's correlations amongst the five VABS 2 scale scores, the Child's CA, and the structural and functional indicators of the home learning environment are shown in Table 3.16.

Table 3.16. Spearman’s correlations for study variables.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. AF - Number skills	—																	
2. AF - Calculation	.18	—																
3. AF - Broader maths	.18	.48**	—															
4. AF - Literacy	.36**	.46**	.44**	—														
5. AF - Everyday	.37**	.26*	.48**	.53**	—													
6. ES - Number skills	.16	.54**	.28*	.47**	.24	—												
7. ES - Calculation	.25	.57**	.34**	.44**	.28*	.76**	—											
8. ES - Broader maths	.13	.46**	.44**	.49**	.38**	.76**	.81**	—										
9. ES - Literacy	.18	.49**	.29*	.51**	.23	.83**	.80**	.87**	—									
10. ES - Everyday	.31*	.37**	.24	.31*	.28*	.54**	.62**	.67**	.67**	—								
11. PA - Maths	.01	.23	.27*	.20	.26*	.18	.10	.21	.12	.11	—							
12. PA - Literacy	-.11	.32*	.16	.11	.08	.32*	.36**	.36**	.37**	.30*	.36**	—						
13. ChA Maths [a]	.29*	.35**	.15	.28*	.25	.34*	.48**	.29*	.29*	.43**	.31*	.10	—					
14. CA child	-.32*	.32*	.16	-.02	.02	-.13	-.29*	-.19	-.17	-.27*	.25	.06	-.11	—				
15. VABS Receptive	-.20	.57**	.16	.23	-.02	.36**	.35**	.33*	.36**	.36**	.12	.42**	.23	.39**	—			
16. VABS Expressive	-.19	.49**	.19	.35**	.06	.36**	.32*	.36**	.27*	.27*	.13	.24	.05	.29*	.71**	—		
17. VABS Written	-.20	.56**	. 28*	.26	.08	.36**	.25	.29*	.21	.21	.31*	.20	.23	.57**	.63**	.67**	—	
18. VABS Community	-.28*	.58**	. 32*	.34**	.12	.39**	.37**	.36**	.29*	.29*	.26*	.31**	.26	.52**	.77**	.78**	.83**	—
19. VABS Fine motor skills	-.26*	.44**	.19	.29*	-.06	.31*	.21	29**	.29*	.29*	.29*	.34**	.14	.40**	.77**	.68**	.70**	.75**

Note: N = 59, [a] N = 56, *p < .05, **p < .001. AF = Activity frequency. ES = Expectation score. PA = Parental attitude. ChA = Child’ s attitude. CA = Child chronological age. VABS = Vineland adaptive behaviour scale.

With regard to the relationship between the variables measuring the general level of functioning and adaptive behaviour of the child and the frequency of home-based learning activities, low negative correlations between the frequency of number-based activities and the scores for the VABS community scale, *rs*(59) = -.28, p = .030 and fine motor skills scale, *rs*(59) = -.28, p = .050, were found. A moderate to strong positive correlation between the frequency of calculation-based activities and the scores for all the VABS scales was observed; receptive *rs*(59) = .57, p < .001; expressive *rs*(59) = .49, p < .001; written: *rs*(59) = .56, p < .001; community: *rs*(59) = .58, p < .001; fine motor skills: *rs*(59) = .44, p < .001. Moreover, moderate positive correlations between the frequency of literacy-based activities and the scores for the VABS expressive scale, *rs*(59) = .35, p = .007, and the community scale, *rs*(59) = .34, p = .009, were found. A low positive correlation was found between the frequency of literacy-based activities and the scores for the VABS fine motor skills scale, *rs*(59) = .29, p = .024.

Low positive correlations were found between all categories of parental expectations and the scores of the VABS scales, except for the instances between the expectations for calculation and the scores for the VABS written and for the VABS fine motor skills scales, and for the instance between the expectations for calculation and the scores for the VABS-II written scale.

No significant correlations were found between the VABS scales and the child's attitude towards mathematics.

A low positive correlation was found between child's CA and the VABS expressive scale, *rs*(59) = .29, p = .025. Moderate positive correlations were found between child's CA and the VABS receptive scale, *rs*(59) = .39, p = .002, and VABS fine motor skills scale, *rs*(59) = .40, p = .002. Strong positive correlations were observed between child's CA and the VABS written scale, *rs*(59) = .57, p < .001, and VABS community scale, *rs*(59) = .52, p < .001.

With regard to the relationship between structural and functional indicators of the home learning environment, a moderate positive correlation between the frequency of home literacy-based activities and the frequency of both everyday life activities: *rs*(59) = .53, p < .001, and all the mathematics-based categories of home learning activities was observed, i.e., number skills: *rs*(59) = .36, p = .005; arithmetic: *rs*(59) = .46, p < .001; and broader mathematical skills: *rs*(59) = .44, p < .001. Moreover, a moderate correlation between the frequency of activities supporting arithmetic skills and the frequency of broader mathematical skills was found; $rs(59) = .48$, $p < .001$. In addition, a low to moderate positive correlation was found between the frequency of everyday life activities and the frequency of mathematics-based categories of home learning activities, i.e., number skills: $rs(59) = .37$, $p = .004$; arithmetic: $rs(59) = .26$, $p = .044$; and broader mathematical skills: $rs(59) = .48$, $p < .001$.

There were strong positive correlations between all the parental expectations categories. Furthermore, positive correlations between parental expectations and frequency for all literacy-based and mathematics-based activities were found, except for the frequency of activities supporting number skills, i.e., arithmetic: rs(59) = .57, $p < .001$; broader mathematical skills rs(59) = .44, $p = .001$; literacy: rs(59) = .51, $p < .001$; and everyday life skills: rs(59) = .28, $p = .032$.

There was a moderate positive correlation between parents' attitudes towards literacy and parents' attitudes towards mathematics: $rs(59) = .36$, $p = .005$, and frequency of arithmetic activities: $rs(59) = .32$, $p = .015$, and all the categories of parental expectations (number skills: $rs(59) = .32$, $p = .014$; arithmetic: rs(59) = .36, $p = .005$; broader mathematical skills: $rs(59) = .36$, $p = .005$; and literacy: $rs(59) = .37$, $p = .004$). Moreover, a weak positive correlation between parental attitudes towards mathematics and frequency of broad mathematical activities was observed $rs(59) = .27$, $p = .041$.

A low positive correlation between child's attitudes towards mathematics and frequency of number skill activities: $rs(56) = .29$, $p = .033$ and literacy activities: $rs(56) = .28$, $p = .034$, was observed, and moderate positive correlations between child's attitudes towards mathematics and frequency of activities targeting arithmetic skills: $rs(56) = .35$, $p = .008$, and parental attitude towards mathematics: $rs(56) = .31$, $p = .019$ were found.

Finally, low to moderate negative correlations were found between child CA and frequency of activities targeting number skills: $rs(59) = -.32$, $p = .005$, parental expectations towards their child's arithmetic skills: $rs(59) = -.30$, $p = .024$, and parental expectations towards their child's everyday life skills: $rs(59) = -.27$, $p = .039$. A moderate positive correlation was observed between child CA and frequency of activities targeting arithmetic skills: $rs(59) = .32$, $p = .005$.

3.4 Discussion

3.4.1 Home learning environment in Down syndrome and in Williams syndrome

The first aim of this study was to explore and compare the home learning environment of primary school children with DS and with WS. A series of Mann–Whitney *U* tests showed that the home learning environment provided by parents of children with DS and children with WS was consistent across the two groups in terms of type and frequency of learning activities occurring at home, level of expectations, and parental attitudes towards learning. The only exception was reported for the frequency of counting and number recognition activities. In fact, parents of children with DS provided this type of mathematics-based activities more often than parents of children with WS. This finding might be explained by the analysis of specific items of the VABS that showed that only parents of children with DS reported this area as challenging for their child. In fact only parents of children with DS reported that their child could not count at least 10 objects and that they could not identify numbers. In the instances where parents of children with DS reported that their child could perform these tasks, a higher number of parents of children with DS than parents of children with WS reported that their child needed support to perform these tasks.

The lack of significant group differences for almost all the structural and functional indicators of the home learning environment investigated in this study suggests that the home learning environment is not syndrome specific. Hence, the discussion of the home learning environment presented in this section is based on the whole sample rather than considering each group separately.

3.4.1.1 Literacy-based activities vs mathematics-based activities

In agreement with studies on TD populations (Blevins-Knabe et al., 2000; LeFevre et al., 2009) and with the study predictions, literacy-based activities were reported to occur significantly more frequently than mathematics-based activities. All the mathematics-based activities were reported to occur less than once a week, while literacy-based activities were reported to occur once a week. The same pattern, with literacy-based activities being more frequent than maths-based activities, was observed in the content of educational TV programmes children were exposed to. This finding showed that the home learning environment of children with DS and children with WS was characterised by a richer HLE as compared to the HME. This finding might be explained by the significantly more positive attitudes that parents reported towards literacy as compared to mathematics. Alternatively, this result could also be explained by the relatively wider availability of literacy-based resources when compared to mathematical-based resources and by the cultural tendency of giving more importance to literacy-based skills rather than to mathematics-based skills.

In agreement with the findings reported by Blevins-Knabe et al. (2000) in TD population, which showed that parents who engaged with their child more often in literacy-based activities offered mathematical activities more frequently, it was found that parents that engaged more in literacy-based activities provided a more varied and richer home learning environment. The same pattern was observed for the everyday life activities, in that, parents who frequently engaged with their child in activities supporting their child's academic skills, also engaged more in activities supporting other areas of their child's development, such as social skills and activities supporting the development of life-skills and higher levels of independence.

3.4.1.2 Formal activities vs informal activities

The results of this study showed that parents engaged with their child in both formal and informal literacy and number-based and calculation activities. The only statistically significant difference between the frequency of formal and informal activities was found for the number skills category, in that both groups reported more frequent informal activities than formal activities. This finding showed that parents were more familiar with informal mathematical activities supporting counting and digit recognition skills than with formal activities supporting these skills.

3.4.1.3 Mathematics-based activities

No significant differences between the frequency of the three mathematics-based activities was found. This finding showed that parents of children with DS and of children with WS provided a varied HME including activities supporting counting and number recognition as well as more advanced mathematical skills, such as calculation and broader mathematical skills. This finding is in contrast with this study's predictions and with previous studies in TD populations, which showed that counting was the most frequent mathematics-based activity occurring at home (Blevins-Knabe et al., 2000). This discrepancy might be explained by the different survey used to investigate the home learning environment or by the sampling technique, which for the current study might have led to the selection of participants particularly interested in mathematical activities. Another explanation could be the different mean CA of the samples included in the studies. While the study investigating TD populations included pre-school children, the current study includes older children attending primary school.

The results also indicated that parents who offered more learning opportunities supporting calculation skills also engaged more frequently in activities supporting broader mathematical skills. This finding seems to suggest that parents that offered activities targeting more advanced mathematical skills, also provided a richer HME that included activities that focused on other aspects of mathematics, such as functional mathematics and geometry.

3.4.1.4 Parental expectations at the end of primary school

Expectations of parents of children with DS and with WS were generally high. This finding is in line with previous studies investigating parental goals and expectations for their child with DS, in relation to literacy (Al Otaiba et al., 2009; Ricci & Osipova, 2012). Moreover, parents with high expectations tended to have high expectations overall, rather than for a specific category of their child's academic abilities. Parents had significantly higher expectations for their child's literacy skills and number-based skills compared to the other categories. Expectations on calculation skills were significantly lower than the other categories. These findings contrast partially with the study's predictions in that parental expectations do not seem to be in line with the maths targets set by the English national curriculum, especially for the calculation component.

Furthermore, parents who had higher expectations for their child were, in general, offered more frequent learning activities at home, however this pattern was not found for number-based activities. This suggests that while the frequency of number-based activities was not influenced by parental expectations, all other types of learning activities were indeed impacted by the level of parental expectations. These findings agree with previous studies in TD population (del Rio et al., 2017) and apply to both advanced mathematical activities and to literacy-based activities.

3.4.1.5 Child and parental attitudes

In contrast with the study predictions, the results related to the child's and parental attitudes towards mathematics and literacy neither support nor disagree with previous studies in TD populations. In fact, while the study by Blevins-Knabe et al. (2000) reported that the frequency with which parents involved their child in mathematical activities was positively correlated to their own attitudes towards mathematics, the current study found a weak positive correlation only between parental attitudes towards mathematics and the frequency of broad mathematical activities. This finding seems to suggest that parental attitudes towards mathematics do not impact the frequency of activities supporting basic mathematical skills and calculations. However, respondents with a more positive attitude towards mathematics were the ones who were more likely to take advantage of opportunities to incorporate maths elements (such as quantities, shapes, and numbers) into their everyday activities.

Results also showed that, overall, children had low attitudes towards mathematics. This finding contrasts with the findings shown by Lusby and Heinz (2020), who reported that children with DS displayed a positive interest in reading, irrespective of their age. This inconsistency could be attributed to several factors, including the higher number of challenges reported by parents in their children's mathematical learning compared to reading. Additionally, it may be influenced by the measurement scale used to assess this construct. Possible improvements to the scale are discussed in the concluding section. Finally, the findings showed that children with more positive attitudes towards mathematics tended to engage more often in both literacy and mathematics-based activities at home. This could be related to children's attitude towards learning in general, rather than their specific attitude towards mathematics and it should be further investigated.

3.4.1.6 Use of technology

Most respondents reported that their child had daily access to technology at home. This finding underscores a heightened use of technology for home-based activities in these populations, even before COVID-19 pandemic, in comparison to what had been documented in prior studies (Trenholm & Mirenda, 2006). In line with the study predictions, the results showed that the usage of technology at home was consistent with the pattern observed in the learning activities occurring without the use of technology. Specifically, children were more frequently reported to engage with literacy-based TV programmes rather than mathematics-based ones.

One of the sub-themes that was identified through thematic analysis of the comments of respondents was the positive aspects related to the use of technology, regarding the advantages and the support that technology provided to children. The positive attitude shown by most of the respondents towards the use of technology to support their child is a noteworthy aspect that educators and researchers should consider when developing educational training and resources. However, as the data was collected before COVID-19 pandemic, it is likely that the use of technology, both in terms of frequency and type of activities, has changed and this aspect of the home learning environment should be re-evaluated. Finally, it would be interesting to further investigate whether parental concerns related to the use of technology, and particularly those related to the time spent in front of the screen, have changed and how after the COVID-19 pandemic.

3.4.2 Academic performance and general level of functioning of the child reported by parents

The second aim of this study was to collect parental perspectives on the general level of functioning and on the academic performance of their child, and to investigate the interactions between the characteristics of the home learning environment and the child's general level of functioning.

One third of the participants reported that their child was working below the standards of the national curriculum in at least one of the three academic areas investigated. Almost half of the respondents reported that their child was working towards the expected standard and very few respondents reported that their child was working at expectations or at greater depth within the expected standard. The analysis of the missing answers reported for this item shows a higher count for mathematics than for reading and writing. The greater number of missing answers in the mathematics category than the others may indicate that parents have a less comprehensive understanding of their child’s performance in this subject at school.

When comparing the target level the child was working at in mathematics with the level of the other academic areas, the analyses showed that most of the children were working at the same level in maths as they were working in reading and in writing. For example, if they were working at expectations in maths they were also working at expectations in reading and in writing. The comparison of the year group that the child was attending at school with the year group the child was working at across the different academic areas showed that most of the children were working at a lower level than the school year they were enrolled in. Moreover, the gaps between the year attended and the year group the child was working at tended to increase with the age of the child. Notably, for mathematics, Y2 targets seemed to be a significant challenge for both children with DS and children with WS.

The analyses of specific items of the VABS revealed that while all parents of children with WS reported that their child were able to count to 10 objects and to recognise numbers, three children with DS failed to perform these tasks. This finding is in line with previous literature reporting challenges with counting tasks for children with DS, but not for individuals with WS (Onnivello et al., 2019; Van Herwegen & Simms, 2020). Since previous studies have shown that individuals with WS have good knowledge of counting names but that understanding of numbers and how they relate to each other is often poor (Van Herwegen & Simms, 2020), these results may suggest that while parents of children with WS recognise the difficulties that their child faces, they might underestimate the difficulties that they encounter specifically with counting. Further studies should investigate if this is the case. Moreover, the analysis of specific items of the VABS showed that even if most of the children could not use money, read the time and use the calendar independently, half of the children in each group demonstrated an understanding of the function of the money, of the clock and of the calendar. These findings together with the observation that a very limited number of respondents chose the option “don’t know” when completing these items (always less than 10%) indicated that respondents performed these activities with their child at home. These findings are in line with the ones reported by Ansari (2003) where only 12% of parents of children with WS reported that their child could read the time. However, while this author interpreted this result as the question being inappropriate for the ages tested (age range: 4;5 – 15;0 years), the fact that a limited number of parents in the current study replied “don’t know” to these items suggests that these learning activities occur at home and that they are considered appropriate by the respondents.

In line with the study predictions, a series of Mann–Whitney *U* tests revealed no significant differences between the two clinical groups concerning the general level of functioning across most of the sub-scales of the VABS. The only exception was the VABS expressive sub-scale, for which the DS group reported significantly lower scores than the WS group. This was in line with previous studies reporting expressive skills as the greatest domain of vulnerability within the language domain in DS (Deckers et al., 2019). These findings underscore the reliability of parental perspectives, as they align closely with previous literature.

Correlational analyses conducted on the whole sample showed that the child’s level of functioning was positively correlated with their CA. This indicates that higher scores on all the VABS scales were associated with older ages.

In line with the study’s predictions, the examination of the relationships between structural and functional indicators of the home learning environment and the VABS scales showed that the type and the frequency of learning activities that occurred at home changed with the child’s general level of functioning, as reported by the parents. Children scoring lower on the VABS were frequently involved in activities supporting number skills, while children with higher levels of functioning were frequently involved in calculation-based and literacy-based activities.

Moreover, these findings revealed that the general level of functioning of the child was positively associated with parental expectations. Specifically, higher levels of adaptive behaviours in all scales were linked to higher parental expectations for maths- and literacy-based competences.

3.4.3 Changes of the home learning environment over time

The third aim of this study was to investigate the relationship between CA and the structural and functional indicators of the home learning environment, and specifically the type and frequency of home-based activities and parental expectations.

A series of Kruskal-Wallis *H* tests showed that the frequency of learning activities occurring at home and parental expectations for children with DS and WS did not change across different age groups. Only one significant difference was found between year groups, with respect to the expectations for domain-general abilities supporting mathematical development. In summary, the current study found that the support that primary school children with DS and WS receive at home, measured as the frequency of maths-based activities and literacy-based activities, and parental expectations remained stable across development. The findings about the frequency of home-based activities are in line with the results reported by Lettington and Van Herwegen (2024) with regards to the HLE of individuals with DS and WS and further highlight the observation that CA does not seem to impact the level of support provided by caregivers of individuals with DS and WS to promote their development. However, since this is a cross-sectional study, it is important to exercise caution when interpreting these findings, and it is advisable to look into replicating them by using longitudinal designs.

3.4.4 Limitations of the study

There are some limitations to consider when interpreting the results of this study. First, the observed drop-out rates impact data quality and this has implications for the generalisation of the results. In fact, only 36% of the respondents who used the anonymous link to access the questionnaire and who completed the consent form completed the survey. However, the response rate increases to 60% when it is calculated based on the respondents who completed the first section of the survey before discontinuing their participation. A recent meta-analysis by Wu et al. (2022) reported that the average online survey response rate in education-related research is 44.1% (95% CI [42.3%; .46%]). In this respect the response rate of this study can be considered slightly below average or well above average, depending on the drop-out rate taken into account. Still, a reflection on the elements that would have increased response rates is needed and is presented in the following section.

Second, the sample of this study mainly consisted of highly educated white families based in the UK, and as such it may not capture the full range of experience of children with DS and WS and their families. This limited the range of analyses performed, as SES was not incorporated in the analyses conducted, contrary to what was originally intended. Moreover, this limitation should be kept in mind when interpreting findings related to the nature and frequency of home learning activities and parents' expectations and attitudes. Additionally, because of the nature of the online survey and of the recruiting method employed, this study could only cover the population with access to the internet and to social media, and those associated with a support group. Again, this limitation should be considered when interpreting the results, as the findings presented in this study might not fully represent all families and experiences of children with DS and WS.

Third, a methodological limitation of this study has been that the pilot did not test the entire research process, but it only tested recruitment, administration, and data collection phases. This was due to the limited number of participants that took part in the pilot test (n = 2).

Fourth, the data collected through the web-based survey might be affected by accuracy issues because it relies on self-reported data, which can be influenced by biases where parents may either overestimate or underestimate their child's abilities or the frequency of the home-based learning activities. The inclusion of OEQs where respondents could further elaborate their responses, and the inclusion of control items in Section 2 and in Section 4 of the survey aimed at reducing and controlling for these biases.

Finally, this study might be limited in the extent to which it was able to accurately measure the home learning environment. Studies integrating different measures of the home learning environment, such as naturalistic observation and interviews, would add weight to the evidence collected. Moreover, this study only considered shared activities between the child and the carer. Further investigation of the home learning environment should include shared activities involving other members of the family (e.g., siblings).

3.4.5 Reflections on the “Maths at Home” questionnaire

As shown in Table 3.2, the highest drop-out rate (30 percentile points) was registered at the start of the survey, between Section 0 and Section 1. Indeed, as argued by Ruel et al. (2016), it is possible that the decision of including the child demographic section at the beginning of the questionnaire might have played a role in this high drop-out rate. In fact, on one hand these questions are not engaging for the respondents, especially carers of children with NDCs that are often asked to complete forms and to provide demographic information of their child. On the other hand, these questions might have set the wrong tone for the whole survey and might have given respondents second thoughts about taking part in the study, especially regarding confidentiality and data protection. In future uses of this survey, it might be preferable to include this section towards the end of the survey.

When designing the questionnaire, several measures were implemented to mitigate the drop-out rate. These included a clear description of the rationale of the study at the start of the survey. In hindsight, providing a more comprehensive explanation of the study's objectives might have proved beneficial not only for helping participants understand the research's rationale but also to clarify why such an extensive number of questions was needed and to prevent subsequent frustration. Also, the number and the type of questions to be asked was carefully evaluated and discussed with the participants of the pilot test – for example, CEQ were more than OEQ, as these require less time to be completed. However, the survey still included a high number of questions (n = 66), and this might have been a crucial factor for non-completion rates. In future, it might be useful to evaluate not only the number of questions but also whether to include or not entire sections. For example, the current version of the questionnaire included a section about the use of technology at home to investigate whether this was used or not and with what frequency. Despite this being related to the overarching aim of exploring the home learning environment of children with DS and children with WS, this section may not be essential and could have been omitted from the questionnaire.

Finally, the questionnaire was designed to incorporate a progress bar at the top of each page to visually indicate the percentage of completion. However, the very low-moving progress bar due to the length of the survey might have deterred respondents from completing the survey. Hence, as reported by Ruel et al. (2016) the decision to include the progress bar in each page, as opposed at specific checkpoints, should be evaluated for future use.

Additional improvements of the survey relate to adjustments to single items. First, participants were asked to provide their child’s date of birth in the form of an OEQ using the format DD/MM/YYYY. However, as mentioned above, 4 participants were excluded from the final sample as they either did not provide DOB or because they reported a wrong DOB, which corresponded with the date when they took the survey. To address this issue, in the future CA could be collected in two different sections of the survey asking the same question but rephrased as, for example, “How old is your child (year, month – e.g., 10,1; 10 years and 1 month)?”. Alternatively, the survey might include a question asking the school year the child is attending, this data could then be used as a proxy of the age of the child in case this data is wrong or missing.

Second, overall, the “Maths at Home” questionnaire proved to be a useful tool to describe the home learning environment of children with DS and WS. However, the use of this tool to measure the mathematical abilities of the child resulted in some issues. To collect this data, respondents were asked to provide information about the school performance of their child on maths. The answer was meant to be used as a proxy of the mathematical abilities of the child. The question was phrased as a double-barrelled question[18] and asked the following:

> “What level of **maths** is your child functioning at in school? E.g., if the child is working towards the national curriculum: a) working towards the expected standard, b) working at expectations, c) working at greater depth within the expected standard. If the child is working below the standard of the national curriculum: P-scales”.

Because of the OEQ format, responses of the participants varied greatly, and this had a negative impact on data analysis. In fact, some of the respondents did not use the categories provided by the researcher to describe the level of the targets that their child was working at. Instead, they reported an example of the targets their child was working on. Future implementation of this item should include the use of a CEQ followed by an OEQ which respondents could use to add further comments. Moreover, because the categories indicated in the question were broad, these have been interpreted differently by respondents. In particular, the categories “working towards the expected standards” was selected by respondents both when their child was working towards the target of their year

[18] A double-barrelled question is a question which asks respondents two questions at the same time – in this case “is the child following the English national curriculum?” and “at what level is the child working?” – but only allows one answer.

group (e.g., a child in Y2 working towards a Y2 target) and when the child was working towards the target of an earlier year group, in accordance with the targets set by the school (e.g., a child in Y5 working towards a Y2 target). Furthermore, some respondents referred to different standards than the English national curriculum, such as the Oxford reading stages, or reported their child's ability according to recent assessment (e.g., the Salford Sentence Reading Test for reading). In future this question could be improved by splitting the question in two parts and by providing a clearer and more well-defined description of the categories listed, for example it should explain the meaning of "expected level" or the use of the word "standard". Alternatively, mathematical abilities could be measured using quantitative scales, such the ones used in the study by Ansari (2003).

Another item that could be improved was the measure of the child's attitude towards mathematics (ChA). This construct was measured asking respondents to rate 5 statements about their child's attitudes on a 5-point Likert scale. The data that was collected in the current study turned out to be relatively uninformative and did not yield interesting results in the analyses. Future implementations of this item should include a broader and more diverse set of statements and should directly capture the child's perspective rather than reporting the respondents' view on their child's attitude. For example, the instructions might ask the parent to ask and report the child's view.

Some changes of the wording used in the demographic section would be needed to make the tool more inclusive. For example, the survey asks for the child's "gender" rather than asking for the child's "sex". Also, the question asking for the ethnic origins of the participant should be rephrased and for example ask: "Which ethnicity best describes you?". Furthermore, all the questions included in the demographic section should provide the option "prefer not to say".

Finally, as suggested by Ruel et al. (2016) it might be beneficial to consider incorporating a final question regarding the respondents' experience of completing the survey. This could serve as a valuable starting point for reflections and could provide insights useful for future research projects.

3.4.6 Implications and future directions

The findings presented in this chapter provide a substantial body of information about the structural and functional components of the home learning environment of primary school children diagnosed with DS and WS. Since these findings provide insights into the activities and the resources used in the household to support mathematical development of children with DS and WS and on the frequency of such activities, they could inform the design and development of parental interventions aimed at highlighting opportunities to enrich the home learning environment. On the other hand, these findings could be used by educators and other professionals working with the child to support consistency between different settings and to support the child's development outside the home environment. In fact, one of the sub-themes that was identified through thematic analysis of the comments of respondents highlighted family-school relationships and the important role of the communication between family and school to support the child's development.

A significant area of exploration for future studies will be the relationship between the home learning environment and mathematical development in neurodivergent populations. Research in TD populations has reported a positive relationship between environmental factors and mathematical development (Daucourt et al., 2021). However, since this study did not collect a measure of mathematical development, this association could not be explored. Further research should investigate this relationship and how structural and functional indicators of the HME influence mathematical development.

3.5 Conclusion

This study explores and compares the home-based learning experiences of primary school children diagnosed with DS and with WS using a web-based parental survey.

Results of this study indicate that the home learning environment of children with DS and WS were similar for all the structural and functional indicators investigated, except for the frequency of activities supporting number recognition and counting, in that parents of children with DS offered these activities more often than parents of children with WS. Conversely, type and frequency of home learning activities as well as parental expectations changed based on the child's general level of functioning. In fact, children scoring lower on the VABS were frequently involved in activities supporting number skills, while children with higher levels of functioning were frequently involved in calculation-based and literacy-based activities. Moreover, higher levels of adaptive behaviours were associated with higher parental expectations for maths- and literacy-based competences. These findings indicate that the home learning environment is not syndrome specific.

The results of this study indicate that various learning activities take place in the homes of primary school children diagnosed with DS or with WS, with literacy-based activities occurring once a week and mathematics-based activities occurring less than once a week. The HME was reported to be varied and characterised by activities supporting different mathematical skills such as counting, calculation, and the use of numbers in everyday life contexts. Parents engaged with their child in both formal and informal literacy and mathematics activities, and informal activities occurred more often only when supporting counting and number recognition skills. Results show that parents who provided a rich home learning environment also engaged more in activities supporting their child's broader development. Additionally, findings suggest that parental attitudes towards maths do not impact the frequency of activities supporting basic mathematical skills. However, parents with a more positive attitude towards mathematics were the ones who were more likely to take advantage of the opportunities to incorporate mathematics elements into their everyday activities.

Finally, this study found that the level of support that children diagnosed with DS and WS receive at home during the primary school years does not change across development. On the contrary, the level of parental expectations changed across development, with lower expectation scores reported by parents of older children for domain-general skills related to mathematical development.

Chapter 4: Mathematical abilities in autism: A systematic review

This chapter presents a systematic review of mathematical abilities in autism. As explained in Chapter 1, when it comes to this topic, the overview provided by the existing reviews of literature (Chiang & Lin, 2007; Dowker, 2020; Tonizzi & Usai, 2023) is partial. In particular, only arithmetic word problem and calculation skills have been reviewed systematically, and the reviewed studies only included autistic individuals without intellectual disabilities (ID). The decision to conduct a systematic review rather than using behavioural research to investigate mathematical development in this population stems from the need to identify the specific areas where additional research is needed before conducting additional experimental studies.

The aims of this systematic review are twofold: 1) to provide a comprehensive overview of the studies investigating basic mathematical processes and specific components of mathematics in autism, as described by the multi-level framework by Gilmore (2023) introduced in section 1.2.1, and 2) to describe mathematical profiles of autistic individuals with ID and without ID and across development.

This will help researchers in addressing gaps in the literature and in designing effective teaching programmes and educational tools to support mathematical development in autism.

4.1 Background and rationale of the study

As reported in section 1.3.3, there are currently 3 literature reviews which investigated the development of mathematical abilities in autism: one systematic review by Chiang and Lin (2007), one narrative review by Dowker (2020), and one meta-analysis by Tonizzi and Usai (2023). Despite the significance of these studies, they present limitations which do not allow to fully address the aims of this study.

With respect to the systematic review by Chiang and Lin (2007), one of its main limitations is the limited scope, which includes only studies which used the

Arithmetic subtest of Wechsler assessment, meaning that it only investigated the specific component of arithmetic word problems. Second, the review includes studies that compared scores with the normed population (rather than with a control group), and by doing so the study does not provide enough information to determine whether mathematical abilities in autism are in line with chronological age (CA) and / or mental age (MA). Third, the review does not provide detailed information about the participants that took part in the included studies. For example, it is not clear how and whether a diagnosis of autism was made. Moreover, the authors report that "many of the studies included children with mental retardation" (Chiang & Lin, 2007, p. 548). However, the number of studies which included individuals with ID is unclear, and it is not specified whether autistic individuals with ID showed a different mathematical profile than the ones without ID. This omission constitutes an important gap. In fact, although some autistic individuals demonstrate high intellectual abilities, a high proportion of autistic individuals meet the criteria for ID (Miller et al., 2017), which for some can be characterised by a delayed and slower development of mathematical abilities (Bashash et al., 2003; Brankaer et al., 2011). Fourth, the review presents some methodological limitations that may impact the validity and the reproducibility of the findings. These include the limited number of databases that were searched (*n* = 2), the limited description of the search process, and the lack of documented procedures to assess the quality of the data extraction process.

The narrative review by Dowker (2020) aims to investigate arithmetic abilities in children with specific "developmental cognitive disabilities" including autism. Hence, autism is not the specific focus of the paper. Moreover, due to the search method used, some relevant studies may have been missed. This is because narrative reviews usually prioritise recent literature and may not be as comprehensive as systematic reviews.

Finally, the meta-analysis by Tonizzi and Usai (2023) used stringent inclusion criteria both in terms of the population and of the design of the studies and ended up including 13 studies with a limited age range of participants (between 6 and 16 years old). Moreover, the meta-analysis only included studies with a matched-group design and with a TD control group that were published after January 2000. The inclusion of matched-group research is motivated by the type of analyses conducted. However, the rationale for the temporal limitation is not clearly explained and, again, limits the scope of the study. As for the mathematical components investigated, the studies included in the meta-analysis only assess calculation and arithmetic word problems. Possibly because of the stringent inclusion criteria, the meta-analysis ended up including only studies investigating mathematical abilities of autistic participants without ID. Hence, the results provided by this review do not help the understanding of the mathematical profile of autistic individuals with ID.

In summary, previous reviews of the literature provide a limited overview of mathematical abilities in autism, as they only focus on two specific components of mathematics (calculation and arithmetic word problems) and, where the cognitive profile of the samples investigated was taken into consideration, only autistic individuals without ID were included. As described in section 1.3.3, the findings reported by Chiang and Lin (2007) and by Tonizzi and Usai (2023) are in agreement and indicate that autistic individuals show average arithmetic word problem skills and calculation skills but perform poorer than their typically developing (TD) peers.

4.1.1 Current study

The first aim of this systematic review was to synthesise the research that has investigated mathematical processes and specific components in the autistic population, with the final aim to identify gaps in the current research. To address the first aim, the following research question (RQ) was asked:

- What basic mathematical processes and specific components of mathematics have been examined in autism?

The second aim of this systematic review was to investigate mathematical profiles in autism through a neuroconstructivist perspective. To address this aim, the following RQs were asked:

- What are the mathematical profiles in autism? In particular, are mathematical abilities in line with CA and / or MA measured through assessment of full-scale IQ (FSIQ), verbal IQ (VIQ), and non-verbal IQ (NVIQ)?
- Do CA, MA, or autistic traits explain mathematical abilities in autism?

4.2 Methodology

The protocol of this systematic review was pre-registered prior to analysis of the data in the Open Science Framework[19] web application (https://doi.org/10.17605/OSF.IO/Q3D7C). The main aim of the pre-registration form was to specify the research plan and to improve the quality and the transparency of the research. The pre-registration form has been updated in January 2024.

4.2.1 Literature search

4.2.1.1 Inclusion criteria

To be included in the systematic review, studies had to meet the following criteria:

(1) Include participants with a clinical diagnosis of autism according to the diagnostic criteria at the time of writing, with or without ID, and with no co-occurring chromosomal conditions. ID was defined as having IQ scores below 75 (Spaniol & Danielsson, 2022). Studies in which the clinical diagnosis of autism was not reported were excluded.

(2) Report for the autistic sample both a quantitative measure of mathematical abilities and the assessment used. No restrictions were placed on the study design. However, studies using solely qualitative research methods (e.g., ethnography, action research, interviews, focus groups) and studies reporting on other reviews of the literature were excluded.

(3) Be published in a peer reviewed journal and be available in English.

4.2.1.2 Electronic database search and hand search

The databases of PsycINFO, Scopus, British Education Index, and ERIC and the registry of Web of Science were searched using the search terms: (autis* OR asperger*) AND (mathematic* OR arithmetic* OR numer*).

The search was initially set to target only the most important and relevant key terms: “autis*” AND “mathematic*”. Then, the search terms were updated

[19] Open Science Framework is a free and open-source project management tool that supports researchers throughout their project lifecycle.

iteratively adding disjointed terms either to the left-hand or right-hand side of the AND to get a more exhaustive set of results. Throughout the iterations, the query was optimised and the desirability of adding new terms to the query was evaluated based on the effect on the number of hits (Bramer et al., 2018). New terms being evaluated were not considered necessary and hence deleted if the number of hits increased greatly and included a high ratio of non-relevant references. More general search terms (e.g., "number"), more specific search terms (e.g., "Asperger"), and synonyms retrieved in relevant references (e.g., "arithmetic") were evaluated. Truncation and Boolean operators were used throughout.

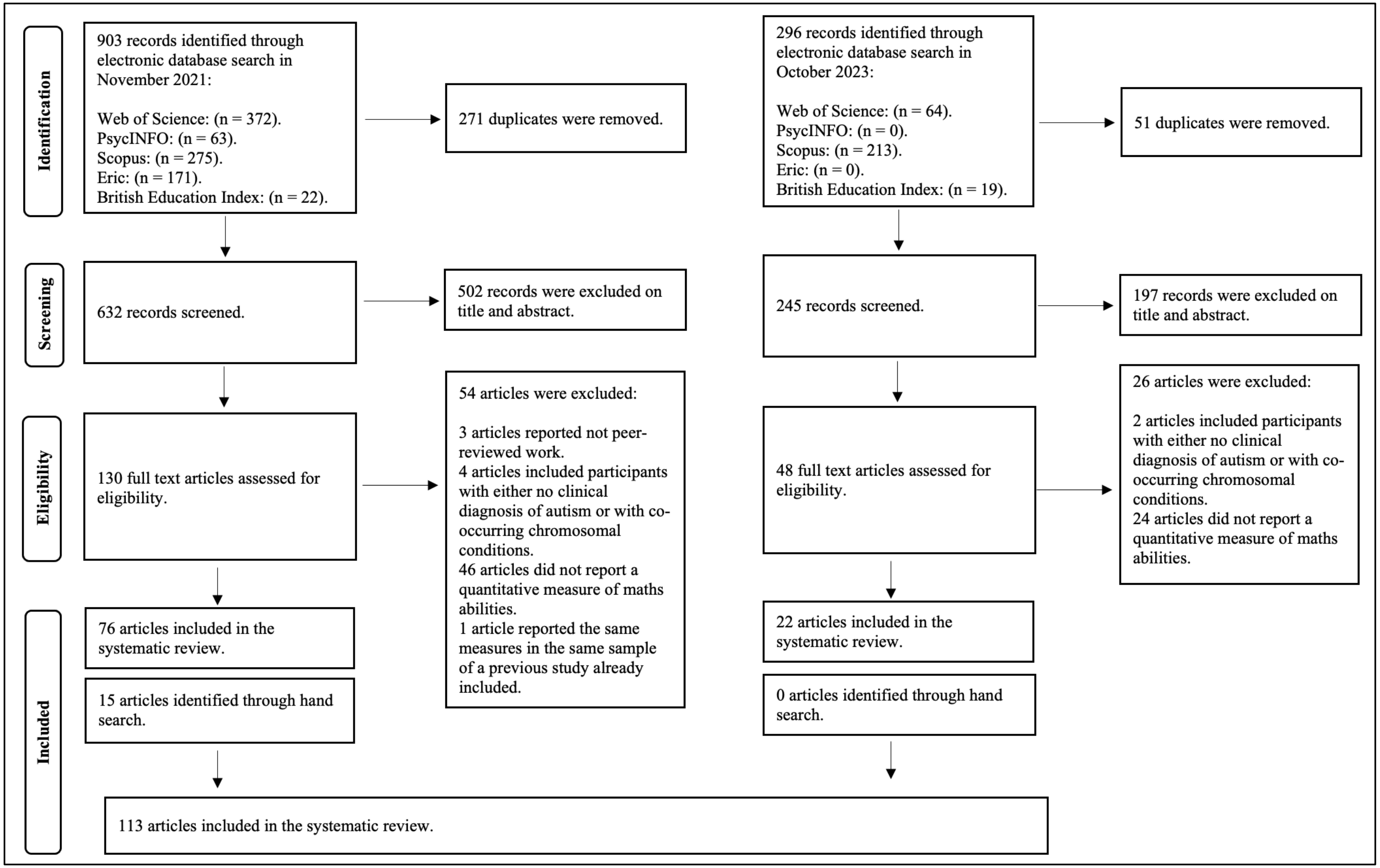
Search terms were limited to the title, the abstract and the keywords of the papers, and date limits were not applied. The search was carried out in the electronic databases under the advice of the specialist librarian from UCL. This search was conducted in November 2021. Reference list of all the studies that ultimately met inclusion criteria from the electronic database search (n = 76) were screened for any additional paper that may have been missed. Additionally, the reference list of the review articles identified through the search process were screened. The hand search process identified 15 additional articles.

The same electronic database search was re-run in October 2023 to check studies that met the inclusion criteria and were published between the date of completion of the first database search (November 2021) and the date of completion of the analysis of the data collected (October 2023). This was done to provide more up-to-date results. The result of this search process led to the inclusion of 22 new articles.

4.2.1.3 Final set of articles

Figure 4.1 provides a summary of the search process. First, in November 2021 903 studies were identified through an electronic database search. Duplicates were removed (n = 271) and title and abstracts of 632 records were screened for relevance against the inclusion criteria. An independent reviewer completed the screening for 10% of all the records (n = 63) to establish reliability. The percentage of agreement was 92% (n = 58). Disagreements were discussed and resolved during an agreement meeting. The full text of 130 eligible articles were then retrieved and screened for eligibility. The rationale for the exclusion of the studies was recorded by the researcher. Overall, the process led to the exclusion of 54 articles. A hand search of the reference list of the 76 included studies led to the inclusion of 15 more articles. The electronic database search run in October 2023 led to the further inclusion of 22 articles. The result of the literature search process was 113 articles, including 114 studies. The total number of studies amounts to 114 because the article by Lincoln et al. (1988) was categorised as two distinct studies, as the paper includes two different studies conducted on two different autistic samples.

Figure 4.1. Flow diagram of the search process.



4.2.2 Data extraction

The data extracted was organised into 3 sections within an Excel dataset: (1) characteristics of the study, (2) characteristics of the participants, (3) mathematical outcomes. A comprehensive list of the data extracted and of the coding rules used to populate the Excel database is provided below.

Section 1: Characteristics of the study. For each study the following data was extracted:

- Author(s).
- Year of publication.
- Country where data was collected.
- Study design, depending on the number of data points collected for the same sample: i) cross-sectional: studies that collect data only for one data point; and ii) longitudinal: studies that collect data from the same participants at multiple data points.
- Type of study, according to the following taxonomy: i) case study: studies which involves an in-depth examination of an individual or a group and which includes a detailed description and analysis of the subject, often based on various data sources, such as clinical reports, interviews, or observations; ii) correlational study: studies measuring a relationship between two variables without any manipulation; iii) case-control study: studies comparing the outcome of a dependent variable (e.g., mathematical outcome) of two or more pre-existing groups; iv) intervention study: studies in which the researcher manipulates a variable to assess the impact of exposure to an intervention; and v) single-case research: study in which the researcher manipulates an independent variable and collects repeated measurement of a dependent variable before (i.e., baseline) and after (i.e., intervention phase) the introduction of the independent variable and where the individual case being studied serves as its own control. Even though single case experimental research falls under the umbrella of interventions, it was differentiated from the category “intervention study” because it uses different data analyses (it typically relies on visual analysis) and because it focuses on the effect of the intervention on the specific case investigated,

with implications on the level of insight and on the generalisation of the study findings.

Section 2: Characteristics of the participants. For each study the following data was extracted for both the autistic sample and for the control group, when included:

- Sample size.
- Group mean (*M)*, group standard deviation (*SD)*, and range of CA.
- Sex of participants.
- Rule for inclusion.
- Diagnosis.
- Diagnostic criteria used for the diagnosis of autism.
- Diagnostic or screening tool used.
- Group *M*, *SD*, and range of measure of autistic traits.
- Type of control group.
- Matching rules.

Section 3: Mathematical outcomes. For each study the following data was extracted for the autistic sample:

- Name of the tool used to assess IQ.
- Group M, SD, and range of FSIQ, VIQ, and NVIQ.
- Name and version of the tool used to assess mathematical performance. The description of the assessments was retrieved from the studies, when available. Otherwise, information about the assessment was retrieved from the test-publishers' website.
- Mathematical basic process or specific component assessed, according to the taxonomy described in section 1.2. The mathematical abilities investigated in the included studies were coded based on the assessment used instead of the mathematical ability named in the study. Studies which did not follow the standard assessment procedure were excluded.
- Group *M*, *SD*, and range of performance for each basic process / specific component assessed.

- Type of score reported. These included: i) standard scores; ii) z-scores; iii) percentiles; iv) raw scores; and v) age-equivalent scores. To compare standard scores from assessments with different mean and standard deviation, all standard scores of mathematical performances were expressed as z-scores. Data points related to studies that only reported percentiles, raw scores or age-equivalent scores were excluded.
- Statistical analyses performed and main findings.

Any information of interest that was not reported in a study was coded as "not reported". In the case of intervention studies and single-case research, only pre-intervention data was extracted. Two independent reviewers verified data extraction for all the included studies. Disagreements were resolved by discussion in agreement meetings and mainly concerned typos made during the data extraction phase.

Charts were produced using Observable, a web-based data visualisation tool, and Microsoft Excel.

4.3 Results

To provide a comprehensive synthesis of the studies included in the systematic review, the results are organised in three sections. The first section includes analyses focused on the metadata and the design of the studies. The second section focuses on the demographic of the autistic samples and reports an overview of autistic traits, sex ratio, mean CA, and mean IQ scores. The final section summarises the findings on mathematical performance categorised using the taxonomy described in Section 1.2.

4.3.1 Characteristics of the studies

One-hundred thirteen articles, including 114 studies were identified in this systematic review. These included 143 autistic samples and reported on 220 measures of mathematical performance.

As shown in Table 4.1, most of the studies were case-control studies (*n* = 53, 46%) followed by single-case research (*n* = 28, 25%), correlational studies (*n* = 23, 20%), intervention studies (*n* = 6, 5%) and case studies (*n* = 4, 4%). Most of the studies (n = 110, 96%) collected data at a single time point, while four studies used a longitudinal design, with only two of them conducting actual longitudinal analyses.

Table 4.1. Studies included in the systematic review categorised by type of study and by study design.

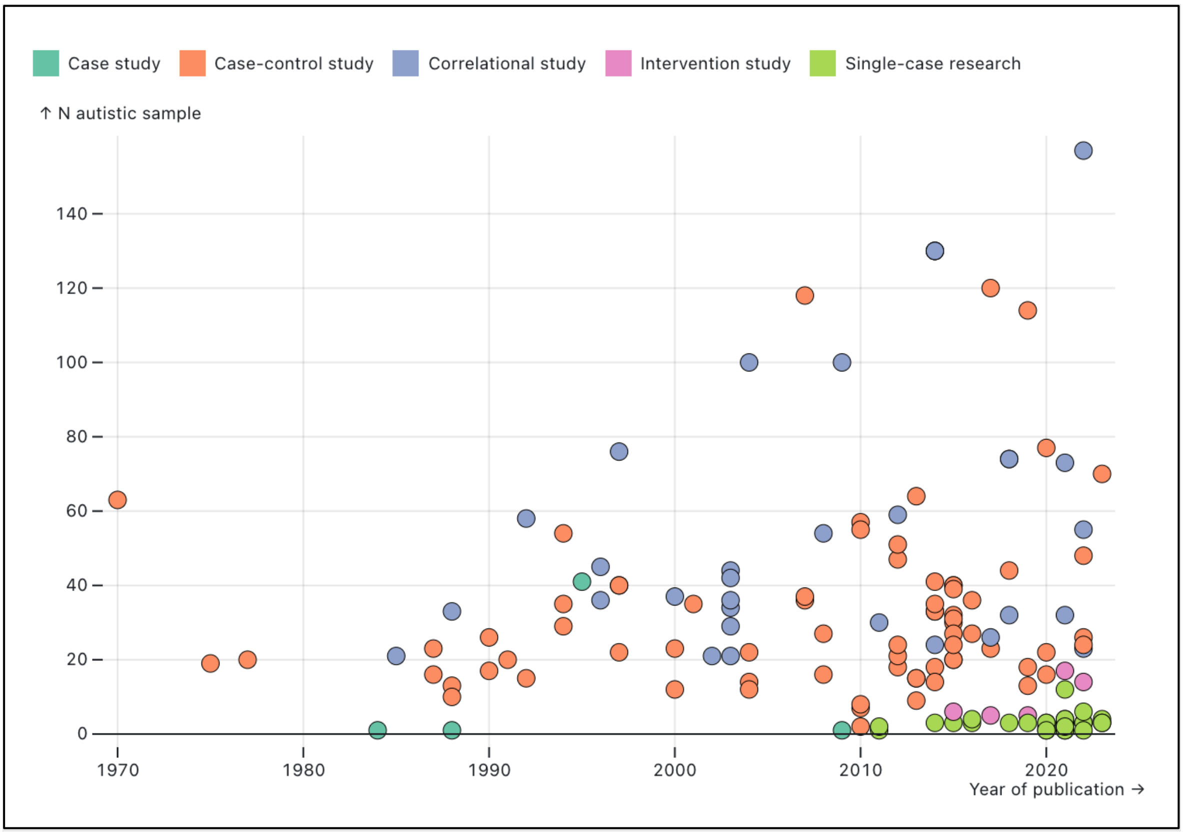
	Cross-sectional	Longitudinal	Total	
	n	n	n	%
Case-study	4	-	4	4
Correlational study	21	2	23	20
Case-control study	51	2	53	46
Intervention study	6	-	6	5
Single-case research	28	-	28	25

Note: N = 114.

The studies collected data across 14 countries. Most of the studies (n = 87, 76%) were run in the United States of America or the United Kingdom (n = 12, 11%).

Figure 4.2 shows the distribution of the 143 autistic samples plotted by year of publication of the study. The included studies were published between 1970 and 2023 and the frequency of the publications has increased over the years. In particular, the distribution of studies is skewed towards the present day, with the first quartile of articles published in 34 years, the second quartile in 10 years, and the remaining 50% of the included studies published in the last 9 years. Moreover, this chart shows that single-case research and intervention studies investigating mathematical abilities in autism have been published only after 2012, and since then single-case research has gained increased adoption in the field.

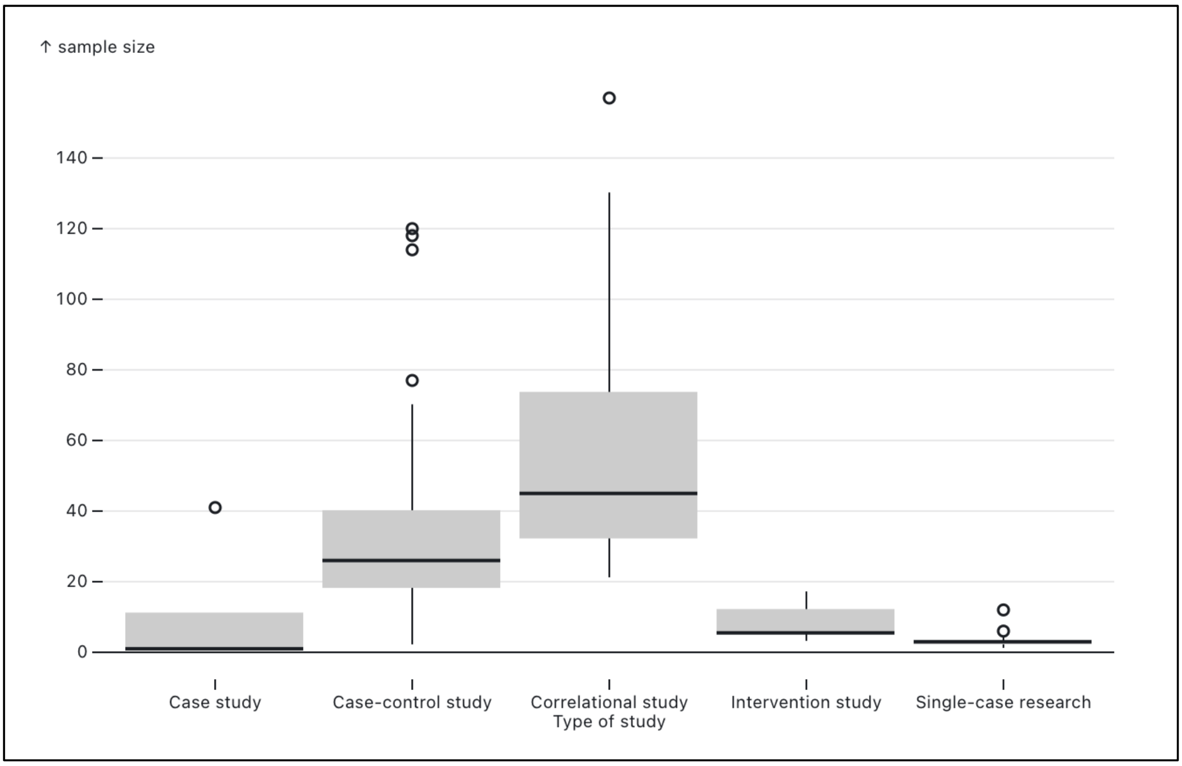
Figure 4.2. Autistic samples plotted by publication year and size and grouped by type of study.



Note: *N* = 143.

As Figure 4.3 shows, the sample size varied considerably across the studies (range: 1 to 157). The figure illustrates the distribution of the size of the autistic samples for different types of study. Across the 143 samples, single case research (mean = 3.00, median = 3.00, SD = 2.13, range: 1 to 12), case studies (mean = 11.00, median = 1.00, SD = 20.00, range: 1 to 41) and intervention studies (mean = 8.33, median = 5.50, SD = 5.72, range: 3 to 17) were characterised by a consistently smaller and less variable number of participants than case-control studies (mean = 32.36, median = 26.00, SD = 23.54, range: 2 to 120), and correlational studies (mean = 57.38, median = 43.00, SD = 37.54, range: 21 to 157).

Figure 4.3. Sample size of the autistic samples grouped by type of study.



Note: N = 143.

4.3.2 Characteristics of the participants

4.3.2.1 Autistic traits

All the studies reported that participants' diagnosis of autism had been confirmed by either a clinical or a parental report. When the diagnosis of autism was assessed using a validated rating scale (*n* = 76 samples, 53%), the diagnostic and screening tools that were used to measure autistic traits varied greatly. Most of the studies used either the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2012) (*n* = 44, 58%) or the Autism Diagnostic Interview (ADI; Rutter et al., 2003) (*n* = 33, 43%). The thresholds applied by different studies when using these tools were not consistent and the measures recorded by different tools could not be compared with each other. This prevented a comparison of autistic traits scores across different autistic samples and, consequently, did not allow the investigation into whether the severity of autistic traits influence mathematical skills. Hence, contrary to the original plan, the RQ around whether the severity of autistic traits influence mathematical skills could not be addressed.

4.3.2.2 Sex ratio

The proportion of male and female participants for the autistic sample was reported for 127 samples (88%). The proportion of females across the 127 samples varied significantly, ranging from no females to all females (min: 0%; max: 100%). On average the percentage of females in the autistic samples was relatively low at 18%, with a quarter of the studies having less than 4% of female participants, half of the studies having less than 15% of female participants, and three quarters of the studies having less than 23% female representation in their samples.

4.3.2.3 Mean CA

Mean CA data for the autistic samples were available for most of the samples (*n* = 128; 90%). The distribution of mean CA is shown in Figure 4.4. Two studies provided mean CA for the whole autistic sample rather than for the two sub-groups that were being compared (Goldstein et al., 1994; Spek et al., 2008). The remaining studies (*n* = 7, including 13 samples) provided age ranges. The mean CA of the autistic samples was 12.75 years (min: 4.63, max: 41.93), with 25% of the studies having autistic participants younger than 9 years old , 50% having autistic participants younger than 11 years old, and 75% having participants not older than 14 years in their autistic samples. The seven studies that reported CA as a range replicated this pattern, with the minimum CA being 3 years (Mayes & Calhoun, 2003a, 2003b; Myles et al., 1994) and the maximum CA 18 years (Myles et al., 1994).

4.3.2.4 Mean IQ scores

IQ measures were collected using a variety of tests. Table 4.2 shows the count and percentages of the studies which reported FSIQ, VIQ and NVIQ scores to describe their autistic sample. IQ scores of the autistic sample were not reported for 29 samples (20%). The most employed test was the Wechsler IQ test, in multiple versions, which accounted for 88 cases, representing 62% of the total of the samples included in the systematic review.

Table 4.2. List of assessments used to measure IQ scores.

Assessment	*n*	%
Wechsler IQ test (Wechsler, multiple versions)	88	62
IQ scores not reported	29	20
Mixed assessments (2 or 3 from the list)	9	6
Stanford–Binet (multiple versions)	4	3
Raven's coloured progressive matrices (RCPM; Raven, 1993)	4	3
Differential ability scale (DAS; Elliot, 1990)	3	2
Kaufman brief intelligence test (Kaufman & Kaufman, 2004)	2	1
Woodcock-Johnson Test of Cognitive Abilities (WJ; Woodcock et al., multiple versions)	2	1
Universal nonverbal intelligence test	1	1
British picture vocabulary (BPVS; Dunn et al., 1982)	1	1

Note: *N* = 143.

Table 4.3 shows the number of studies which reported IQ scores, with FSIQ having the highest number of reports (*n* = 98, 69%), followed by NVIQ (*n* = 83, 58%) and VIQ (*n* = 81, 57%). The quartile and the minimum and maximum values provide an insight into the distribution of the mean IQ scores of the samples investigated in the included studies. For all IQ measures, less than 25% of the samples reported mean IQ scores under the threshold used in this study to define ID (IQ < 75).

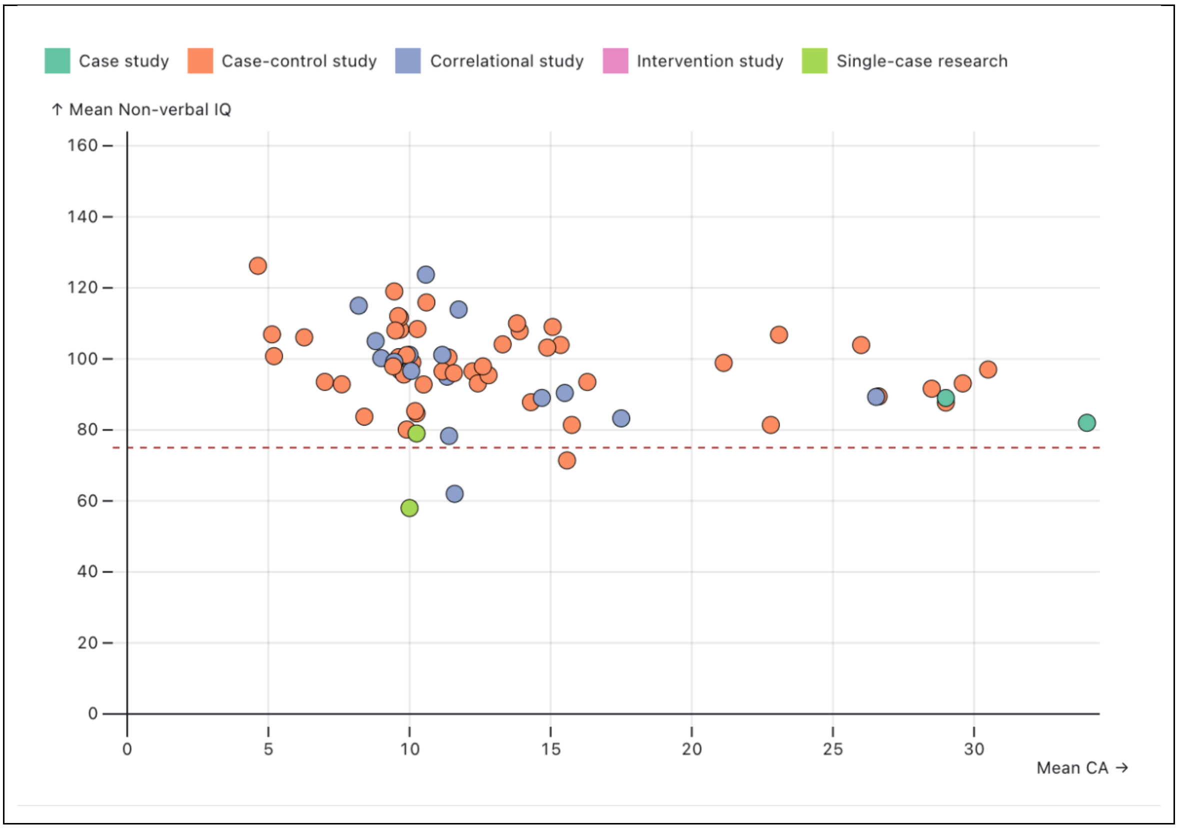
Table 4.3. Count, quartiles, and range of IQ scores reported for the autistic samples.

IQ measure	*n*	%	Q1	Q2	Q3	Range
FSIQ	98	69	79.40	96.78	104.15	43.00 – 160.00
VIQ	81	57	91.31	99.40	105.88	56.70 – 130.60
NVIQ	83	58	89.34	97.90	104.78	58.00 – 126.21

Note: *N* = 143. FSIQ = Full scale IQ. VIQ = Verbal IQ. NVIQ = Non-verbal IQ.

Figure 4.4 visually represents the distribution of mean IQ scores and mean CA of all the samples which reported both data. These charts show that the three measures of IQ present a similar relationship with mean CA and that only a few studies reported mean IQ scores below the threshold for ID. This may be explained by the fact that, in most studies (*n* = 78, 54%), the inclusion criteria required IQ scores higher than 70.

Figure 4.4. Association between mean IQ scores and mean CA for autistic samples by study design.



Note: n FSIQ vs CA = 86; n VIQ vs CA = 68; n NVIQ vs CA = 70. Standard Scores (M = 100; SD = 15). The horizontal red dotted line y = 75 represents the threshold used in the current study to define ID.

4.3.3 Mathematical performance

Mathematical performance was reported differently in different studies. It took the form of standard scores (*n* = 114, 76%), grade-equivalent scores (*n* = 9, 6%), age-equivalent scores (*n* = 1, < 1%), percentile scores (*n* = 8, 5%), raw scores (*n* = 12, 7%), and z-scores (*n* = 9, 6%). One study (Mayes & Calhoun, 2007) only reported the results of the analyses performed and omitted to report the scores of mathematical outcomes.

Most of the samples included in the systematic review reported the measure for a single mathematical basic process or component and assessed such mathematical component using only one assessment (*n* = 109; 70%).

Table 4.4 reports the frequency of the studies investigating different basic processes and specific components of mathematics based on the multi-level framework described in Chapter 1. Most of the studies investigated the specific components of arithmetic word problems (*n* = 84; 38%) and calculation (*n* = 61; 29%).

Table 4.4. Measures of mathematical outcomes reported for the autistic samples.

Mathematical process / component	*n*	%
Arithmetic word problems	84	38
Calculation	61	27
Overall mathematics achievement	52	24
Enumeration	10	5
Magnitude comparison	8	4
Number line	5	2

Figure 4.5 shows the distribution of mean CA of the samples included in the systematic review categorised by the mathematical components being measured.

The chart shows that the median mean CA was similar for most of the mathematical components (arithmetic word problems: 11.25 years, calculation: 9.96 years, overall maths achievement: 10.00 years, enumeration: 9.50 years) except for magnitude comparison and number line components which were assessed in slightly younger samples (magnitude comparison: 6.27 years; number line: 6.50 years).

Mathematical abilities were investigated not earlier than 4 years, with the minimum mean CA (4.63 years) reported in the studies investigating overall mathematics achievement and magnitude comparison skills. This is not surprising, as the average age of autism diagnosis is ~4–5 years of age (Lord et al., 2020). A few outliers in the data indicated that some studies investigating arithmetic word problems, calculation, overall mathematics achievement, and magnitude comparison assessed mathematical abilities of adult participants.

Studies investigating arithmetic word problems reported samples with the wider mean CA range (35.57 years) while studies examining number line skills reported samples with a narrower mean CA range (4.01 years). This is likely due to the small number of studies using this task with the autistic population (n = 5).

Figure 4.5. Mean CA of the autistic samples by mathematical component.

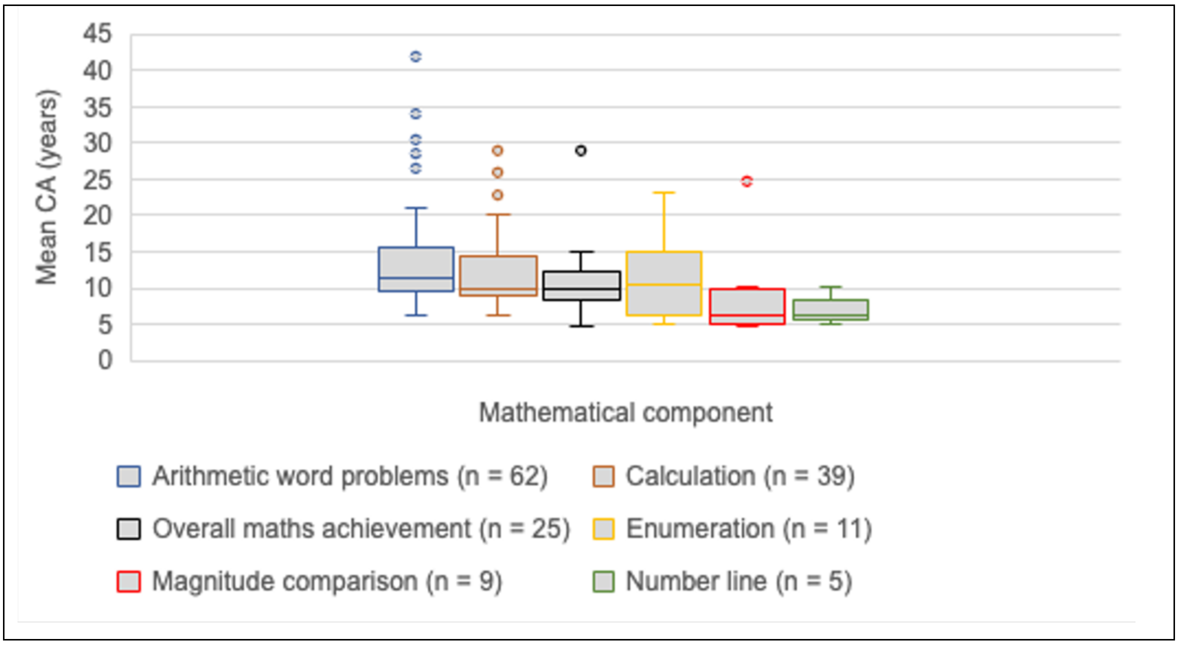
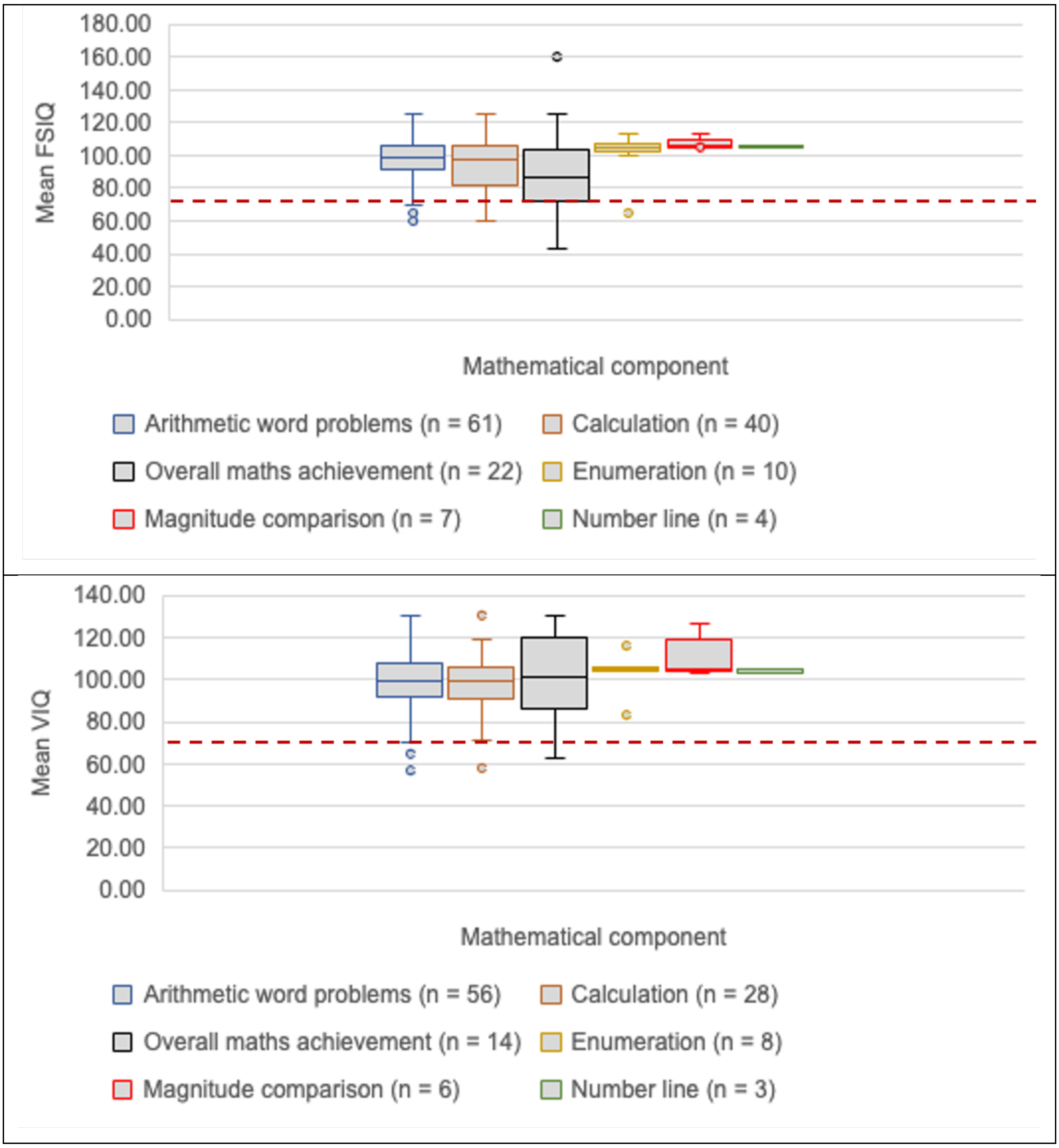


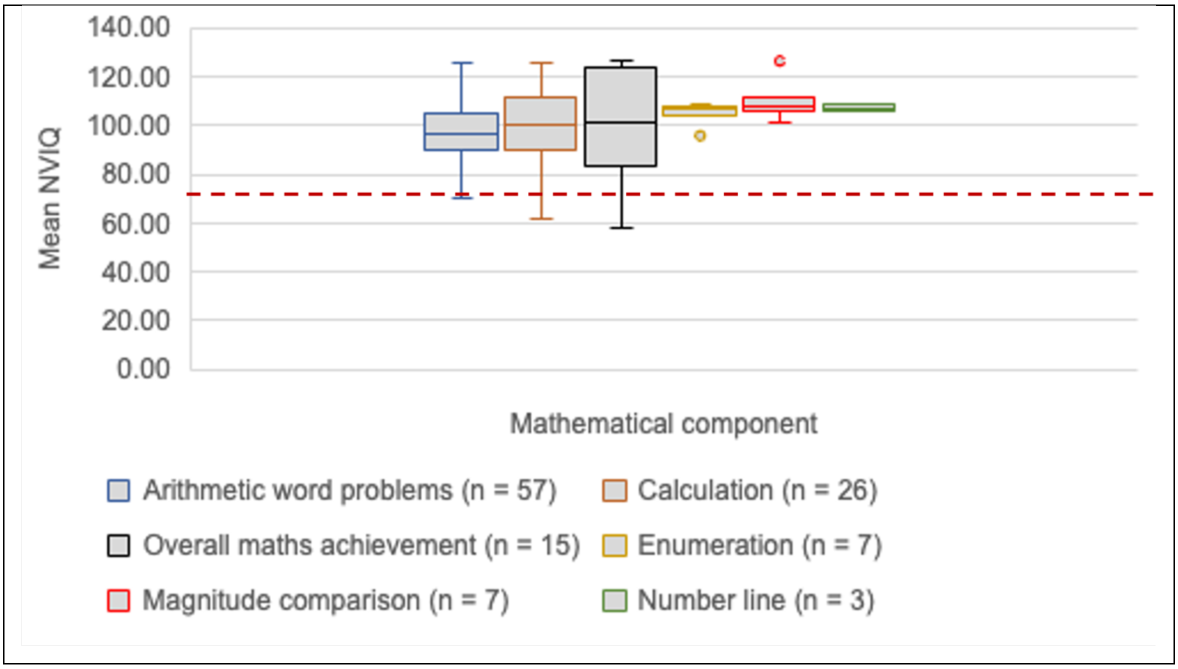
Figure 4.6 shows the distribution of mean FSIQ, mean VIQ and mean NVIQ of the autistic samples included in the systematic review categorised by mathematical components.

With the only exception of the samples which reported measure of overall mathematics achievement (median mean FSIQ = 86.11), the median mean FSIQ, VIQ and NVIQ for all mathematical components were close to or higher than the standardised mean score of 100, indicating that most of the studies investigated samples with mean IQ within the typical range.

While for overall mathematics achievement, 25% of the samples reported mean FSIQ (but not VIQ and NVIQ) below the threshold of ID, only a smaller percentage of samples of studies investigating arithmetic word problems and calculation fall below the ID threshold of 75 points. The distributions of mean IQ scores of the studies reporting measures of enumeration, magnitude comparison and number line skills fall above the ID threshold, with the exception of the case-control study by Kirk et al. (2017), which assessed counting abilities of autistic participants.

Figure 4.6. Mean IQ scores of the autistic samples by mathematical component.





Note: Standard scores (M = 100; SD = 15). The horizontal red dotted line y = 75 represents the threshold used in the current study to define ID.

4.3.3.1 Arithmetic word problems

Overview. Fifty-seven studies assessed arithmetic word problem skills across 78 samples and reported 84 outcome measures. These included two case studies, 12 correlational studies, 36 case-control studies and seven single-case research studies. The included studies used nine different assessments to measure this ability. The most used test was the Arithmetic scale of the Wechsler IQ test (*n* = 38, 45%; Table 4.5).

Table 4.5. Assessment tools used to measure arithmetic word problems.

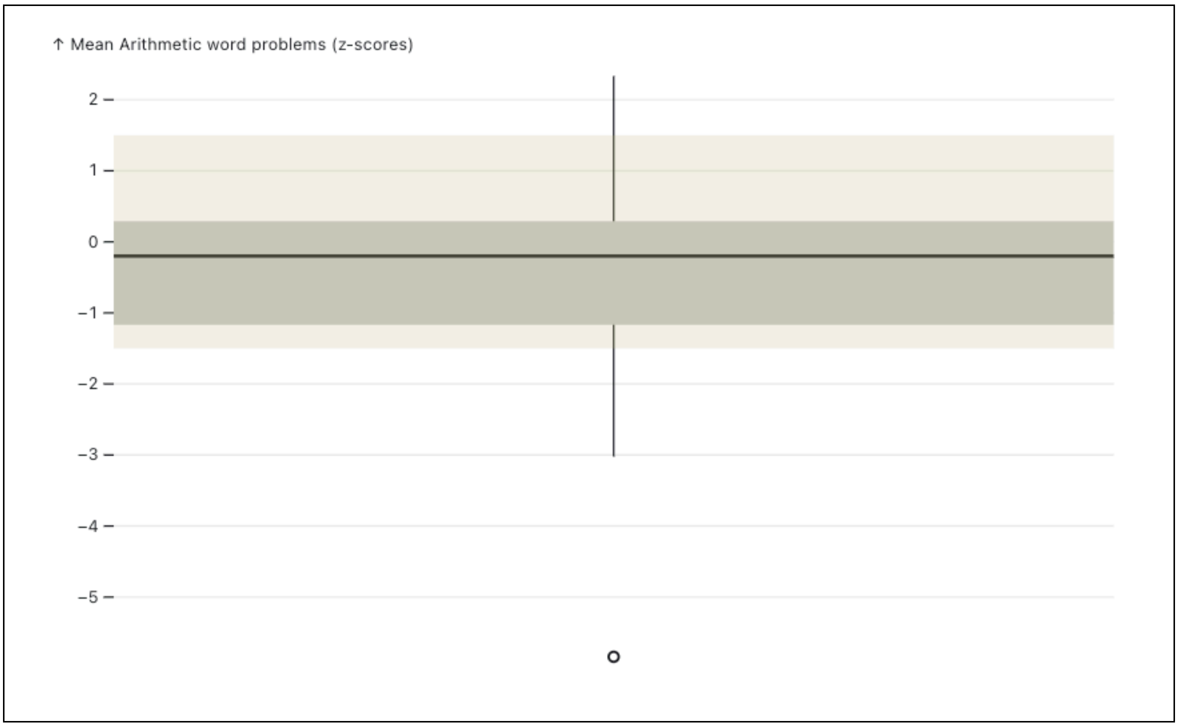
Assessment	Scale	*n*	%
Wechsler (Wechsler, multiple versions)	Arithmetic	38	45
Wechsler individual achievement test (Wechsler, multiple versions)	Mathematical reasoning (WIAT-MR)	15	18
Woodcock–Johnson test of achievement (Woodcock et al., multiple versions)	Applied problems (WJ-AP)	11	13
Cognitive developmental skills in arithmetic (Desoete & Roeyers, 2006)	Linguistic Mental representation Contextual (CDR-LMC)	5	6
Test of mathematical abilities (Brown et al., multiple versions)	Story problem (TOMA-SP)	5	6
Kaufman test of educational achievement (Kaufman & Kaufman, 1985)	Math applications (KTEA-MA)	4	5
Test for the diagnosis of mathematical competencies (Gregoire, Noel, & Van Nieuwenhoven, 2004)	Subtest 5.1 (TEDI-MATH 5)	3	4
Mathematical word problem solving (MWPS; Griffin & Jitendra 2009)	n/a	1	1

Mathematical problem instrument (MPI; Mulligan & Mitchelmore, 1997)	n/a	1	1
Mixed assessments (WIAT and WJ)	n/a	1	1

Note: *n* = 84. n/a: not applicable.

Figure 4.7 illustrates how the mean performance in arithmetic word problem skills is distributed across the autistic samples for which standardised scores were reported (*n* = 70). The boxplot shows an unbalanced distribution of the data around the median, with a skew towards lower z-scores. The spread of the middle 50% of the samples reported a performance within the typical range, with the median of the mean scores of the autistic samples slightly below the standardised mean *z*-score of 0. The whiskers extend over the typical range, with maximum score at just over 2 standard deviations and minimum score at -3 standard deviations. One outlier is plotted more than 5 standard deviations below the standardised mean. This represents a sample from the study by Kurth and Mastergeorge (2010) made of autistic students with a mean FSIQ score below the ID threshold. Overall, 89% of the samples z-scores (*n* = 62) lied above the lower end of the typical range. The remaining eight samples reported mean FSIQ below the score of 75.42.

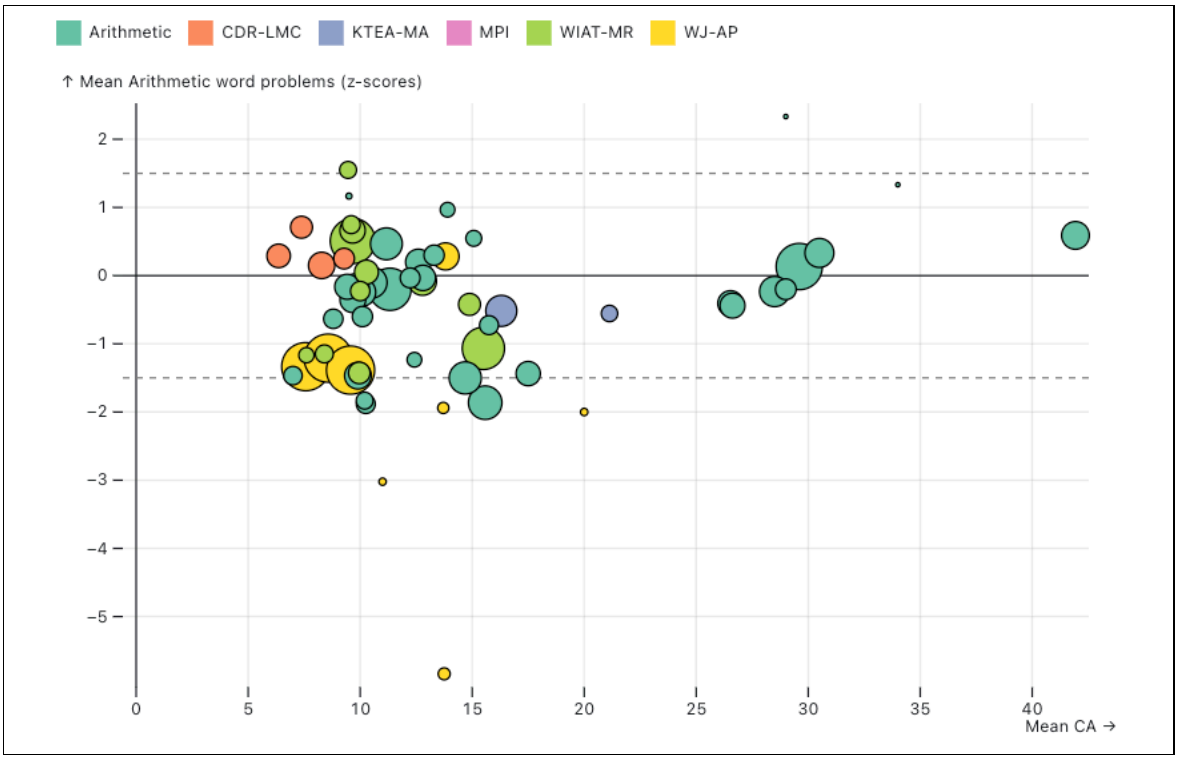
Figure 4.7. Distribution of mean arithmetic word problem *z*-scores.



Note: *n* = 70. The light grey section between -1.5 and 1.5 standard deviations from the standardised mean shows the range usually considered “typical” in mathematical cognition within the clinical and the educational contexts (e.g., in Santos et al., 2022).

Relationship with CA. Figure 4.8 shows the distribution of mean arithmetic word problem scores (transformed into *z*-scores) plotted against CA. Arithmetic word problem skills were investigated in studies including autistic samples with a wide mean CA range (average mean CA = 14.30 years; mean CA range = 5.13 - 41.93). Visual analysis of the chart shows that most of the mean z-scores lie within 1.5 standard deviations of the standardised mean. Furthermore, the mean scores from samples with higher mean CA are close to or above the standardised mean score (z = 0) and lie within a narrower range of scores than younger samples.

Figure 4.8. Association between mean arithmetic word problem *z*-scores and mean CA by assessment tool.

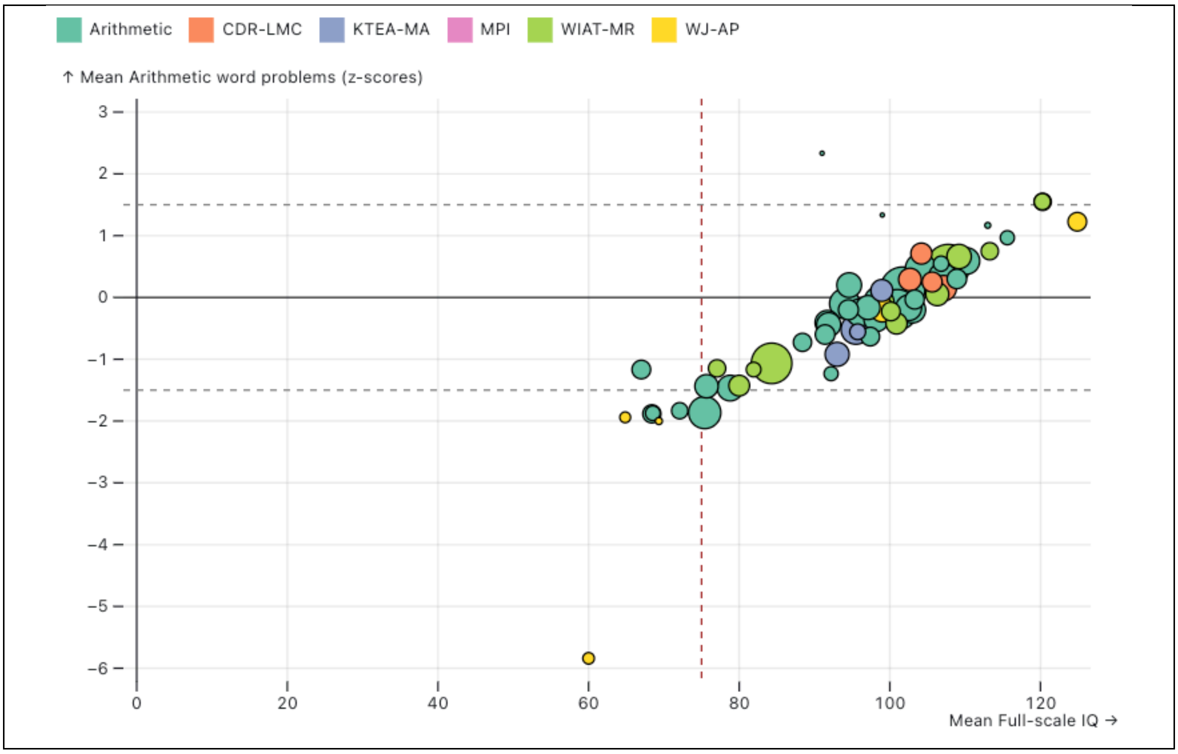


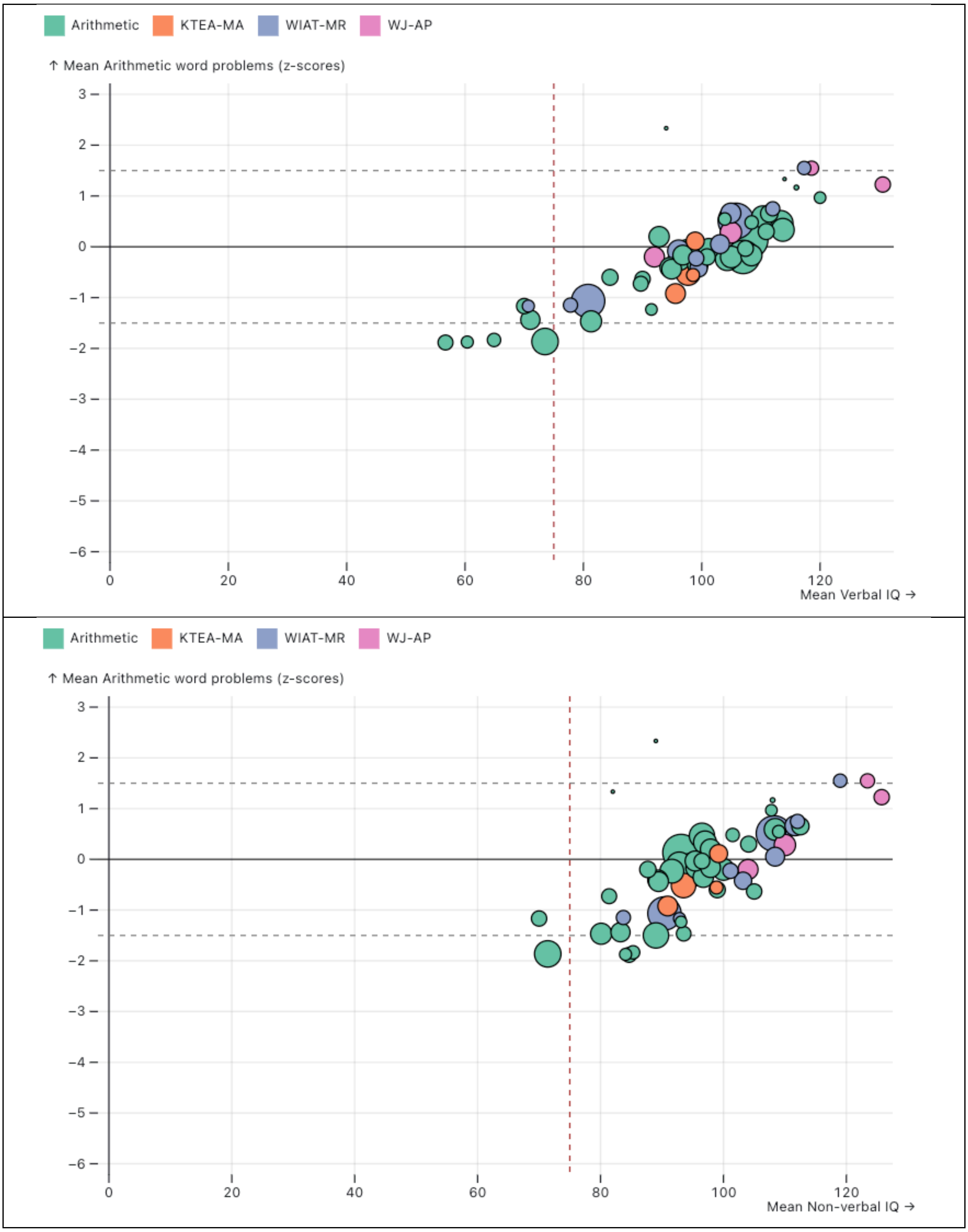
Note: *n* = 62. CDR-LMC = Cognitive developmental skills in arithmetic – Linguistic scale, Mental representation scale and Contextual scale. KTEA-MA = Kaufman test of educational achievement – Math applications scale. MPI = Mathematical problem instrument. WIAT-MR = Wechsler Individual Achievement Test – Mathematical reasoning scale. WJ-AP = Woodcock–Johnson test of achievement – Applied problems scale. The horizontal solid black line y = 0 represents the standardised mean. The horizontal grey dotted lines y = -1.5 and y = 1.5 encompass the range usually considered “typical” in mathematical cognition within the clinical and the educational contexts (e.g., in Santos et al., 2022). The dot size is scaled based on the sample size.

Relationship with FSIQ, VIQ and NVIQ. Figure 4.9 shows the distribution of mean arithmetic word problem scores (transformed into *z*-scores) plotted against FSIQ, VIQ, and NVIQ. Arithmetic word problem skills were investigated in studies including autistic samples with high mean FSIQ (average mean FSIQ = 95.63; mean FSIQ range = 60.00 - 124.89). Visual inspection of the charts shows a similar pattern for all IQ measures, where higher arithmetic word problem scores correspond to higher mean IQ scores. All but one of the seven samples with mean FSIQ lower than the ID threshold lie below the lower end of the typical range.

To investigate how performance on arithmetic word problem tasks relates to FSIQ, VIQ, and NVIQ, separate Spearman's correlations were calculated between the mean score of mathematical performance and the mean score of IQ measures. This showed statistically significant results and indicated a strong and positive association between arithmetic word problem scores and FSIQ, VIQ, and NVIQ scores; FSIQ: $rs(62) = .87$, $p < .001$; VIQ: $rs(56) = .85$, $p < .001$; NVIQ: $rs(57) = .69$, $p < .001$. This further supports the trend observed in the chart, indicating that arithmetic word problem scores in autism develop in conjunction with IQ measures.

Figure 4.9. Association between mean arithmetic word problems *z*-scores and mean IQ scores by assessment tool.





Note: *n* FSIQ = 62; *n* VIQ = 56; *n* NVIQ = 57. KTEA-MA = Kaufman test of educational achievement – Math applications scale. CDR-LMC = Cognitive developmental skills in arithmetic – Linguistic scale, Mental representation scale and Contextual scale. MPI = Mathematical problem instrument. WIAT-MR = Wechsler Individual Achievement Test – Mathematical reasoning scale. WJ-AP = Woodcock–Johnson test of achievement – Applied problems scale. The vertical red dotted line x = 75 represents the threshold used in the current study to define ID. The horizontal solid black line y = 0 represents the standardised mean. The horizontal grey dotted lines y = -1.5 and y = 1.5 encompass the range usually considered "typical" in mathematical cognition within the clinical and the educational contexts (e.g., in Santos et al., 2022). The dot size is scaled based on the sample size.

Case-control studies with a TD control group. To examine how arithmetic word problem abilities in autism compare to TD population, results from case-control studies were examined. Out of the 36 case-control studies investigating arithmetic word problem skills, 23 studies compared 29 autistic samples with a TD group, for a total of 30 measures. This subgroup of samples included relatively young autistic participants with a wide mean CA range (average mean CA = 10.32 years; mean CA range = 5.13 - 21.13) and high mean FSIQ scores (average mean FSIQ = 102.95; mean FSIQ range = 72.10 - 120.25). The variability in FSIQ scores within each sample (shown by the high values of SD for mean FSIQ scores), along with the indication that the lowest reported mean FSIQ falls below the ID threshold, indicates the presence of individuals with ID in some of these samples. These studies measured arithmetic word problem skills using a range of different assessments, that included KTEA-MA (*n* = 4), MPI (*n* = 1), MWPS (*n* = 1), TEDI-MATH 5 (*n* = 3), TOMA-SP (*n* = 1), Arithmetic (*n* = 6), WIAT-MR (*n* = 8), CDRL-MC (*n* = 5), and WJ-AP (*n* = 1). Table 4.6 summarises the findings from these studies.

Table 4.6. Overview and interpretation of findings from case-control studies with a TD control group investigating arithmetic word problem abilities.

Article	*N*	Mean CA (*SD*)	Mean FSIQ (*SD*)	Matching criteria	Maths Measure	Reported Findings	Interpretation of findings
Aagten-Murphy et al. (2015)	32	10.28 (1.30)	106.30 (9.60)	CA, FSIQ (Wechsler)	WIAT-MR	AG scored significantly lower than TDG; F (1, 63) = 4.62, p = .040.	Not in line with CA and not in line with MA (FSIQ).
Bae et al. (2015)	20	10.60 (0.94)	109.60 (15.85)	CA, FSIQ (KBIT)	MWPS	AG scored significantly lower than TDG; F (1, 38) = 3.66, p = .001.	Not in line with CA and not in line with MA (FSIQ).
Bae et al. (2015)	20	10.60 (0.94)	109.60 (15.85)	CA, FSIQ (KBIT)	TOMA-SP	AG scored significantly lower than TDG; F (1, 38) = 2.92, p = .006.	Not in line with CA and not in line with MA (FSIQ).
Bullen et al. (2020)	77	11.38 (2.20)	98.60 (n/r)	CA	WIAT-MR	AG scored significantly lower than TDG; p < .05.	Not in line with CA.
Chen et al. (2019)	96	9.67 (1.49)	107.75 (17.91)	CA, FSIQ (Wechsler)	WIAT-MR	AG scored significantly lower than TDG; p = .032.	Not in line with CA and not in line with MA (FSIQ).
Gagnon et al. (2004)	14	15.07 (3.04)	106.80 (9.80)	CA, sex, FSIQ (Wechsler)	Arithmetic	No statistically significant differences found between AG and TDG; p not reported.	In line with CA and in line with MA (FSIQ).
Goldstein et al. (1994)	29	Younger than 13 y	98.93 (12.48)	CA, sex, FSIQ (Wechsler), SES	KTEA-MA	No statistically significant differences found between AG and TDG; p not reported.	In line with CA and in line with MA (FSIQ).
Goldstein et al. (1994)	29	Older than 13 y	93.00 (15.16)	CA, sex, FSIQ (Wechsler), SES	KTEA-MA	No statistically significant differences found between AG and TDG; p not reported.	In line with CA and in line with MA (FSIQ).
Hiniker et al. (2016)	36	9.66 (1.60)	109.19 (20.45)	CA, FSIQ (Wechsler)	WIAT-MR	No statistically significant differences found between AG and TDG; t = -.151, p = .881.	In line with CA and in line with MA (FSIQ).
Iuculano et al. (2014)	18	9.60 (1.64)	113.27 (15.25)	CA, sex, FSIQ (Wechsler)	WIAT-MR	No statistically significant differences found between AG and TDG; t = 1.520, p = .138.	In line with CA and in line with MA (FSIQ).
Iuculano et al. (2020)	16	9.46 (1.80)	120.25 (15.25)	CA, sex, FSIQ (Wechsler)	WIAT-MR	No statistically significant differences found between AG and TDG; p = .170.	In line with CA and in line with MA (FSIQ).
McCauley et al. (2018)	44	12.78 (2.10)	98.77 (14.21)	CA, school grade	WIAT-MR	AG scored significantly lower than TDG; p < .001.	Not in line with CA.
Minshew et al. (1992)	15	21.13 (8.02)	95.73 (13.61)	CA, sex, FSIQ (Wechsler), race	KTEA-MA	No statistically significant differences found between AG and TDG; p not reported.	In line with CA and in line with MA (FSIQ).
Minshew et al. (1994)	54	16.30 (10.16)	95.50 (15.54)	CA, sex, FSIQ (Wechsler), SES	KTEA-MA	No statistically significant differences found between AG and TDG; t = -1.80, p > .05.	In line with CA and in line with MA (FSIQ).

Ohta (1987)	16	10.20 (2.93)	72.10 (16.20)	CA, NVIQ (Wechsler)	Arithmetic	AG scored significantly lower than TDG; *p* < .05.	Not in line with CA and not in line with MA (NVIQ).
Oswald et al. (2016)	27	14.88 (1.68)	100.89 (11.10)	CA, FSIQ (Wechsler), NVIQ (Wechsler)	WIAT-MR	AG scored significantly lower than TDG; *p* = .002.	Not in line with CA and not in line with MA (FSIQ; NVIQ).
Ozonoff et al. (2000)	23	13.30 (3.90)	108.90 (13.80)	CA, FSIQ (Wechsler), VIQ (Wechsler), NVIQ (Wechsler)	Arithmetic	No statistically significant differences found between AG and TDG; *p* not reported.	In line with CA and in line with MA (FSIQ; VIQ; NVIQ).
Ozonoff et al. (2000)	12	13.90 (4.50)	115.60 (15.60)	CA, FSIQ (Wechsler), VIQ (Wechsler), NVIQ (Wechsler)	Arithmetic	No statistically significant differences found between AG and TDG; *p* not reported.	In line with CA and in line with MA (FSIQ; VIQ; NVIQ).
Polo-Blanco et al. (2024)	26	9.35 (2.06)	89.88 (11.78)	CA, sex, SES	MPI	No statistically significant differences found between AG and TDG; t (50) = 0.44; p = .339, d = −0.27.	In line with CA.
Soulieres et al. (2010)	2	9.50 (n/r)	113.00 (2.80)	CA, sex, FSIQ	Arithmetic	AG reported “similar scores” to the TDG.	n/a – statistical analyses not reported.
Titeca et al. (2014)	33	6.27 (0.38)	105.38 (13.27)	FSIQ (Wechsler), SES, and sex	TEDI-MATH 5	No statistically significant differences found between AG and TDG; F (5, 81) = 1.17, p = .330.	In line with MA (FSIQ).
Titeca et al. (2014)	33	6.87 (0.29)	n/r	FSIQ (Wechsler), SES, sex	CDR-LMC	AG scored significantly lower than TDG; F (1, 83) = 8.18, p = .005.	Not in line MA (FSIQ).
Titeca, Roeyers, Ceulemans, et al. (2015)	30	5.98 (0.31)	104.83 (12.36)	CA, FSIQ (Wechsler), SES, and sex ratio	TEDI-MATH 5	No statistically significant differences found between AG and TDG; U = 449.50, p = .994.	In line with CA and in line with MA (FSIQ).
Titeca, Roeyers and Desoete (2015)	20	5.13 (0.33)	105.30 (13.90)	CA, FSIQ (Wechsler), SES, and sex ratio	TEDI-MATH 5	No statistically significant differences found between AG and TDG; U = 184.50, p = .678.	In line with CA and in line with MA (FSIQ).
Titeca, Roeyers, Loeys, et al. (2015)	31	6.36 (0.24)	102.67 (12.31)	Normed sample	CDR-LMC	AG reported similar scores to the normed sample; *p* not reported.	In line with CA.
Titeca, Roeyers, Loeys, et al. (2015)	27	7.38 (0.26)	104.19 (15.41)	Normed sample	CDR-LMC	AG scored significantly above the normed sample; *p* not reported.	Not in line with CA.

Titeca, Roeyers, Loeys, et al. (2015)	39	8.28 (0.31)	107.16 (13.80)	Normed sample	CDR-LMC	AG reported similar scores to the normed sample; *p* not reported.	In line with CA.
Titeca, Roeyers, Loeys, et al. (2015)	24	9.28 (0.30)	105.62 (13.57)	Normed sample	CDR-LMC	AG scored significantly above the normed sample; *p* not reported.	Not in line with CA.
Troyb et al. (2014)	41	13.81 (2.67)	104.93 (14.55) (*)	CA, sex, NVIQ (Wechsler)	WJ-AP	No statistically significant differences found between AG and TDG; *p* =.060.	In line with CA and in line with MA (NVIQ).
Zielińska et al. (2014)	35	9.42 (2.45)	97.10 (16.40)	CA, sex, FSIQ (Wechsler)	Arithmetic	No statistically significant differences found between AG and TDG; *p* not reported.	In line with CA and in line with MA (FSIQ).

Note. *n* = 30. (*) = VIQ was reported when the FSIQ was missing. (n/r) = not reported. n/a: not applicable. AG = Autistic group. TDG = Typically developing group. WIAT-MR = Wechsler Individual Achievement Test – Mathematical reasoning scale. MWPS = Mathematical word problem solving. TOMA-SP = Test of mathematical abilities – Story problem scale. KTEA-MA = Kaufman test of educational achievement – Math applications scale. MPI = Mathematical problem instrument. TEDI-MATH 5 = Test for the diagnosis of mathematical competencies – Sub-test 5.1. CDR-LMC = Cognitive developmental skills in arithmetic – Linguistic scale, Mental representation scale and Contextual scale. WJ-AP = Woodcock–Johnson test of achievement – Applied problems scale.

The interpretation of the findings through the neuroconstructivist perspective shows that 14 samples (74%) out of the 19 using a TD control group matched for both CA and MA (FSIQ) reported autistic performance in line with the TD control group. Three samples out of the seven (43%) using a TD control group matched only for CA reported autistic performance in line with the TD control group. One sample out of the two (50%) using a TD control group matched only for MA (FSIQ) reported autistic performance in line with the TD control group.

Out of the 10 samples which reported statistically significant differences with the TD control group, only two samples (20%) reported higher arithmetic word problem scores for the autistic population. Both these samples come from the same study (Titeca, Roeyers, Loeys, et al., 2015) and featured autistic samples with high mean FSIQ scores. The vast range of assessments used in the studies and the small number of available data points with the same assessment and outcome made it not possible to analyse the relationship between arithmetic word problem outcomes and the type of assessment used.

<u>Summary.</u> In summary, arithmetic word problem skills were investigated in samples with a wide mean CA range (average mean CA = 14.30 years; mean CA range = 5.13 - 41.93) and high mean FSIQ (average FSIQ = 95.63; mean FSIQ range = 60.00 - 124.89).

The average performance of most of the autistic samples on standardised assessment falls within the typical range. The samples lying below the lower end of the typical range (*n* = 7) are characterised by mean FSIQ scores below the ID threshold.

The analysis of the developmental trajectory of standardised scores from cross-sectional samples shows that arithmetic word problem skills develop in line with the normed population and that scores from older samples vary less than younger samples.

Arithmetic word problem skills are positively correlated with IQ scores.

Findings from case-control studies with a TD control group are mixed. However, when the performance of the autistic sample was not in line with the TD control group, most of the studies reported lower scores for the autistic sample.

4.3.3.2 Calculation

Overview. Thirty-seven studies assessed calculation abilities across 50 samples and reported 61 outcome measures. Within the 37 studies, there were one case study, nine correlational studies, 19 case-control studies and eight single-case research studies. The included studies used eight different assessments to measure calculation abilities, and the most used test was the Numerical Operations scale of the WIAT (*n* = 20; 33%, Table 4.7).

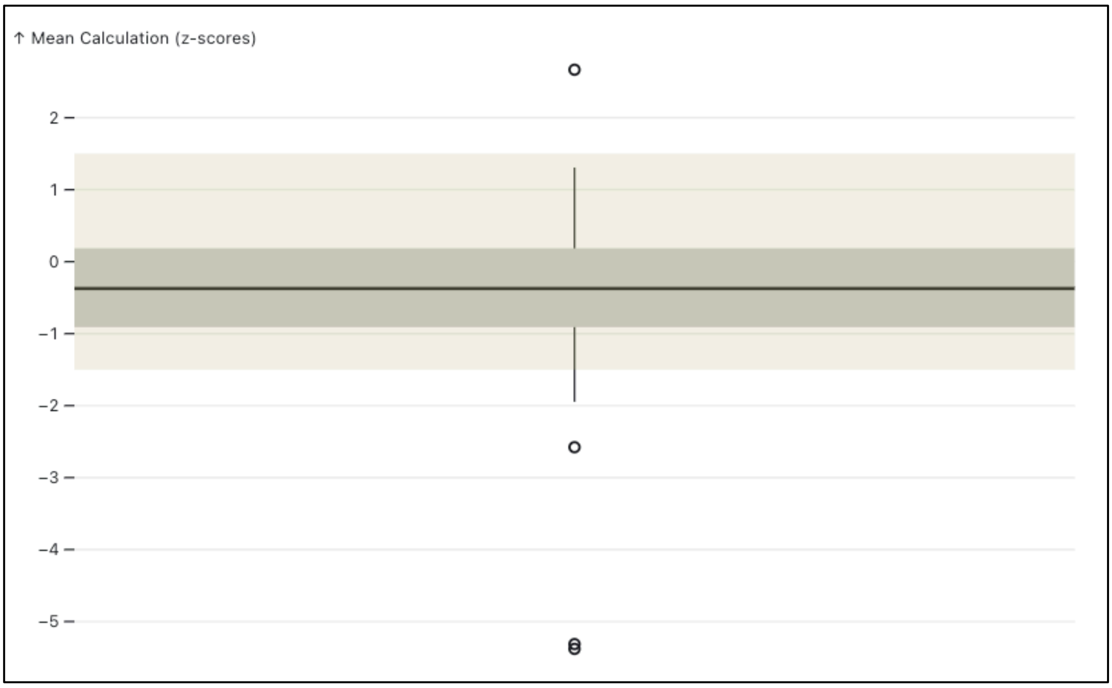
Table 4.7. Assessment tools used to assess calculation abilities.

Assessment	Scale	*N*	%
Wechsler individual achievement test (Wechsler, multiple versions)	Numerical operations (WIAT-NO)	20	33
Woodcock–Johnson test of achievement (Woodcock et al., multiple versions)	Calculation (WJ-C)	9	15
Wide range achievement test (Jastak & Wilkinson, multiple versions)	Computations (WRAT)	9	15
Arithmetic number facts test (TTR; De Vos, 1992)	n/a	5	8
Test of mathematical abilities (Brown et al. multiple versions)	Computation (TOMA-C)	5	8
Woodcock–Johnson test of achievement (Woodcock et al., multiple versions)	Fluency (WJ-F)	4	6
Cognitive developmental skills in arithmetic (Desoete & Roeyers, 2006)	Procedural calculation (CDR-P)	4	6
Kaufman test of educational achievement (Kaufman & Kaufman, 1985)	Math Computation (KTEA-MC)	3	5
Mixed assessments (WIAT and WJ)	n/a	2	4

Note: *n* = 61. n/a: not applicable.

Figure 4.10 illustrates how the mean performance in calculation abilities is distributed across the autistic samples for which standardised scores were reported (*n* = 47). The boxplot symmetry indicates a balanced distribution of the data around the median. The spread of the middle 50% of the samples reported a performance within the typical range, with the median of the mean scores of the autistic samples slightly below the standardised mean *z*-score of 0. The lower whisker extends over the lower end of the typical range, reaching a minimum z-score around 2 standard deviations below the standardised mean. There are four outliers, one above the typical range and three below it, reporting scores as low as -5 standard deviations. The outlier lying above the typical range represents a sample without ID (mean FSIQ = 91.00) from the study by Steel et al. (1984). The three outliers lying below the typical range represent two samples with ID (mean FSIQ < 65) from the study by Kurth and Mastergeorge (2010) and one sample for which no IQ scores were reported (Root et al., 2018). Overall, 89% of the samples z-scores (*n* = 42) lied above the lower end of the typical range. The remaining five samples reported mean FSIQ below the score of 72.7.

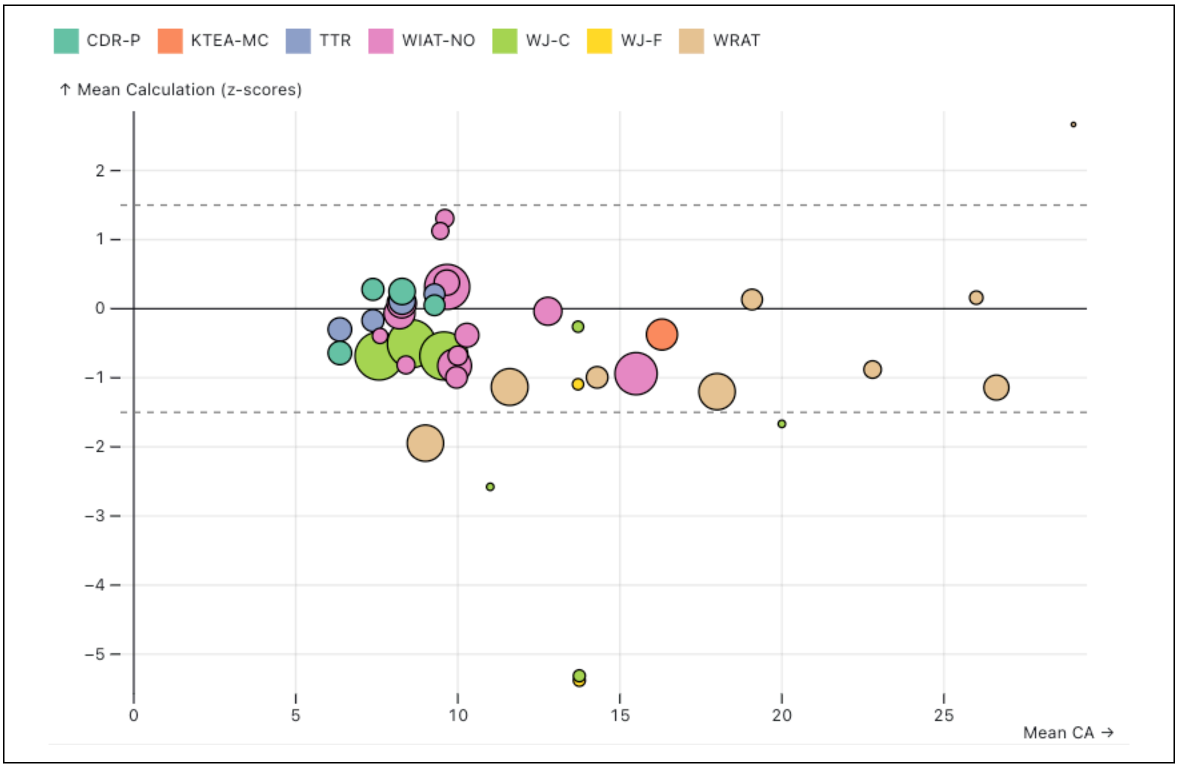
Figure 4.10. Distribution of calculation *z*-scores.



Note: *n* = 47. The light grey section between -1.5 and 1.5 standard deviations from the standardised mean shows the range usually considered "typical" in mathematical cognition within the clinical and the educational contexts (e.g., in Santos et al., 2022).

<u>Relationship with CA.</u> Figure 4.11 shows the distribution of mean calculation scores (transformed into *z*-scores) plotted against CA. Calculation skills were investigated in studies including autistic samples with a wide mean CA range (average mean CA = 12.71 years; mean CA range = 6.36 - 29.00). Visual examination of the chart shows that most of the mean *z*-scores lie within 1.5 standard deviations of the standardised mean. The mean scores from the samples with higher mean CA (mean CA > 10 years) fall very close to or below the expected mean score (*z* = 0), while more variability is observed in younger samples.

Figure 4.11. Association between mean calculation *z*-scores and mean CA by assessment tool.

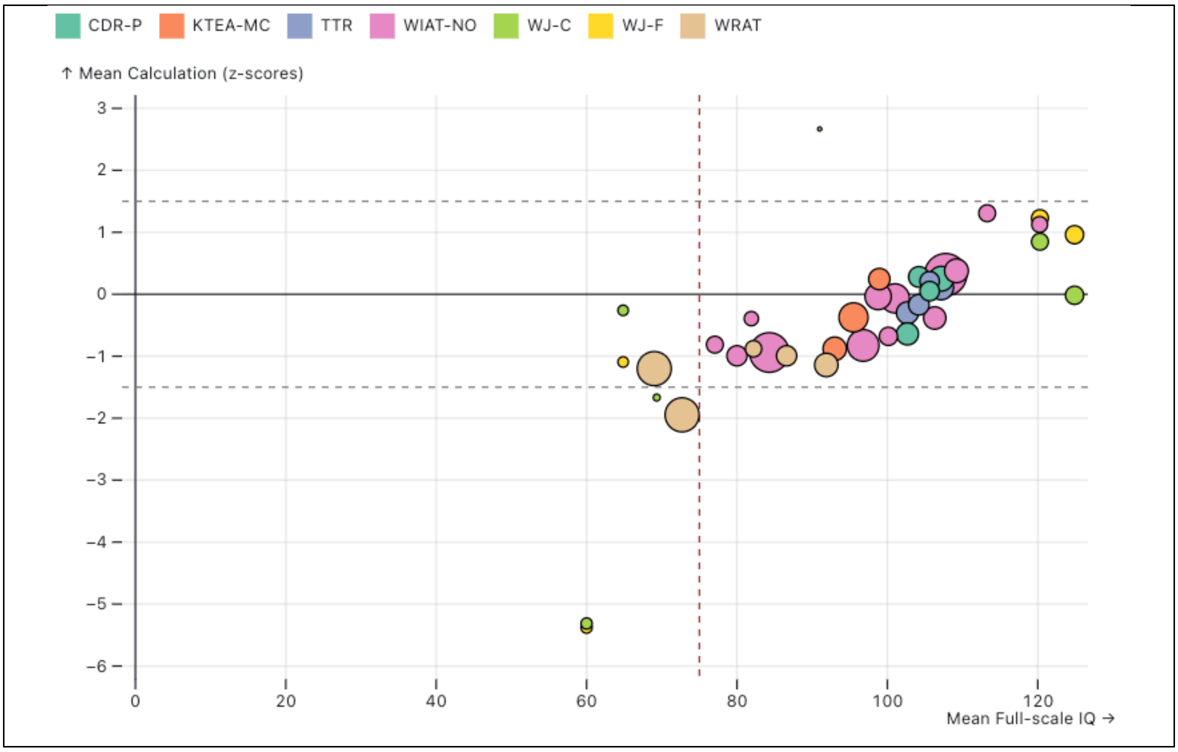


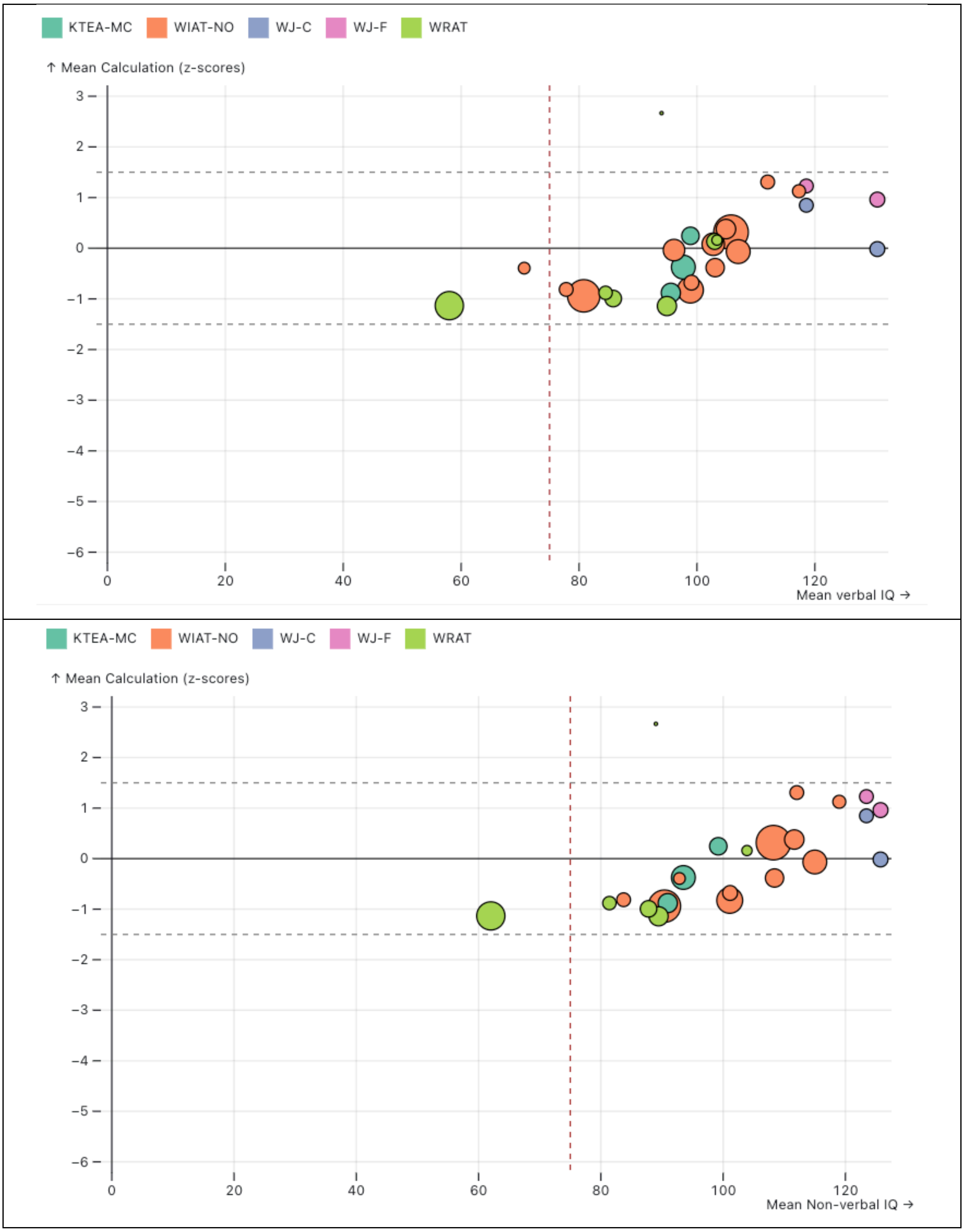
Note: *n* = 42. CDR-P = Cognitive developmental skills in arithmetic – Procedural calculation scale. KTEA-MC = Kaufman test of educational achievement – Math applications scale. TTR = Arithmetic number facts test. WIAT-NO = Wechsler Individual Achievement Test – Numerical operation scale. WJ-C = Woodcock–Johnson test of achievement – Calculation scale. WJ-F = Woodcock–Johnson test of achievement – Fluency scale. WRAT = Wide range achievement test. The horizontal solid black line y = 0 represents the standardised mean. The horizontal grey dotted lines y = -1.5 and y = 1.5 encompass the range usually considered "typical" in mathematical cognition within the clinical and the educational contexts (e.g., in Santos et al., 2022). The dot size is scaled based on the sample size.

<u>Relationship with FSIQ, VIQ and NVIQ</u>. Figure 4.12 shows the distribution of mean calculation scores (transformed into *z*-scores) plotted against mean FSIQ, VIQ, and NVIQ. Calculation skills were investigated in studies including autistic samples with high mean FSIQ (average mean FSIQ = 94.64; mean FSIQ range = 60.00 - 124.89). Visual examination of the charts shows a trend where higher mean calculation scores correspond to higher mean IQ scores, for all IQ measures. Four out of the seven samples with mean FSIQ lower than the ID threshold lie below the lower end of the calculation typical range. Three out of the seven samples with mean FSIQ lower than the ID threshold lie within the lower half of the typical range.

To investigate how performance on calculation tasks relates to FSIQ, VIQ, and NVIQ, separate Spearman's correlations were calculated between the mean score of calculation and the mean score of IQ measures. This showed statistically significant results and indicated a strong and positive association between calculation and FSIQ, VIQ, and NVIQ scores, FSIQ: *rs* (40) = .82, *p* < .001; VIQ: *rs* (28) = .72, *p* < .001; NVIQ: *rs* (26) = .70, *p* < .001. This further supports the trend observed in the chart, indicating that calculation scores in autism develop in conjunction with IQ measures.

Figure 4.12. Association between mean calculation *z*-scores and mean IQ scores by assessment tool.





Note: *n* FSIQ = 39. *n* VIQ = 27. *n* NVIQ = 24. KTEA-MC = Kaufman test of educational achievement – Math applications scale. WIAT-NO = Wechsler Individual Achievement Test – Numerical operation scale. WJ-C = Woodcock–Johnson test of achievement – Calculation scale. WJ-F = Woodcock–Johnson test of achievement – Fluency scale. WRAT = Wide range achievement test. The vertical red dotted line x = 75 represents the threshold used in the current study to define ID. The horizontal solid black line y = 0 represents the standardised mean. The horizontal grey dotted lines y = -1.5 and y = 1.5 encompass the range usually considered “typical” in mathematical cognition within the clinical and the educational contexts (e.g., in Santos et al., 2022). The dot size is scaled based on the sample size.

Case-control studies with a TD control group. To examine how calculation skills compare to TD population, results from case-control studies were examined. Out of the 22 case-control studies investigating calculation abilities, 17 studies compared an autistic sample with a TD group, for a total of 21 samples and 25 measures. This subgroup included autistic samples across a broad CA range (average mean CA = 10.65; mean CA range = 6.36 - 26.00) and exhibiting high mean FSIQ scores (average mean FSIQ = 104.10; mean FSIQ range = 93.00 - 120.25). Notably, none of these studies featured a sample with mean FSIQ lower than the threshold for ID. However, the variability in FSIQ scores within each sample (shown by the high values of SD for mean FSIQ scores in Table 4.8), suggests substantial variability in the IQ measures and the possible inclusion of individuals with ID in the samples. These studies tested calculation abilities using a range of different assessments, that included CDR-P (*n* = 4), KTEA-MC (*n* = 3), TOMA (*n* = 1), TTR (*n* = 5), WIAT-NO (*n* = 11) and WRAT (*n* = 1). Table 4.8 summarises the findings from these studies.

Table 4.8. Overview and interpretation of findings from case-control studies with a TD control group investigating calculation abilities.

Article	*N*	Mean CA (*SD*)	Mean FSIQ (*SD*)	Matching criteria	Maths measure	Reported Findings	Interpretation of findings
Aagten-Murphy et al. (2015)	32	10.28 (1.30)	106.30 (9.60)	CA, FSIQ (Wechsler)	WIAT-NO	AG scored significantly lower than TDG; F (1, 63) = 9.76, p =.003.	Not in line with CA and not in line with MA (FSIQ).
Bae et al. (2015)	20	10.60 (0.94)	109.60 (15.85)	CA, FSIQ (KBIT)	TOMA-C	No statistically significant differences found between AG and TDG; F (1, 38) = 1.43, p = .161.	In line with CA and in line with MA (FSIQ).
Bullen et al. (2020)	77	11.38 (2.20)	98.60 (n/r)	CA	WIAT-NO	AG scored significantly lower than TDG; p < .05.	Not in line with CA.
Chen et al. (2019)	96	9.67 (1.49)	107.75 (17.91)	CA, FSIQ (Wechsler)	WIAT-NO	AG scored significantly lower than TDG; p = .032.	Not in line with CA and not in line with MA (FSIQ).
Goldstein et al. (1994)	29	Younger than 13 y	98.93 (12.48)	CA, sex, FSIQ (Wechsler), SES	KTEA-MC	No statistically significant differences found between AG and TDG; p not reported.	In line with CA and in line with MA (FSIQ).
Goldstein et al. (1994)	29	Older than 13 y	93.00 (15.16)	CA, sex, FSIQ (Wechsler), SES	KTEA-MC	No statistically significant differences found between AG and TDG; p not reported.	In line with CA and in line with MA (FSIQ).
Hiniker et al. (2016)	36	9.66 (1.60)	109.19 (20.45)	CA, FSIQ (Wechsler)	WIAT-NO	No statistically significant differences found between AG and TDG; t = -.61, p = .541.	In line with CA and in line with MA (FSIQ).
Iuculano et al. (2014)	18	9.60 (1.64)	113.27 (15.25)	CA, sex, FSIQ (Wechsler)	WIAT-NO	AG scored significantly higher than TDG; t (34) = 2.64, p = .012.	Not in line with CA and not in line with MA (FSIQ).
Iuculano et al. (2020)	16	9.46 (1.80)	120.25 (15.25)	CA, sex, FSIQ (Wechsler)	WIAT-NO	No statistically significant differences found between AG and TDG; p = .140.	In line with CA and in line with MA (FSIQ).
Kljajevic (2023)	48	8.27 (1.45)	102.77 (17.27) (*)	CA	WIAT-NO	AG scored significantly lower than TDG; t (54.112) = -3.33, p = .002, d = -.749.	Not in line with CA.
May et al. (2013)	64	9.90 (1.84)	96.78 (13.16)	CA, sex	WIAT-NO	AG scored significantly lower than TDG; t = -3.49, p = .001.	Not in line with CA.
May et al. (2015)	40	9.62 (1.57)	n/r	CA, sex, NVIQ (Wechsler)	WIAT-NO	No statistically significant differences found between AG and TDG; F (1,78) = 3.93, p = .051.	In line with CA and in line with MA (NVIQ).
Mayes and Calhoun (2007)	118	n/r	n/r	no matching rules	WIAT-NO	No statistically significant differences found between AG and TDG; p = .300.	n/a – no matching rules.
McCauley et al. (2018)	44	12.78 (2.10)	98.77 (14.21)	CA, school grade	WIAT-NO	AG scored significantly lower than TDG; p < .01.	Not in line with CA.
Minshew et al. (1994)	54	16.30 (10.16)	95.50 (15.54)	CA, sex, FSIQ (Wechsler), race, SES	KTEA-MC	No statistically significant differences found between AG and TDG; t = -0.82, p > .05.	In line with CA and in line with MA (FSIQ).

Rumsey and Hamburger (1988)	10	26.00 (7.00)	103.40 (9.47) (*)	CA, VIQ (Wechsler), NVIQ (Wechsler), WRAT scores	WRAT	No statistically significant differences found between AG and TDG; *p* not reported.	In line with CA and in line with MA (VIQ; NVIQ; WRAT).
Titeca et al. (2014)	33	6.87 (0.29)	n/r	CA, FSIQ (Wechsler), SES, sex	TTR	AG scored significantly lower than TDG; F (1, 83) = 4.44, p = .038.	Not in line with CA and not in line with MA (FSIQ).
Titeca, Roeyers, Loeys, et al. (2015)	31	6.36 (0.24)	102.67 (12.31)	Normed sample	TTR	No statistically significant differences found between AG and the normed sample; *p* not reported.	In line with CA.
Titeca, Roeyers, Loeys, et al. (2015)	27	7.38 (0.26)	104.19 (15.41)	Normed sample	TTR	No statistically significant differences found between AG and the normed sample; *p* not reported.	In line with CA.
Titeca, Roeyers, Loeys, et al. (2015)	39	8.28 (0.31)	107.16 (13.80)	Normed sample	TTR	No statistically significant differences found between AG and the normed sample; *p* not reported.	In line with CA.
Titeca, Roeyers, Loeys, et al. (2015)	24	9.28 (0.30)	105.62 (13.57)	Normed sample	TTR	No statistically significant differences found between AG and the normed sample; *p* not reported.	In line with CA.
Titeca, Roeyers, Loeys, et al. (2015)	31	6.36 (0.24)	102.67 (12.31)	Normed sample	CDR-P	AG scored significantly lower than the normed sample; *p* not reported.	Not in line with CA.
Titeca, Roeyers, Loeys, et al. (2015)	27	7.38 (0.26)	104.19 (15.41)	Normed sample	CDR-P	No statistically significant differences found between AG and the normed sample; *p* not reported.	In line with CA.
Titeca, Roeyers, Loeys, et al. (2015)	39	8.28 (0.31)	107.16 (13.80)	Normed sample	CDR-P	No statistically significant differences found between AG and the normed sample; *p* not reported.	In line with CA.
Titeca, Roeyers, Loeys, et al. (2015)	24	9.28 (0.30)	105.62 (13.57)	Normed sample	CDR-P	No statistically significant differences found between AG and the normed sample; *p* not reported.	In line with CA.

Note. n = 25. (*) = VIQ has been reported when the FSIQ was missing. (n/r) = not reported. (n/a) = not applicable. AG = Autistic group. TDG = Typically developing group. CDR-P = Cognitive developmental skills in arithmetic – Procedural calculation scale. KTEA-MC = Kaufman test of educational achievement – Math applications scale. TOMA = Test of mathematical abilities. TTR = Arithmetic number facts test. WIAT-NO = Wechsler Individual Achievement Test – Numerical operation scale. WRAT = Wide range achievement test.

The interpretation of the findings through the neuroconstructivist perspective shows that eight samples (67%) out of the 12 using a TD control group matched for both CA and MA (FSIQ) reported autistic performance in line with the TD control group. Eight autistic samples were compared to a CA-matched TD control group. For four of these samples two different assessments were used to measure calculation skills (TTR and CDRP; Titeca, Roeyers, Loeys, et al, 2015). A total of 12 calculation measures were reported. For seven of these measures (58%) a performance in line with the TD control group was reported.

Out of the nine samples which reported statistically significant differences with the TD control group, only one sample (11%) reported higher calculation scores for the autistic population (Iuculano et al., 2014). The findings reporting higher scores for the autistic group were explained by the strategy used by participants to perform the calculation task. In fact, Iuculano et al. (2014) reported that the autistic group showed greater use than the TD group of a sophisticated strategy which involved breaking down the original calculation into two or more simpler subproblems. The vast range of assessments used in the studies and the small number of available data points with the same assessment and outcome made it not possible to analyse the relationship between calculation outcomes and the type of assessment used.

<u>Summary.</u> In summary, calculation skills were investigated in samples with a wide mean CA range (average mean CA = 12.71 years; mean CA range = 6.36 - 29.00) and high mean FSIQ (average FSIQ = 94.64; mean FSIQ range = 60.00 - 124.89).

The average performance of most of the autistic samples on standardised assessment falls within the typical range and most of them lie below the standardised mean *z*-score. All the samples reporting calculation scores below the lower end of the typical calculation range (*n* = 4) were characterised by mean FSIQ scores below the ID threshold. However, there were samples characterised by mean FSIQ scores below the ID threshold which reported calculation scores inside the typical calculation range (*n* = 3).

The analysis of the developmental trajectory of standardised scores from cross-sectional samples shows that calculation skills develop in line with the normed population and that scores from older samples vary less than in younger samples.

Calculation skills are positively correlated with IQ scores.

Findings from case-control studies with a TD control group are mixed. However, when the performance of the autistic sample was not in line with the TD control group, most of the studies reported lower scores for the autistic sample.

4.3.3.3 Overall mathematics achievement

Overview. Forty-five studies assessed overall mathematics achievement across 50 autistic samples and reported 52 outcome measures. Within the 45 studies, there were three case studies, eight correlational studies, eight case-control studies, four interventions and 22 single-case research studies. The included studies used 12 different measures to assess this ability, and the most used test was the WJ (*n* = 13, 25%; Table 4.9).

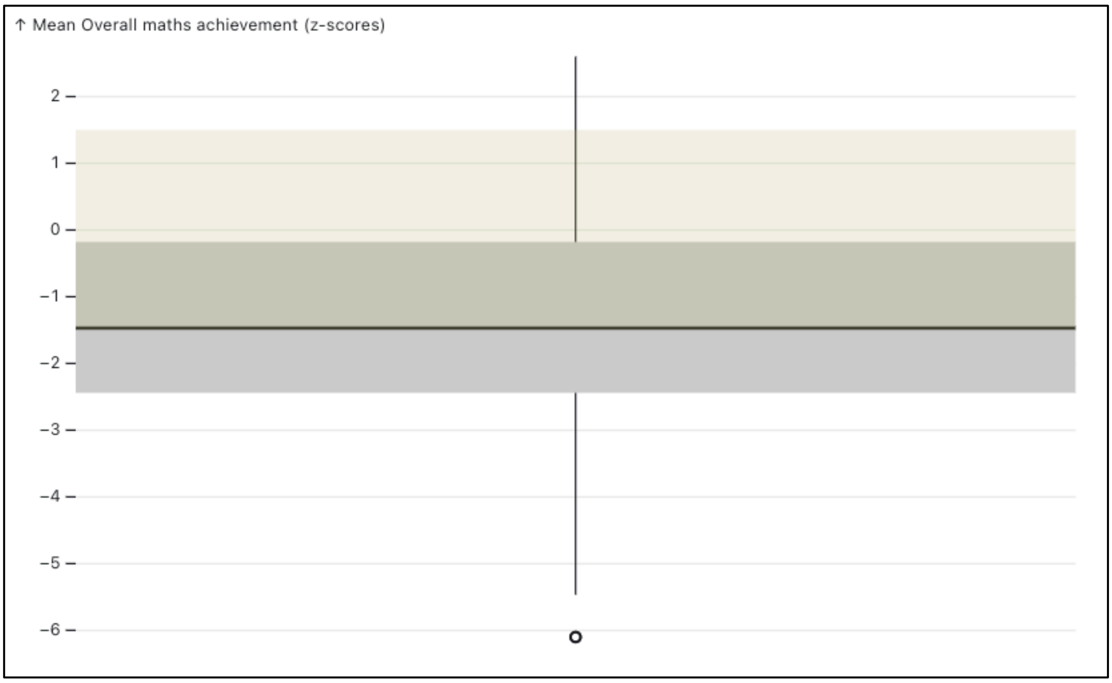
Table 4.9. Assessment tools used to assess overall mathematics achievement.

Assessment	Scale	*N*	%
Woodcock–Johnson test of achievement (Woodcock et al., multiple versions)	Broad math (WJ)	13	25
Test of early mathematics ability (TEMA; Ginsburg & Baroody, 2003)	n/a	12	23
KeyMath diagnostic assessment (KeyMath; Connolly, 2007)	n/a	9	17
Wechsler individual achievement test (Wechsler, multiple versions)	Mathematics composite (WIAT)	7	13
Not standardised researcher-developed task	n/a	3	6
Differential ability scales (Elliott, C. D., 1990)	Basic number skills (DAS)	2	4
Test of mathematical abilities (Brown et al., 2013)	Maths ability index (TOMA)	2	4
Star maths test (STAR; Renaissance, 2019)	n/a	1	2
Peabody individual achievement test (Dunn & Markwardt, 1970)	Mathematics (PIAT)	1	2
School performance test (SPT; Knijnik, Giacomoni, & Stein, 2013)	Math	1	2
Mixed assessments (WIAT and WJ)	n/a	1	2

Note: *n* = 52. n/a = not applicable.

Figure 4.13 illustrates how the mean performance in assessments measuring overall mathematics achievement is distributed across the autistic samples for which standardised scores were reported (*n* = 32). The boxplot symmetry indicates a balanced distribution of the data around the median. The spread of the middle 50% of the samples lies below the standardised mean *z*-score of 0, with less than 25% of the scores falling above the standardised mean level. The median of the mean score of the autistic samples lies below the expected mean score at -1.5 standard deviations. The whiskers extend over the typical range, with maximum score at just over 2 standard deviations and minimum score below -5 standard deviations. Two outliers are plotted 6 standard deviations below the standardised mean *z*-score. These represent a sample from the study by Kurth and Mastergeorge (2010) made of autistic students with a mean FSIQ score below the ID threshold, and one single-case study by Vostanis et al. (2023) which did not report any IQ measure. Overall, 50% of the samples z-scores (*n* = 16) lied above the lower end of the typical range. Only 10 of the remaining 16 samples reported mean FSIQ scores. Out of these 10 samples, five were characterised by mean FSIQ scores below the ID threshold, while five reported mean FSIQ scores above or equal to it (mean IQ score between 75.00 and 89.88).

Figure 4.13. Distribution of mean overall mathematics achievement *z*-scores.

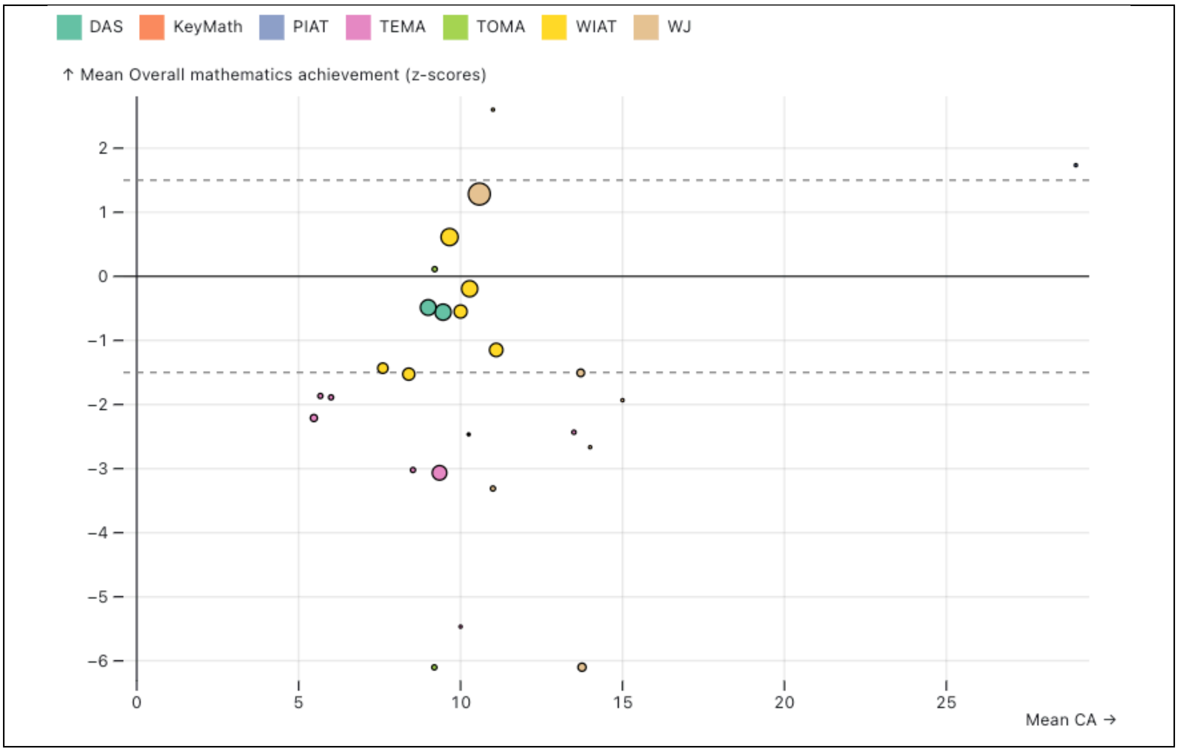


Note: *n* = 32. The light grey section between -1.5 and 1.5 standard deviations from the standardised mean shows the range usually considered “typical” in mathematical cognition within the clinical and the educational contexts (e.g., in Santos et al., 2022).

<u>Relationship with CA.</u> Figure 4.14 shows the distribution of mean scores of overall mathematics achievement (transformed into *z*-scores) plotted against CA. Overall mathematics achievement was investigated in studies including autistic samples with a wide mean CA range (average mean CA = 10.57 years; mean CA range = 4.63 - 29.00), with most studies investigating overall mathematics achievement in children and adolescents. Visual analysis of the chart shows that

most of the mean z-scores lie below the standardised mean, independently of the assessment being used. However, the mean scores reported for the WIAT, WJ, and DAS assessments fall within the typical range, while mean scores measured using the other tests, fall further away from the standardised mean *z*-score. In particular, the mean scores reported for the TEMA fall at least two standard deviations below the standardised mean.

Figure 4.14. Association between mean overall mathematics achievement *z*-scores and mean CA by assessment tool.



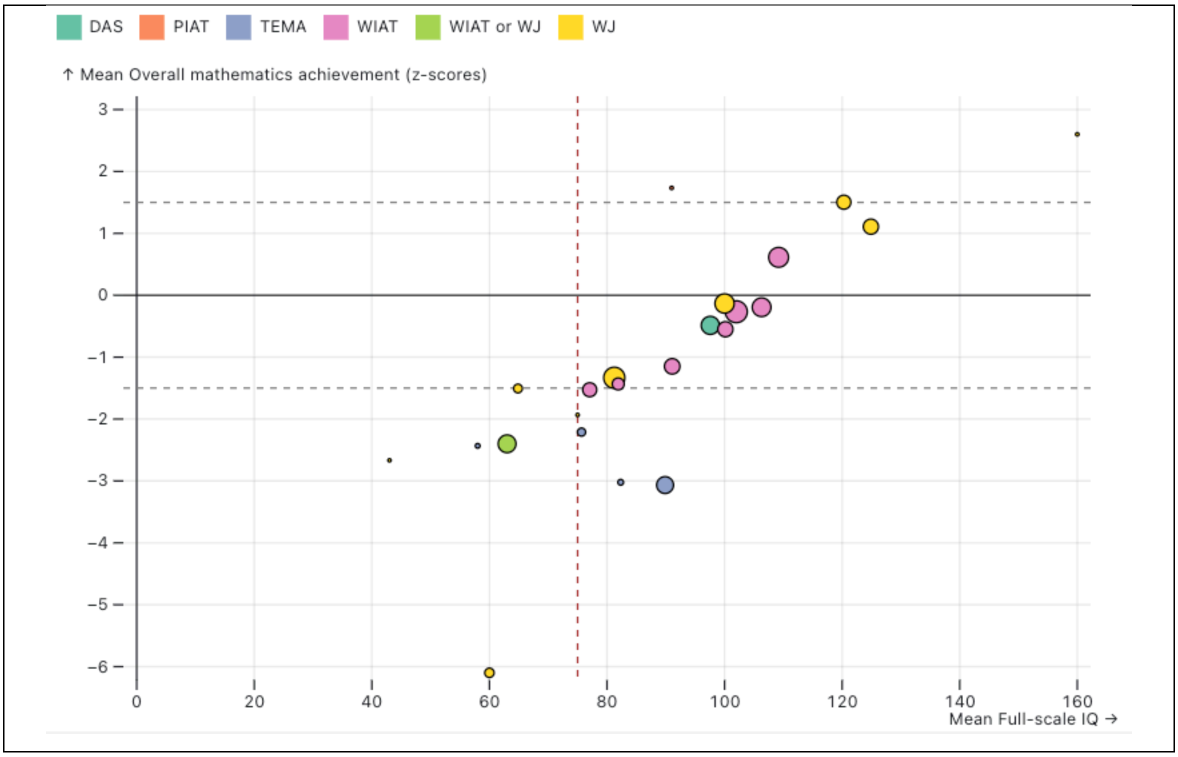
Note: *n* = 26. DAS = Differential ability scales. KeyMath = KeyMath diagnostic assessment. PIAT = Peabody individual achievement test. TEMA = Test of mathematical abilities. TOMA = Test of mathematical abilities. WIAT = Wechsler Individual Achievement Test. WJ = Woodcock–Johnson test of achievement. The horizontal solid black line y = 0 represents the standardised mean. The horizontal grey dotted lines y = -1.5 and y = 1.5 encompass the range usually considered “typical” in mathematical cognition within the clinical and the educational contexts (e.g., in Santos et al., 2022). The dot size is scaled based on the sample size.

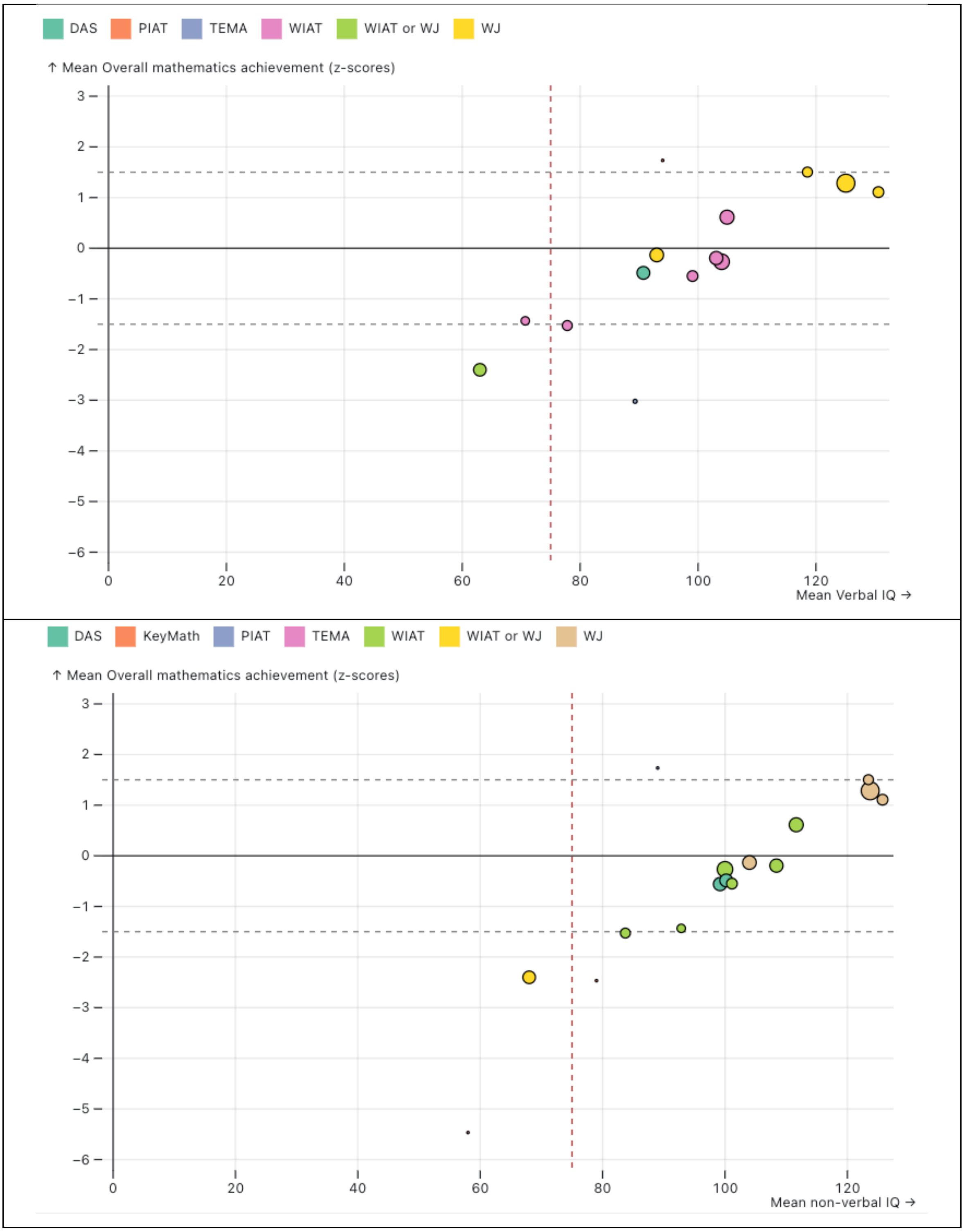
Relationship with FSIQ, VIQ and NVIQ. Figure 4.15 shows the distribution of mean scores of overall mathematics achievement (transformed into *z*-scores) plotted against mean FSIQ, VIQ, and NVIQ. Overall mathematics achievement was investigated in studies including autistic samples with high mean FSIQ (average mean FSIQ = 89.32; mean FSIQ range = 43.00 - 160.00). Visual inspection of the charts shows a similar pattern for all IQ measures, where higher scores of overall mathematics achievement correspond to higher scores of mean

IQ. Half of the 12 samples with mean overall mathematics achievement *z*-score below the typical range reported a measure of IQ scores below the ID threshold, the other half reported measures of IQ scores above the ID threshold.

To investigate how performance on overall mathematics achievement relates to FSIQ, VIQ, and NVIQ, separate Spearman's correlations were calculated between the mean score of mathematical performance and the mean score of IQ measures. This showed a statistically significant result only for FSIQ that indicated a strong and positive association between the two variables, FSIQ: *rs* (22) = .82, $p < .001$; VIQ: *rs* (14) = .43, $p = .126$; NVIQ: *rs* (16) = .49, $p = .064$. The lack of statistically significant results for VIQ and NVIQ can be explained by the small number of the samples reporting those measures (VIQ $n = 14$; NVIQ $n = 16$).

Figure 4.15. Association between overall mathematics *z*-scores and mean IQ scores by assessment tool.





Note: *n* FSIQ = 23. *n* VIQ = 14. *n* NVIQ = 16. DAS = Differential ability scales. KeyMath = KeyMath diagnostic assessment. PIAT = Peabody individual achievement test. TEMA = Test of mathematical abilities. WIAT = Wechsler Individual Achievement Test. WJ = Woodcock–Johnson test of achievement. The vertical red dotted line x = 75 represents the threshold used in the current study to define ID. The horizontal solid black line y = 0 represents the standardised mean. The horizontal grey dotted lines y = -1.5 and y = 1.5 encompass the range usually considered “typical” in mathematical cognition within the clinical and the educational contexts (e.g., in Santos et al., 2022). The dot size is scaled based on the sample size.

Case-control studies with a TD control group. To examine how overall mathematics achievement compares to TD population, results from case-control studies were investigated. Out of the 6 case-control studies investigating overall mathematics achievement, 5 studies compared an autistic sample with a TD group. This subgroup included autistic samples characterised by a small mean CA and a narrow mean CA range (average mean CA = 9.00 years; mean CA range = 4.63 - 11.10) and exhibiting high mean FSIQ scores (average mean FSIQ = 104.64; mean FSIQ range = 89.88 - 126.75). None of these studies included a sample with mean FSIQ lower than the threshold for ID. Nevertheless, the considerable variability in FSIQ scores within each sample (shown by the high values of SD for mean FSIQ scores), suggest substantial variability in the measures of IQ. These studies investigated overall mathematics achievement using two assessments, TEMA (*n* = 2) and the Mathematics Composite scale of the WIAT (*n* = 3). Table 4.10 summarises the findings from these studies.

Table 4.10. Overview and interpretation of findings from case-control studies with a TD control group investigating overall mathematics achievement.

Article	*N*	Mean CA (*SD*)	Mean FSIQ (*SD*)	Matching criteria	Maths assessment	Reported Findings	Interpretation of findings
Aagten-Murphy et al. (2015)	32	10.28 (1.30)	106.30 (9.60)	CA, FSIQ (Wechsler)	WIAT	AG scored significantly lower than TDG; F (1,62) = 6.66, p = .010 eta square = 0.10.	Not in line with CA and not in line with MA (FSIQ).
Hiniker et al. (2016)	36	9.66 (1.60)	109.19 (20.45)	CA, FSIQ (Wechsler)	WIAT	No statistically significant differences found between AG and TDG; t = -.26, p = .793.	In line with CA and in line with MA (FSIQ).
McDougal et al. (2020)	22	11.10 (2.81)	91.09 (14.82)	CA, FSIQ (Wechsler)	WIAT	AG scored significantly lower than TDG; t = 4.62, p < .001, d = 1.03.	Not in line with CA and not in line with MA (FSIQ).
Polo-Blanco et al. (2024)	26	9.35 (2.06)	89.88 (11.78)	CA, sex, SES	TEMA	AG scored significantly lower than TDG; t (50) = -2.89, p = .006, d = -0.81.	Not in line with CA.
Wang et al. (2023)	24	4.63 (0.65)	126.75 (11.95) (*)	CA, sex	TEMA	AG scored significantly lower than TDG; t = 5.32, p < .001, d = -0.81.	Not in line with CA.

Note. *n* = 5. (*) = VIQ has been reported when the FSIQ was missing. AG = Autistic group. TDG = Typically developing group. TEMA = Test of mathematical abilities. WIAT = Wechsler Individual Achievement Test.

The interpretation of the findings through the neuroconstructivist perspective shows that one sample (33%) out of the three using a TD control group matched for both CA and MA (FSIQ) reported autistic performance in line with the TD control group. Two samples out of two using a TD control group matched only for CA reported autistic performance not in line with the TD control group.

All the samples which reported statistically significant differences with the TD control group (*n* = 4), reported lower overall mathematics achievement scores for the autistic population. The small number of case-control studies with a TD control group investigating overall mathematics achievement made it not possible to analyse the relationship between overall mathematics achievement and the type of assessment used.

Summary. In summary, overall mathematics achievement was investigated in samples with a wide mean CA range (average mean CA = 10.57 years; mean CA range = 4.63 - 29.00) and high mean FSIQ (average mean FSIQ = 89.32; mean FSIQ range = 43.00 - 160.00).

The average performance of 50% of the autistic samples on standardised assessment falls below the typical range. The samples lying below the lower end of the typical range and reporting IQ scores (*n* = 12) are characterised by mean IQ scores both above and below the ID threshold (mean IQ score range = 43.00 – 89.88).

The analysis of the developmental trajectory of standardised scores from cross-sectional samples shows that overall mathematics achievement does not develop in line with the normed population.

Arithmetic word problem skills are positively correlated with FSIQ scores.

The findings from all case-control studies reporting outcomes that are not in line with the TD control group report lower scores for the autistic sample.

4.3.3.4 Enumeration

Enumeration skills were assessed in eight case-control studies, seven of which with a TD control group and one with a clinical control group. Table 4.11 reports an overview of these studies and the interpretation of their findings.

All the experiments used computer-based tasks where sets of stimuli were presented to the participants, who were asked to report “how many” items were on the screen as quickly and as accurately as possible.

Enumeration skills were measured using different types of outcomes, which included accuracy scores, RTs, or measures based on RTs that took into consideration the speed-accuracy trade-off. Although variations existed in methodology regarding stimulus type, time of stimulus presentation, instructions given, duration of the task, and inclusion of practice trials, the studies consistently reported congruent results when assessing accuracy scores. However, inconsistencies arose when utilising RTs as an outcome measure.

Perceptual subitizing skills were investigated in seven studies which included autistic samples with mean CA between 5.13 and 23.09 years and with mean FSIQ scores higher than 100. All the studies reported perceptual subitizing skills in line with CA and with MA, based on accuracy scores. However, studies relying on RTs as an outcome measure presented conflicting results and reported both performance in line with CA and MA and performance not in line with CA and MA, characterised by longer RT time for the autistic group. For example, the study by O'Hearn et al. (2013) reported that the younger autistic group (mean CA = 9.00 years) did not perform perceptual subitizing but rather counted the items presented, even when up to 3 objects were shown on the screen. The use of a serial strategy to perform the enumeration task even with numerosity within the subitizing range might be explained by the different instructions that were given to participants. In fact, in the study by O'Hearn et al. (2013) the researcher suggested participants to “count in their head” in case they were not able to provide their answer quickly, and in the study by Jarrold and Russell (1997) participants were explicitly instructed to count. Conversely, for all the studies that reported abilities in line with CA and with MA the researcher asked participants to tell “how many” stimuli were in the screen.

Finally, both O'Hearn et al. (2013) and Gagnon et al. (2004) reported a narrower subitizing range for the autistic group compared to the TD group (i.e., 1-3 rather than 1-4).

Conceptual subitizing skills were investigated in two of the studies described above, which included in their design an experimental condition where stimuli were arranged in a dice pattern (Jarrold & Russell, 1997; O'Hearn et al., 2013). The study by O'Hearn et al. (2013) concluded that the autistic groups reported conceptual subitizing skills in line with CA and MA based on the accuracy scores, but they exhibited longer RTs than the TD group which were not in line with CA and MA. Similarly, Jarrold and Russell (1997), reported longer RTs for the autistic participants that showed reduced advantage for numerosity 5 and 6 when items were presented in a dice pattern rather than randomly. As explained for the perceptual subitizing task, the longer RTs observed for the autistic group to perform the enumeration task might be explained by the choice of the language used to give instructions to the participants.

Counting skills were assessed in eight studies, which included autistic samples with mean CA between 5.13 and 23.09 years and with mean FSIQ scores higher than 100. All the studies which used accuracy scores to measure counting skills reported counting abilities in line with CA and MA except for the study by Titeca, Roeyers, Ceulemans, et al. (2015) for the numerosity of 4 only. When counting abilities were measured through measures based on RTs, the results were inconsistent, with some studies reporting performance in line with CA and MA (Gagnon et al., 2004; Titeca, Roeyers, Ceulemans, et al., 2015; Titeca, Roeyers, & Desoete, 2015) and other studies reporting a performance characterised by significantly higher RTs for the autistic group than for the TD group (Jarrold & Russell, 1997; O'Hearn et al., 2013).

Three of the studies assessing counting abilities also assessed procedural and conceptual counting abilities of young autistic children (mean CA range: 5.13; 6.27 years) with no ID (mean FSIQ range: 104.83; 105.38) using two sub-tests of the TEDI-MATH assessment and reported performance in line with CA and MA for the autistic group (Titeca, Roeyers, Ceulemans, et al., 2015; Titeca, Roeyers, & Desoete, 2015; Titeca et al., 2014).

In summary, the case-control studies with a TD control group investigating enumeration skills and using accuracy scores reported perceptual subitizing, conceptual subitizing and counting skills in autistic samples with high FSIQ scores in line with CA and with MA. However, when the enumeration skills were assessed using RTs, the findings from the included studies were inconsistent. When the performance was not in line with CA and MA, this was characterised by longer RTs, suggesting a shift towards using counting strategies instead of subitizing strategies to perform the enumeration task, even within the subitizing range. This could be explained by the instructions provided to the participants. Finally, a narrower subitizing range for the autistic group has been reported, reduced to 1-3 rather than 1-4 or 1-5, as observed in the TD population.

Table 4.11. Overview and interpretation of findings from case-control studies with a TD control group investigating enumeration skills.

Article	Task	*N*	Mean CA (*SD*)	Mean FSIQ (*SD*)	Matching criteria	Assessment description	Reported Findings	Interpretation of findings
Gagnon et al. (2004)	PS	14	15.07 (3.04)	106.8 (9.8)	CA, sex, FSIQ (Wechsler)	**Stimuli:** Display on a computer screen of 2-9 white squares on a black background randomly arranged. **Total number of trials:** 320. **Presentation Time:** 600 ms. **Instructions:** "Tell how many white squares you see on the screen as fast and as accurate as you can". **Practice trials:** Yes. **Outcome variables:** Accuracy and RT.	No statistically significant differences found between AG and TDG in accuracy (*F* < 1; ns). No statistically significant differences found between AG and TDG in RT (*p* not reported). Significant main effect of group on subitizing range (*F* (1,26) = 17.512; p < .001), that was found smaller for autistic participants (1-3 rather than 1-4).	PS in line with CA and MA (accuracy). PS in line with CA and MA (RT). Narrower PS range.
	C						No statistically significant differences found between AG and TDG in accuracy (*F* < 1; ns) and in RT (*p* not reported).	C in line with CA and MA (accuracy). C in line with CA and MA (RT).
O'Hearn et al. (2013)	PS	**S1**: 9 **S2**: 15 **S3**: 15	**S1**: 11.57 (1.11) **S2**: 15.35 (1.77) **S3**: 23.09 (4.54)	**S1**: 100.25 (6.69) **S2**: 105.08 (14.26) **S3**: 107.43 (12.51)	CA, sex, FSIQ (Wechsler)	**Stimuli:** Display on a computer screen of 1-8 dark grey squares of different size on a light grey background in two arrangements: 1) dice patterns (only 1-6); 2) random. **Total number of trials:** 48. **Presentation Time:** Untimed. **Instructions:** Tell how many squares are presented on the screen as quickly and as accurate as possible. If you don't know the answer quickly you can count the squares in your head". **Practice trials:** Yes. **Outcome variables:** Accuracy, RT, and corrected RT (mean RT / % correct answers).	No statistically significant differences found between AG and TDG in accuracy; *p* not reported. AG reported higher RT than TDG; *F* (1,72) = 16.58, *p* < .001. Significant main effect of group on subitizing range (χ^2 (4) =13.71, *p* =.008, Cramer's *V* = .42) that was found smaller for autistic participants (1-3 rather than 1-4).	PS in line with CA and MA (accuracy). PS not in line with CA and MA (RT). Narrower PS range.
	CS						No statistically significant differences found between AG and TDG in accuracy; *F* (1, 72) = 3.86, *p* = .060.	CS in line with CA and MA (accuracy). CS not in line with CA and MA (RT).

							AG reported higher RT than TDG; F (1,72) = 16.58, p <.001.	
	C						No statistically significant differences found between AG and TDG in accuracy; p not reported. Autistic reported higher RT than TD; F (1,72) = 16.58, p < .001.	C in line with CA and MA (accuracy). C not in line with CA and MA (RT).
Jarrold and Russell (1997)	PS	22	12.47 (2.82)	83.09 (17.92) (*)	VIQ (BPVS)	**Stimuli:** Display on a computer screen of 3-6 black dots on a white background in two arrangements: 1) dice patterns; 2) random with white squares as visual distractors (twice as numerous as the black dots). **Total number of trials:** 24. **Presentation Time:** Untimed. The next stimulus was only shown once the participant had counted the number of dots correctly. **Instructions:** "Count the black dots on the screen as quickly as possible". **Practice trials:** Not reported. **Outcome variables:** Difference of RT between the two arrangements (diff RT).	No statistically significant differences found between AG and TDG in diff RT; p not reported.	PS in line with MA (RT).
	CS						No statistically significant differences found between AG and TDG in diff RT; p not reported.	CS in line with MA (RT)
	C						AG reported smaller diff RT scores than TDG for 5 and 6 dots; p < .001 in both cases.	C not in line with MA (RT)
Kirk et al. (2017)	C	23	7.20 (1.74)	64.95 (8.51)	No matching criteria	**Give a number task**	AG performed better than those with DS; p not reported.	n/a – no matching rules reported.
Titeca et al. (2014)	PS	33	6.27 (0.38)	105.38 (13.27)	CA, FSIQ (Wechsler) SES	**Stimuli:** Display on a computer screen of 1 - 9 black squares randomly arranged. Task controlled for continuous variables. **Total number of trials:** 72. **Presentation Time:** 120 ms. **Instructions:** "Tell how many squares are presented on the screen as quickly and as accurate as possible". **Practice trials:** Yes. **Outcome variables:** Accuracy.	No statistically significant differences found between AG and TDG in accuracy; F (5, 81) = 1.17, p = .330.	PS in line with CA and MA (accuracy).
	C						No statistically significant differences found between AG and TDG in accuracy; F (5, 81) = 1.17, p = .330.	C in line with CA and MA (accuracy).
	PCC					**TEDI-MATH subtest 1 and 2**	No statistically significant differences found between AG and TDG in; p not reported.	PCC in line with CA and MA.
Titeca, Roeyers,	PS	30	5.98 (0.31)	104.83 (12.36)			No statistically significant differences found between AG and	PS in line with CA and MA (accuracy)

Ceulemans, et al. (2015)							TDG in accuracy; U = 419.50, p = .987. No statistically significant differences found between AG and TDG in RT; F (1, 56) = 0.33, p = .570.	PS in line with CA and MA (RT)
	C				CA, FSIQ (Wechsler) SES, sex	**Stimuli:** Display on a computer screen of 1 - 9 black squares on white background randomly arranged. Task controlled for continuous variables. **Total number of trials:** 72. **Presentation Time:** 120 ms. **Practice trials:** Yes. **Instructions:** “Tell how many squares are presented on the screen as quickly and as accurate as possible”. **Outcome variables:** Accuracy and RT.	Statistically significant group difference found between AG and TDG in accuracy for numerosity 4; U = 289.00, p = .039. No statistically significant differences found between AG and TDG in RT; F (1, 29) = 2.09, p = .159.	C for numerosities 5-9 in line with CA and MA (accuracy). C in line with CA and MA (RT).
	PCC					**TEDI-MATH subtest 1 and 2**	No statistically significant differences found between AG and TDG (subset 1: U = 345.00, p = .111; subset 2: U = 329.00, p = .067)	PCC in line with CA and MA.
Titeca, Roeyers and Desoete (2015)	PS	20	5.13 (0.33)	105.30 (13.90)	CA, FSIQ (Wechsler) SES, sex	**Stimuli:** Display on a computer screen of 1 - 9 black squares randomly arranged. **Total number of trials:** Not reported. **Presentation Time:** Not reported. **Practice trials:** Not reported. **Instructions:** “Tell how many squares are presented on the screen as quickly and as accurate as possible”. **Outcome variables:** Accuracy and RT.	No statistically significant differences found between AG and TDG in accuracy; U = 68.00, p = .112. No statistically significant differences found between AG and TDG in RT; U = 84.00, p = .377.	PS in line with CA and MA (accuracy) PS in line with CA and MA (RT)
	C						No statistically significant differences in accuracy; U = 69.50, p = .123. No statistically significant differences found between AG and TDG in RT; U = 102.00, p = .914.	C in line with CA and MA (accuracy) C in line with CA and MA (RT)
	PCC					**TEDI-MATH subtest 1 and 2**	No statistically significant differences found between AG and TDG (subtest 1: t (38) = −0.43, p = .673; subtest 2: t (38) = 0.12, p = .903).	PCC in line with CA and MA.

Soulieres et al. (2010)	PS	2	9.50 (n/r)	113.00 (2.80)	CA, FSIQ (Wechsler) sex	**Stimuli:** Display on a computer screen of 2- 9 white squares on black background randomly arranged. **Total number of trials:** 160. **Presentation Time:** 600 ms. **Practice trials:** Yes. **Instructions:** “Tell how many squares are presented on the screen. Give a fast and accurate response”. **Outcome variables:** % of error and RT.	Autistic participants and TD group displayed a similar pattern of error and of RT; *p* not reported.	n/a – accuracy data and RT data not reported
	C						Autistic participants and TD group displayed a similar pattern of error and of RT; *p* not reported.	n/a – accuracy data and RT data not reported

Note: *n* = 8. (PS): Perceptual Subitizing, (CP): Conceptual Subitizing, (C): Counting, (PCC): Procedural and Conceptual Counting. (*) = VIQ has been reported when the FSIQ was missing. (n/r) = not reported. (n/a) = not applicable. (ns) = not significant. AG = Autistic group. TDG = Typically developing group. DS = Down syndrome.

4.3.3.5 Magnitude comparison

Magnitude comparison skills were assessed with both symbolic (*n* = 1) and non-symbolic (*n* = 8) tasks in eight case-control studies with a TD control group. Table 4.12 reports an overview of these studies and the interpretation of their findings.

Participants were asked to either tell the researcher, point, or touch the side of the screen with the largest set of dots or to use a keyboard and press the button corresponding to the largest set.

Across the studies, mean CA of the autistic samples ranged between 4.63 and 10.28 years and mean FSIQ scores were higher than 100 with considerable variability across the samples shown by the high SD scores.

The findings from the studies investigating non-symbolic magnitude comparison skills varied depending on the outcome measure used by the researcher. This was not surprising because, as discussed in Chapter 1, the study by Dietrich et al. (2016) highlighted the variability in results derived from diverse measures used to explore magnitude comparison skills within the TD population and that these could not be used interchangeably. Research employing *w* and RTs as measures for magnitude comparison skills revealed consistent findings, while studies relying on accuracy as an outcome measure presented conflicting results. The three studies that used *w* as the outcome measure reported that autistic participants needed a larger ratio than TD to accurately discriminate numerosity. These findings were reported as autistic individuals showing either non-symbolic magnitude comparison skills not in line with CA and MA (Aagten-Murphy et al., 2015; Hiniker et al., 2016) or not in line with CA (Wang et al., 2023). Moreover, all the studies that used RTs as outcome measure (*n* = 3) reported non-symbolic magnitude comparison skills in line with CA and with MA (Hiniker et al., 2016; Titeca, Roeyers, Ceulemans, et al., 2015; Titeca, Roeyers, & Desoete, 2015). The six studies that reported accuracy scores as outcome measure showed conflicting findings. Hiniker et al. (2016) and Li et al. (2023) reported performance not in line with CA and with MA. The results by Wang et al. (2023) reported a performance not in line with CA. Conversely, all the studies published by Titeca (Titeca, Roeyers, Ceulemans, et al., 2015; Titeca, Roeyers, & Desoete, 2015; Titeca et al., 2014) reported a performance in line with CA and with MA for the autistic group.

These contrasting results lack a clear explanatory factor, as the studies involved similar samples in terms of mean CA and mean FSIQ but differed significantly in test methodologies, including the ratio between the two sets, presentation duration and length of the task, and inclusion of practice trials with feedback. Notably, Titeca and colleagues' studies were based in Belgium, whereas the others involved autistic participants based in the USA and in China. The variation in educational systems across these countries might contribute to the divergent outcomes. One out of the two autistic participants included in the study by Soulieres et al. (2010) (CA = 9.5 years, FSIQ = 115) reported no significant differences with the control group in line with the studies by Titeca (Titeca, Roeyers, Ceulemans, et al., 2015; Titeca, Roeyers, & Desoete, 2015; Titeca et al., 2014), while the other (CA = 9.5 years, FSIQ = 111) showed significantly higher accuracy scores compared to the TD group. This led the authors to conclude that there is some evidence that autistic children might show superior magnitude comparison abilities.

In contrast to the findings reported for the non-symbolic task, the only study investigating symbolic magnitude comparison reported abilities for the autistic group in line with CA and with MA, regardless of the measure used (Hiniker et al., 2016).

In summary, studies examining symbolic magnitude comparison skills in autistic individuals revealed abilities in line with CA and MA in terms of accuracy, RT, and ratio between the two stimuli. Non-symbolic magnitude comparison skills revealed that autistic participants needed a larger ratio than TD individuals for accurate numerosity discrimination and reported RTs in line with the TD population. As for accuracy scores, conflicting results emerged across the six studies, with some reporting performance in line with CA and MA and others not.

Table 4.12. Overview and interpretation of findings from case-control studies with a TD control group investigating numerical magnitude processing skills.

Article	Task	*N*	Mean CA (*SD*)	Mean FSIQ (*SD*)	Matching criteria	Assessment description	Reported Findings	Interpretation of findings
Aagten-Murphy et al. (2015)	NS	32	10.28 (1.30)	106.30 (9.60)	CA, FSIQ (Wechsler)	**Stimuli:** 2 sets of 48 black and white dots on grey background. Task controlled for continuous variables. **Total number of trials:** 200. **Presentation Time:** 500 ms. **Instructions:** Touch the side of the screen with more marbles. **Practice trials:** n/r. **Outcome variables:** w.	AG reported significantly higher *w* than TDG; $F(1,58) = 7.93$, $p = .007$, partial eta square = 0.12.	Not in line with CA and MA (*w*).
Hiniker et al. (2016)	NS	36	9.66 (1.60)	109.19 (20.45)	CA, FSIQ (Wechsler)	**Stimuli:** 2 sets of 2–9 green dots on black background. Ratios: .50, .67, .75, .80, .83, .86, .88, and .89. Task controlled for continuous variables. **Total number of trials:** 52. **Presentation Time:** 1500 ms. **Instructions:** Press the response key corresponding to the larger set using a left- or right-handed button-box. **Practice trials:** n/r **Outcome variables:** Accuracy, RT, and *w.*	AG reported significantly lower accuracy than TDG; $p < .05$. No statistically significant differences found between AG and TDG in RT; *p* not reported. Autistic group reported higher *w* than TD; $p < .05$.	Not in line with CA and MA (accuracy). In line with CA and MA (RT). Not in line with CA and MA (*w*).
	S					**Stimuli:** 2 Arabic numerals are displayed simultaneously on the screen. Ratios: .50, .67, .75, .80, .83, .86, .88, and .89. Task controlled for continuous variables. **Total number of trials:** 52. **Presentation Time:** 1500 ms. **Instructions:** Press the response key corresponding to the larger set using a left- or right-handed button-box. **Practice trials:** Not reported. **Outcome variables:** Accuracy, RT, and *w.*	No statistically significant differences found between AG and TDG in accuracy; *p* not reported. No statistically significant differences found between AG and TDG in RT; *p* not reported. No statistically significant differences found between AG and TDG in *w*; *p* not reported.	In line with CA and MA (accuracy). In line with CA and MA (RT). In line with CA and MA (*w*).
Li et al. (2023)	NS	70	5.20 (0.54)	100.77 (13.98) (*)	CA, NVIQ (RCPM)	**Stimuli:** 2 sets of 8–22 blue and red dots on white background. Ratios: .50, .67, .80. Task controlled for continuous variables. **Total number of trials:** 36. **Presentation Time:** n/r	AG reported significantly lower accuracy than TDG; $F(1, 185) =$	Not in line with CA and MA (accuracy).

						Instructions: Press the response key corresponding to / point to the larger set of dots. **Practice trials:** 2 **Outcome variables:** Accuracy	42.50, *p* < .001, partial eta square = 0.187.	
Titeca et al. (2014)	NS	33	6.27 (0.38)	105.38 (13.27)	CA, FSIQ (Wechsler), SES	**Stimuli:** 2 sets of black dots on white background. Ratios: .33, .50, .67, .75, .80, and .83. Task controlled for continuous variables. **Total number of trials:** 72. **Presentation Time:** 1200 ms. **Instructions:** Press the response key corresponding to the larger set on a five-button response box. **Practice trials:** Yes. **Outcome variables:** Accuracy.	No statistically significant group differences found between AG and TDG in accuracy; *p* not reported.	In line with CA and MA (accuracy).
Titeca, Roeyers, Ceulemans, et al. (2015)	NS	30	5.98 (0.31)	104.83 (12.36)	CA, FSIQ (Wechsler), SES, sex	**Stimuli:** 2 sets of black dots on white background. Ratios: .33, .50, .67, .75, .80, and .83. Task controlled for continuous variables. **Total number of trials:** 72. **Presentation Time:** 1200 ms. **Instructions:** Press the response key corresponding to the larger set on a response box. **Practice trials:** Yes. **Outcome variables:** Accuracy and RT.	No statistically significant differences found between AG and TDG in accuracy; *U* = 399.50, *p* = .590. No statistically significant differences found between AG and TDG in RT; *U* = 403.00, *p* = .628.	In line with CA and MA (accuracy). In line with CA and MA (RT).
Titeca, Roeyers and Desoete (2015)	NS	20	5.13 (0.33)	105.30 (13.90)	CA, FSIQ (Wechsler), SES, sex	**Stimuli:** 2 sets of black dots on white background. Ratios: .33, .50, .67, .75, .80, and .83. **Total number of trials:** n/r. **Presentation Time:** n/r. **Instructions:** Press the response key corresponding to the larger set on a response box. **Practice trials:** n/r. **Outcome variables:** Accuracy and RT.	No statistically significant group differences found between AG and TDG in accuracy; *F* (1, 28) = 0.26, *p* = .614. No statistically significant differences found between AG and TDG in RT; *F* (1, 28) = 0.12, *p* = .728.	In line with CA and MA (accuracy). In line with CA and MA (RT).
Soulieres et al. (2010)	NS	2	9.50 (n/r)	113.00 (2.80)	CA, sex, FSIQ (Wechsler)	**Stimuli:** 2 sets of white squares on black background. **Total number of trials:** 60. **Presentation Time:** 5,000 ms. **Instructions:** Press the response key corresponding to the larger set. **Practice trials:** Yes. **Outcome variables:** % of correct responses, % of error.	One autistic participant reported accuracy scores not statistically different from the TDG; *p* > .05. One autistic participant reported significantly higher accuracy scores compared to the TDG; *p* < .05.	n/a – accuracy data and RT data not reported.

Wang et al. (2023)	NS	24	4.63 (0.65)	126.21 (8.66) (*)	CA, sex	**Stimuli:** 2 sets of 5–21 blue and yellow dots on white background. Ratios: 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.7, 2.8, 2.9, and 3. Task controlled for continuous variables. **Total number of trials:** 88. **Presentation Time:** 2128 ms. **Instructions:** Tell the researcher which side has more dots. **Practice trials:** Yes. **Outcome variables:** Accuracy and *w*.	AG reported significantly lower accuracy than TDG; t (46) = 11.83, p < .001, Cohen's d = 3.68. AG reported significantly higher *w* than TDG; t (46) = − 6.94, p < .001, Cohen's d = 2.20.	Not in line with CA (accuracy). Not in line with CA (*w*).

Note: *n* = 8. (NS): Non-symbolic; (S): Symbolic. (*) = NVIQ was reported when the FSIQ was missing. (n/r) = not reported. (n/a) = not applicable. AG = Autistic group. TDG = Typically developing group.

4.3.3.6 Number line

Number line skills were assessed in four case-control studies with a TD control group and in one intervention study. Table 4.13 reports an overview of these studies and the interpretation of the findings of the case-control studies.

The studies used both pen-and-paper and computer-based number-to-position tasks. The experiments presented stimuli using different formats (Arabic digit, analogue magnitude, or verbal) and focused on different intervals (0-10, 1-10, 0-100, 0-1,000).

Across the studies, the mean CA of the autistic sample ranged between 5.13 and 10.28 years with small SDs and the mean FSIQ score was higher than 100, when reported.

The findings from the case-control studies with a TD control group (n = 4) suggest number line performance in line with CA and with MA for primary school age autistic individuals with IQ scores within the typical range. Moreover, Titeca, Roeyers, Ceulemans, et al. (2015) and Aagten-Murphy et al. (2015) identified a trend resembling the development observed in the TD population, especially for the Arabic digit format and verbal format (Titeca, Roeyers, Ceulemans, et al., 2015) and for the 0-1,000 task (Aagten-Murphy et al., 2015). As described in Chapter 1, this trend denotes a developmental shift from a logarithmic to a more linear representation in precision of the number line responses. However, the authors cautioned against definitive conclusions, expressing reservations about the strength of evidence supporting this pattern in autistic individuals. Finally, the results showed that the different format used to present the stimuli did not influence the performance of the autistic group, in line with what observed in the TD group (Titeca, Roeyers, Ceulemans, et al., 2015; Titeca, Roeyers, & Desoete, 2015).

Table 4.13. Overview and interpretation of findings from case-control studies with a TD control group investigating number line skills.

Article	Task	*N*	Mean CA (SD)	Mean FSIQ (SD)	Matching criteria	Assessment description	Reported Findings	Interpretation of findings
Aagten-Murphy et al. (2015)	AD	32	10.28 (1.30)	106.30 (9.60)	CA, FSIQ (Wechsler)	**Interval:** 1 – 100 and 1 – 1,000. **Stimuli:** Computer-based task. **Total number of trials:** 40. **Practice trials:** n/r. **Outcome variables:** Square root of the average squared difference between the selected location and the actual location of the number	No statistically significant differences found between AG and TDG in accuracy; 1 – 100: F (1, 62) = 3.75, p = .057, eta square= 0.06; 1–1,000: F (1, 62) = 3.36, p = .070, eta square = 0.05.	In line with CA and MA.
Satsangi and Bofferding (2017)	AD	5	6.50 (1.60)	n/r	No matching criteria	**Interval:** 0 – 10. **Stimuli:** Pen and paper task. The participant was asked to place an arrow shaped card showing a number between 1 and 9 on the number line. **Total number of trials:** 18. **Practice trials:** n/r. **Outcome variables:** Linearity, slope, and accuracy.	Participants included in the experimental group reported statistically significant improvement in their understanding of numerical relationships of numbers on the number line task.	n/a – no matching rules reported.
Titeca et al. (2014)	AD, AM, V	33	6.27 (0.38)	105.38 (13.27)	CA, FSIQ (Wechsler), SES	**Interval:** 0 – 100. **Stimuli:** Pen and paper task. For the AM task, the stimuli consisted of black dots on a white disc, controlled for perceptual variables. **Total number of trials:** 30. **Practice trials:** Yes. **Outcome variables:** PAE.	No statistically significant differences found between AG and TDG in accuracy; p not reported.	In line with CA and MA.
Titeca, Roeyers, Ceulemans, et al. (2015)	AD, AM, V	30	5.98 (0.31)	104.83 (12.36)	CA, FSIQ (Wechsler), SES, sex	**Interval:** 0 – 100. **Stimuli:** Pen and paper task. For the AM task, the stimuli consisted of black dots on a white disc, controlled for perceptual variables. **Total number of trials:** 30. **Practice trials:** Yes. **Outcome variables:** PAE.	No statistically significant differences found between AG and TDG for total task (U = 315.00, p = .146) and for the separate formats (p > .05).	In line with CA and MA.
Titeca, Roeyers and Desoete (2015)	AD, AM, V	20	5.13 (0.33)	105.30 (13.90)	CA, FSIQ (Wechsler), SES, sex	**Interval:** 0 – 10. **Stimuli:** Pen and paper task. For the AM task, the stimuli consisted of black dots on a white disc, controlled for perceptual variables. **Total number of trials:** n/r. **Practice trials:** n/r. **Outcome variables:** PAE.	No statistically significant differences found between AG and TDG for accuracy for total task (U = 104.00, p = .170) and for separate formats (χ^2 (2) = 1.41, p = .494).	In line with CA and MA.

Note: n = 5. (AD): Arabic digit; (AM): Analogue Magnitude; (V): Verbal. n/r = not reported. n/a = not applicable. AG = Autistic group. TDG = Typically developing group.

4.4 Discussion

This systematic review included 114 studies reporting on 143 autistic samples and 220 measures of mathematical performance. The following sections discuss the main findings found with reference to the three RQs of the current study.

4.4.1 What basic processes and specific components of mathematics have been examined in autism?

The majority of the studies investigating mathematical skills in autism reported measures for arithmetic word problem abilities (*n* = 84). Next came calculation abilities (*n* = 61), overall mathematics achievement (*n* = 52), followed by studies focusing on enumeration (*n* = 10), magnitude comparison (*n* = 8), and number line skills (*n* = 5). It is worth noting that the most frequently studied component – i.e., arithmetic word problems – is an advanced mathematical skill (as described in Chapter 1). The high interest observed for the investigation of arithmetic word problem skills may be explained by the presence of the linguistic component in these tasks and by the fact that, historically, research in autism has focused on the impairments of the language domain (Fletcher-Watson & Happé, 2019). On the other hand, basic skills described by the literature on TD population as the domain-specific abilities serving as precursors of mathematical performance (Nogues & Dorneles, 2021), such as counting, subitizing, and magnitude comparison, were investigated the least. While the few studies investigating these basic mathematical skills in autistic samples with high FSIQ scores reported performance in line with CA and with MA (IQ), the investigation of such components should be further investigated to provide additional insights into how the basic building blocks of mathematical learning need to be supported to enable the development of more complex mathematical abilities in autism.

Most of the studies included in the systematic review used standardised assessments, with subscales that measured arithmetic word problems, calculation, and overall mathematics achievement. Ad hoc researcher-made tests were used mainly to assess enumeration, magnitude comparison and number line abilities. The constructs measured by the standardised assessments were complex, but this complexity was conveyed only through an overall score. In the future, the use of analyses on single items or subsets of items of the standardised tests might enable researchers to capture the construct's complexity and to explain it by clarifying which steps or items the autistic individuals found challenging, if any. Moreover, generally, standardised assessments focus on a narrow and quantitative interpretation of mathematical abilities and ignore the occurrence of behaviours that are indicative of mathematical competence – e.g., the use of specific strategies. For this reason, the combination of qualitative observations, such as the level of support needed by the autistic participants to complete a task, together with standardised assessment and the inclusion of qualitative findings in the field of mathematical cognition is desirable. This practice is already present in some single-case research and could be extended to other designs.

Some researchers reported the impossibility to complete the mathematical assessment with some autistic participants. However, no details were provided in relation to whether there was a specific aspect of the assessment that prevented its completion (e.g., length, format) and how the experimenter managed to overcome these challenges. The inclusion of this information would be useful for the research community to confirm whether the assessments used for the TD populations are appropriate to measure mathematical skills in autism, especially in case of ID.

Finally, as reported in Section 4.2.2, mathematical abilities were coded based on the assessment tool used instead of using the label used by the author of the study. It was observed that the inconsistent use of terms to refer to different basic processes and specific components of mathematics and the mismatch between the claimed measurement (i.e., the mathematical component named by the researcher) and the actual measurement (i.e., the component measured by the assessment used) was not unusual. Furthermore, only a minority of the studies included in the current systematic review justified the choice of the assessment used to measure mathematical performance with reference to either the age range or the mean IQ scores of their participants or by referring to previous studies, as it happened for the tools used to measure intellectual and cognitive abilities. In fact, while for the IQ measures some studies administered different assessments with a clear justification for the hierarchy used, this only happened once for the mathematical outcomes, where the choice of the mathematical assessment was linked to the IQ test that was conducted (Mayes & Calhoun, 2008). These practices reflect the relatively young stage of the field of mathematical cognition when it comes to NDCs (Nogues &

Dorneles, 2021). In future studies, the inclusion of a description of the mathematical components being assessed – instead of a general definition of “mathematical achievement” – together with a rationale for the choice of the assessment tool chosen is advisable.

4.4.2 What are the mathematical profiles in autism?

To answer this question the following paragraphs present a discussion of the findings organised by mathematical component. For each component, the findings are discussed separately for samples with mean IQ scores below and above the ID threshold.

Only six studies investigating arithmetic word problem skills included samples with mean IQ scores below the ID threshold. All but one of these samples reported mean arithmetic word problem scores below the lower end of the typical range. This highlights low arithmetic word problem skills for autistic individuals with ID. Unfortunately, none of these studies featured a TD control group design, hence it is not possible to determine whether arithmetic word problem skills in autistic individuals with ID are in line with CA and / or MA. The results from the studies investigating arithmetic word problem skills in samples with mean IQ scores above the ID threshold showed that their average performance was within the typical range, with three samples reporting average performance above the typical range. The analysis of the findings from case-control studies with a TD control group were mixed, showing performances both in line and not in line with CA and / or MA. However, when the performance of the autistic sample was not in line with the TD control group, only one study reported higher scores for the autistic samples without ID (Titeca, Roeyers, Loeys, et al., 2015). However, the autistic samples (mean CA_1 = 7.38; mean $FSIQ_1$ = 104.19; mean CA_2 = 9.28; mean $FSIQ_2$ = 105.62) were compared against normed scores rather than a control group matched on specific criteria, and the assessment used to measure arithmetic word problem abilities was not used by any other study included in the review, so measurement bias could not be ruled out. These results are in line with previous reviews of literature which reported average arithmetic word problem skills for autistic individuals without ID (Chiang & Lin, 2007; Tonizzi & Usai, 2023).

In line with the pattern described for the studies investigating arithmetic word problem skills, only seven studies investigated calculation skills in samples with mean IQ scores below the ID threshold. However, contrary to what was observed for the arithmetic word problem component, only half of these samples reported calculation skills below the typical range, while the other half reported calculation performance within the typical range band and below the standardised mean. This highlights that performance on calculation skills of autistic individuals with ID is mixed and should be further investigated. Unfortunately, none of these studies featured a TD control group design, hence it is not possible to determine whether calculation skills in autistic individuals with ID are in line with CA and / or MA. The results from the studies investigating calculation skills in samples with mean IQ scores above the ID threshold showed that their average performance was within the typical range and skewed towards lower scores, with only one sample reporting average performance above the typical range. The analysis of the findings from case-control studies with a TD control group were mixed, showing performances both in line and not in line with CA and / or MA. However, when the performance of the autistic sample was not in line with the TD control group, only one study reported higher scores for a young (*M* = 113.27, *SD* = 15.25) autistic group with high IQ scores (*M* = 113.27, *SD* = 15.25) (Iuculano et al., 2014). These results are in line with previous reviews of literature which reported average calculation skills for autistic individuals without ID (Tonizzi & Usai, 2023).

Only six studies investigating overall mathematics achievement included samples with mean IQ scores below the ID threshold. In line with the pattern observed for arithmetic word problem skills, all but one of these samples reported overall mathematics achievement below the typical range, indicating a poor performance. Unfortunately, none of these studies featured a TD control group design, hence it is not possible to determine whether overall mathematics achievement in autistic individuals with ID is in line with CA and / or MA. The results from the studies investigating overall mathematics achievement in samples with mean IQ scores above the ID threshold showed that their average performance was shifted towards lower scores, with some samples reporting average performance below the typical range. The number of studies with a TD control group-matched design was small and their findings mixed. However, all case-control studies reporting outcomes not in line with the TD control group reported lower scores for the autistic sample. It is worth noting that, within the assessments used to measure overall mathematics achievement, the ones where autistic participants reported higher scores were the ones where the final score is computed as a composite score between two separate scales (e.g., WIAT and WJ), while the lower scores were reported when overall mathematics achievement was measured through assessments characterised by alternating mixed tasks (e.g., TEMA and PIAT). With this in mind, lower scores might be due to the type of the assessment used, and, for example, to difficulties related to switching between different tasks rather than to a difficulty with the specific mathematical tasks. The small size and the small number of samples included in this subgroup limits the extent to which these observations can be generalised, and further research is needed in this area.

Enumeration skills were investigated only in case-control studies with a TD control group featuring autistic samples with mean FSIQ above the ID threshold. The findings reported measures of perceptual subitizing, conceptual subitizing, and counting skills based on accuracy in line with CA and MA. A few studies reporting measures of RT showed mixed outcomes, with some studies reporting the use of counting strategies with small numerosities. However, these mixed results may be explained by the way instructions were given to participants, and therefore should not detract from the evidence of enumeration skills being in line with CA and with MA in autistic individuals without ID.

Numerical magnitude processes were investigated only in case-control studies with a TD control group featuring autistic samples with mean FSIQ above the ID threshold. The only case-control study examining symbolic magnitude comparison skills reported abilities of the autistic sample to be in line with MA and CA. Findings from studies investigating non-symbolic magnitude comparison skills showed that autistic participants needed a larger ratio than TD for accurate numerosity discrimination and reported mixed results when using accuracy-based performance measures. These findings are in line with the findings reported in the narrative review by Dowker (2020).

Finally, the findings from the case-control studies with a TD control group investigating number line skills in autistic samples with mean FSIQ above the ID threshold reported a performance in line with CA and with MA.

In summary, the current systematic review provides a partial description of the mathematical profiles of autistic individuals with ID and a description of the mathematical profiles of autistic individuals without ID. Autistic individuals with ID report poor arithmetic word problem skills and overall mathematics achievement, and mixed calculation skills. Autistic individuals without ID report arithmetic word problem and calculation skills in the typical range, overall mathematics achievement shifted towards the lower end of the typical range, and enumeration skills, magnitude comparison skills and number line skills in line with CA and with MA. The limited number of available studies and the lack of studies with a matched-group design with a TD control group, made it not possible to fully address the question of whether autistic individuals with ID show a different mathematical profile than the ones without ID. Unfortunately, as reported by Russell et al. (2019) the exclusion of people with ID from autism research is not limited to the field of mathematical cognition and needs to be addressed. The limited representation of autistic individuals with ID in these studies might stem from challenges in recruitment, especially in relation to individuals who are less likely to attend mainstream schools (Whitby & Mancil, 2009). Additionally, difficulties in assessing mathematical abilities in autistic individuals with lower cognitive abilities or communication barriers using standard measures may contribute to this gap (Keen et al., 2016; Mayes & Calhoun, 2003b). However, mathematical abilities in autistic people with ID should be further investigated. A better understanding of mathematical abilities in this group could highlight the need for differentiated interventions. The recent surge in single-case research methods observed in the last ten years appears promising in addressing this gap, as these studies tend to generally include autistic participants with lower IQ scores.

Finally, these findings challenge some of the prejudices that come with the association of autism with strong mathematical outcomes and giftedness, and show a pattern of performance which is similar or lower than the TD population rather than exceptional mathematical abilities, even in autistic individuals without ID.

4.4.3 Do CA, MA, or autistic traits explain mathematical abilities in autism?

4.4.3.1 Development of mathematical abilities in autism

Findings from the current systematic review show that most of the research on mathematical skills in autism is based on cross-sectional data, with only a small number of studies adopted a longitudinal approach (*n* = 4). Moreover, among the few studies using a longitudinal design, only a subset conducted comprehensive longitudinal analyses on samples including children and adolescents (*n* = 2). Both studies focused on calculation skills and reported improvements over a 1-year period between 7-12 years and 8-13 years (May et al., 2015), and over a 9-year period, from ages 9 to 18 (Kim et al., 2018).

This systematic review tried to fill this gap by building developmental trajectories from the cross-sectional autistic samples taken from the included studies with autistic samples. Due to the limited number of studies where the autistic sample has mean IQ score below the ID threshold and due to the narrow age range of the samples, it was not possible to build cross-sectional developmental trajectories for this subgroup. Due to the nature of the scores reported in the studies it was possible to build developmental trajectories only for 3 mathematical components: arithmetic word problems, calculation, and overall mathematics achievement.

Visual examination of these cross-sectional developmental trajectories shows that, for autistic samples with mean IQ score above the ID threshold, arithmetic word problem and calculation skills develop in line with CA. Visual examination of the developmental trajectory of overall mathematics achievement for autistic samples with mean IQ score above the ID threshold showed that this component does not develop in line with CA. This could be due to the smaller number of autistic samples used to plot the developmental trajectory for mathematical achievement (n = 25) compared to the other specific components of mathematics (arithmetic word problem: n = 62; calculation: n = 39). Also, this finding might be explained by the fact that the CA range for mathematical achievement is narrower than the others, hindering the possibility to identify a trend.

4.4.3.2 Relationship with IQ measures

The results from the studies investigating arithmetic word problems, calculation and overall mathematics achievement show a strong and positive relationship between general intellectual abilities and the development of these specific mathematical components.

This relationship between mathematical components and overall intellectual abilities implies that interventions or educational strategies aimed at enhancing general intellectual abilities might positively impact the development of specific mathematical skills, and the other way around. Conversely, challenges or delays in overall intellectual development might affect the pace or the extent of improvement of these mathematical abilities. Future studies are needed to further understand this relationship and increase the effectiveness of interventions to support mathematical development in autistic populations.

4.4.3.3 Relationship with level of autistic traits

It was not possible to determine whether different levels of autistic traits impact mathematical performance because the studies did not provide enough information related to the diagnosis as well as to the diagnostic criteria used to assess autistic participants, and because of the use of a wide range of different assessments to measure autistic traits, each one using different clinical cut offs.

4.4.4 Limitations

The main limitation of this study is that the results are reported at group means level instead of at the individual data point level. So, for example, the average scores in calculation could be explained by a large percentage of participants in the sample having average calculation skills or by some individuals having low calculation abilities, some having average abilities, and others having strong calculation abilities. Because the data is presented at group level, the current review cannot explain which is the case – although inspection of the standard deviation of each sample can give a sense of the variability within the sample – and all the findings should be treated with caution.

Another limitation of this study comes from the fact that the findings should be interpreted through temporal and spatial lenses that could not be provided to the reader. Even if a brief description of the shift of diagnostic criteria has been provided in Chapter 1, the findings reported in the results section do not carry this information with them and this might lead to erroneous interpretations. For example, because diagnostic criteria were much narrower in the 1960s and ’70s than today, the picture we see in autistic 60-year-olds diagnosed in childhood is likely to be quite different from the one of a newly diagnosed 60-year-old, or for today’s autistic children when they reach 60 years. As for the spatial contextual information, this review only provides information on the country where the study was conducted. The lack of detailed information provided by the single studies about the educational system and the implementation of educational policies where the studies took place limits our understanding of these findings.

4.5 Conclusion

This chapter presented a systematic review focusing on the research investigating mathematical abilities in the autistic population and included 113 studies published between 1970 and 2023. The main aims were to create a comprehensive overview of existing literature to address the fragmented nature of research in this area and to provide a description of mathematical profiles in autism.

The results can be summarised in four main points:

1) Autistic individuals with ID reported poor arithmetic word problem skills and overall mathematics achievement, and mixed calculation skills.
2) Autistic individuals without ID reported arithmetic word problem and calculation skills within the typical range, overall mathematics achievement shifted towards the lower end of the typical range, and enumeration skills, magnitude comparison skills and number line skills in line with CA and in line with MA.
3) Arithmetic word problem and calculation skills develop in line with CA and IQ scores. Overall mathematics achievement develops in line with IQ scores only.
4) The association of autism with strong mathematical outcomes and giftedness was not found in the studies included in this systematic review.

Chapter 5: Teaching mathematics to primary school students with Down syndrome: Findings from focus groups

The main aim of this qualitative study was to explore the maths school environment of primary school students with Down Syndrome (DS). This study employed focus group interviews and reflexive thematic analysis to explore the experiences of primary school educators working in English mainstream settings when supporting mathematical abilities of a student with DS. Specifically, this study investigated the challenges faced by the participants and the teaching strategies and learning resources employed. Finally, this study aimed at exploring whether the teaching strategies reported by the educators aligned with the cognitive profile of individuals with DS, as it is described by the literature.

5.1 Background and rationale of the study

Despite the increasing recognition of the needs of learners with DS and how to support them in the classroom (Hargreaves et al., 2021), literature on teaching mathematics to students with DS is scant and the area largely unexplored (Clarke & Faragher, 2015). Indeed, while it is widely acknowledged that teachers and teaching assistants (TA)[20] play a significant role in motivating and supporting their students' mathematical learning (Fauzy & Hosshan, 2023), there are few studies investigating educators' needs, their experiences in the mathematics inclusive classroom and the challenges they face. Furthermore, research offers little evidence about the teaching practices and the learning materials used in the maths inclusive classroom and their effectiveness.

The following sections present a brief overview of the current literature. The concluding paragraph presents the aims of the current study.

[20] The specific title used to define the role of the adult supporting the teacher in various aspects of classroom management varies between schools and countries and has also evolved over time. In this thesis, the term to define this role is Teaching Assistant (TA).

5.1.1 Teaching and learning experiences in the maths inclusive classroom

A few studies have explored teaching and learning experiences in the maths inclusive classroom involving a student with DS. These studies employed various methodologies and school settings.

Two studies investigated the experiences in the maths inclusive classroom through observation: one study reported on students' experiences in a primary school in New Zealand (Rietveld, 2005) and another study reported on teacher's experiences in a secondary school in Italy (Monari Martinez & Benedetti, 2011). Rietveld (2005) described the learning experiences of three students with DS aged between 5 and 6 years. Their findings described the lack of effective mathematics instructions, the different level of abilities between the student with DS and the rest of the class, inconsistent display of mathematical skills of students with DS, and challenging behaviours such as slow or non-existent task persistence, opting out strategies, and throwing or scattering learning materials as barriers to learning. The study by Monari Martinez and Benedetti (2011) described the experiences of learning algebra of two students with DS in the mainstream secondary education setting. The results of this study highlighted that TAs were not prepared to teach such specific topics. Hence, the collaboration of the class teacher with their TA was described as a key element in carrying out mathematical interventions for these students.

The experience of educators teaching mathematics in the inclusive classroom has been explored through surveys and interviews with educators in two studies, both based in Australia. The study by Clarke and Faragher (2015) used a survey to investigate the needs identified at the beginning of the school year by 16 primary school teachers and 12 TAs when supporting a student with DS in their class for the first time. Their results reported that participants made a strong endorsement of inclusion as an appropriate practice for teaching mathematics in primary education. However, the study also highlighted the importance of acknowledging the complexities of teaching mathematics within this context, especially in relation to session planning, learning resources, and teaching strategies to be employed. The second phase of this study reported participants' responses at the end of the school year and highlighted an increased emphasis of the participants on challenges linked to students' attitude and behaviour, rather than their academic achievement (Clarke & Faragher, 2015). Finally, Clarke and Faragher (2015) reported that participants' responses highlighted the complexity of measuring student's knowledge and predicting their responses to mathematics lessons on any given day. In the second study investigating the experience of educators teaching mathematics in the inclusive classroom through interviews, Faragher and Clarke (2020) observed 15 Australian primary school classrooms including a student with DS during maths classes. The themes which emerged from the qualitative analysis of the interviews with teachers (*n* = 19) and TAs (*n* = 19) focused on the challenges faced during teaching and planning. These included 1) how to foster independence in learning, 2) student's readiness to learn advanced mathematical topics, 3) adjustments to the lessons to support inclusion, 4) social and behavioural challenges, and 5) management of staff and resources.

Finally, a recent study by Hargreaves et al. (2021) investigated the educational experiences of primary and secondary school students with DS in the UK through a parental online survey (*n* = 569). The results of the study highlighted that in UK most of the students with DS attend mainstream schools (95%) and that they receive additional support for a large part of the school day, with over half of the students receiving over 30 hours of 1:1 support each week[21]. Also, the study showed that over 80% of the students with DS were taught maths in the classroom alongside their peers and that in general most of the students participated in academic and social activities but were commonly not accessing all opportunities, such as foreign language classes. Furthermore, findings showed high levels of collaboration between parents and educators, characterised by regular meetings that could take the form of either informal gatherings or more formal opportunities to collaborate.

5.1.2 Teaching practices employed in the maths inclusive classroom

A few studies have reported the practices used by teachers when supporting a student with DS in the inclusive classroom. These studies highlight the teaching strategies employed and the adjustments made by the teachers to include a student with DS in their class.

[21] English mainstream schools are expected to provide a compulsory week of at least 32.5 hours (Department for Education, 2023).

Kabashi and Kaczmarek (2019) observed a 5-year-old child with DS in an American inclusive classroom and described the various strategies that were effectively used by the educators to facilitate the child’s participation, transitioning between activities, academic outcomes, and social development. Within the strategies that proved to be successful in supporting the student’s academic outcome, the authors reported breaking down tasks into small and incremental steps paired with the use of visuals to represent each step, choice making, and scaffolding. Choice making was described as the act of the student selecting an activity from several familiar options at a particular moment in time. To be effective, the authors reported that the choices needed to be encouraging and meaningful to the student. As for the scaffolding strategies, these included the teacher providing clues, demonstrations, and modelling using both verbal feedback and gestures.

Faragher et al. (2017) reported the adjustments made by a teacher to the curriculum to include an 11-year-old student with DS in the mathematics class in Australia. These included the use of technology (Talking Calculator app), enlarged worksheets, and the involvement of the TA in the session.

Finally, an article by Bird (2016) reported a list of broad methods that were found to support success of students with DS, which included: 1) setting up and scaffolding situations for learning by imitation, 2) daily practice of specific skills built into engaging class activities, 3) planned activities that become part of the child’s routine, 4) support for positive behaviour, and 5) adaptations that use learning strengths and support areas of difficulty.

Overall, the teaching strategies and methods described in these studies emphasise the value of student engagement, scaffolding of learning through demonstration and modelling, and adaptation of the curriculum based on the areas of strengths and difficulties of students with DS. While these taxonomies acknowledge areas of strengths and areas of difficulties of students with DS, they do not explicitly link the recommended teaching strategies to the cognitive profile of students with DS (refer to Section 1.3.1). Moreover, rather than providing a rationale on why some teaching strategies are effective for supporting students with DS, these taxonomies offer educators an action-oriented guide and tend to focus on student’s performance only. The prescriptive nature of these taxonomies highlights a tendency of providing educators with partial knowledge, leaving them without the necessary tools to fully appreciate the value of the educational practices which they employ.

5.1.3 Current study

Previous literature offers valuable insights into the learning and teaching experiences in inclusive maths classrooms including a student with DS. However, these studies are diverse in terms of the educational context, as they encompass teachers' and students' experiences from nursery to secondary school settings and cover three different continents. These factors are crucial to consider when comparing studies and drawing conclusions, especially in educational research with neurodivergent populations. Firstly, the level and the type of support provided can vary significantly across different educational levels. Secondly, it is important to recognise the unique policies, practices, and cultural contexts inherent in each country's educational system because these factors can have a profound impact on the experiences and outcomes of teachers and students.

Moreover, the existing literature underscores that there is a gap in knowledge regarding the teaching strategies and the learning resources employed by educators. This gap has significant implications, as it limits our ability to comprehend classroom dynamics and develop effective educational programs. Without insight into what happens within the classroom and how students are supported, it becomes challenging to devise meaningful and impactful educational interventions.

More importantly, there is no research on the learning and teaching experiences in inclusive maths classrooms and on the practices employed in the UK. In fact, the study conducted by Hargreaves et al. (2021), albeit investigating inclusive practices in the UK, lacked specific details regarding the basic processes and specific components of mathematics of the education of students with DS. Moreover, as this study was based on a parental survey, the educators' viewpoint was missing. These considerations reinforce the need for the current study, particularly given the significant increase in the enrolment of students with DS in mainstream primary schools across the UK over the last 30 years (Hargreaves et al., 2021; Van Herwegen et al., 2018).

The current study employed focus group interviews to explore the experiences of teachers and TAs (educators here onwards) when teaching mathematics to primary school students with DS in mainstream schools in England. This study aimed to answer the following research questions (RQs):

1. What are the challenges related to teaching mathematics to a student with DS in the inclusive classroom?
2. What are the teaching strategies and resources employed in the maths inclusive classroom?
3. Are the teaching strategies used by educators aligned with the cognitive profile of students with DS?

5.2 Participants and Methods

5.2.1 Participants

Participants were recruited through purposive sampling, using online networks and social media platforms frequented by parents of children with DS and professionals supporting those students. Additionally, the recruitment process tapped into the personal network of the researcher. Specific inclusion criteria included (a) being a primary school class teacher or TA working in an English mainstream school; (b) supporting a student with DS at the time.

The original sample consisted of six participants. However, one participant was excluded because they did not attend all the sessions. Hence, the final sample counted five participants. These included three class teachers and two TAs from primary mainstream schools located in England. All the participants were female. Their level of experience in supporting students with DS was varied. One participant had a previous experience supporting a student with DS in a different teaching role, one participant had been working with the same student with DS for 5 years, and the remaining participants were going through their first year of supporting a student with DS in their class (Table 5.1). Two participants reported to have received some guidance on how to support mathematical abilities of students with DS, and one participant reported to have received some specific training to support mathematical learning in general, but not specifically for the DS learning profile. As reported in Table 5.1, most of the schools adopted the White Rose programme and the Maths for Life programme. Only one school had adopted more than one mathematical programme as part of their curriculum.

Table 5.1. Characteristics of participants.

	Role	Sex	Teaching experience supporting students with DS	Maths Programme
P1	TA	Female	First year supporting a student with DS	Maths for Life
P2	Teacher	Female	First year supporting a student with DS	White Rose
P3	Teacher	Female	Previous experience as 1:1 TA of a student with DS	White Rose; Mastering Number Approach; Mathletics
P4	TA	Female	Has been supporting this student since Reception year	Maths for Life
P5	Teacher	Female	First year supporting a student with DS	White Rose

5.2.2 Materials and procedure

Participants took part in three 60-minute online sessions during the summer term, that is the last of the three terms of the school year[22]. Therefore, by the first session, all participants had been supporting a student with DS for at least two-thirds of the school year, gaining substantial experience that they could share with the other educators.

Table 5.2 provides an overview of the main topics discussed in each session alongside the corresponding activities.

Table 5.2. Topics covered and activities ran during the focus group sessions.

Date	Main topic	Activities
		Reflective Journal: Demographic survey
Session 1 (April 2022)	Teaching challenges and mathematical learning targets	• Teaching maths to students with DS (quiz) • What maths learning targets are you working on? (whole group)
		Reflective Journal: Task 1
Session 2 (May 2022)	Learning resources and teaching strategies	• Learning resources (whole group) • Successful teaching strategies (whole group)
		Reflective Journal: Task 2
Session 3 (June 2022)	Lesson plan	• Write a maths lesson plan (small groups) • Successful teaching strategies (whole group) • Final feedback

Prior to the first session, participants were sent a link to access their individual reflective journal. Each reflective journal included a section with an overview of the project and contact details for the primary researcher, and a survey to collect their demographic details (including name, sex, current role, and prior experience in supporting students with DS). Additionally, the journal comprised two tasks and a

[22] In English primary schools, the length of the summer term can vary in length depending on the school and local authority. It usually begins after the Easter break and lasts until the end of the school year, typically in July.

final feedback form. Participants were encouraged to complete the demographic section before the first session and to work on each of the two tasks between the workshop sessions. These tasks were designed to prompt reflection on specific aspects of their work and were intended to facilitate the discussion during the following focus group sessions. Task 1 asked participants to provide an example of maths learning resource that they found useful and one example of a maths learning resource that was not effective in supporting the student they were working with. Task 2 asked participants to track and report the maths teaching activities that took place with the student with DS on a daily basis. Between Session 2 and Session 3 participants were asked to report for each day the type of learning resources used, the setting where mathematical learning took place, the challenges faced, and their level of confidence when teaching (Figure 5.1).

Figure 5.1. Replica of task 2 included in the reflective journal.

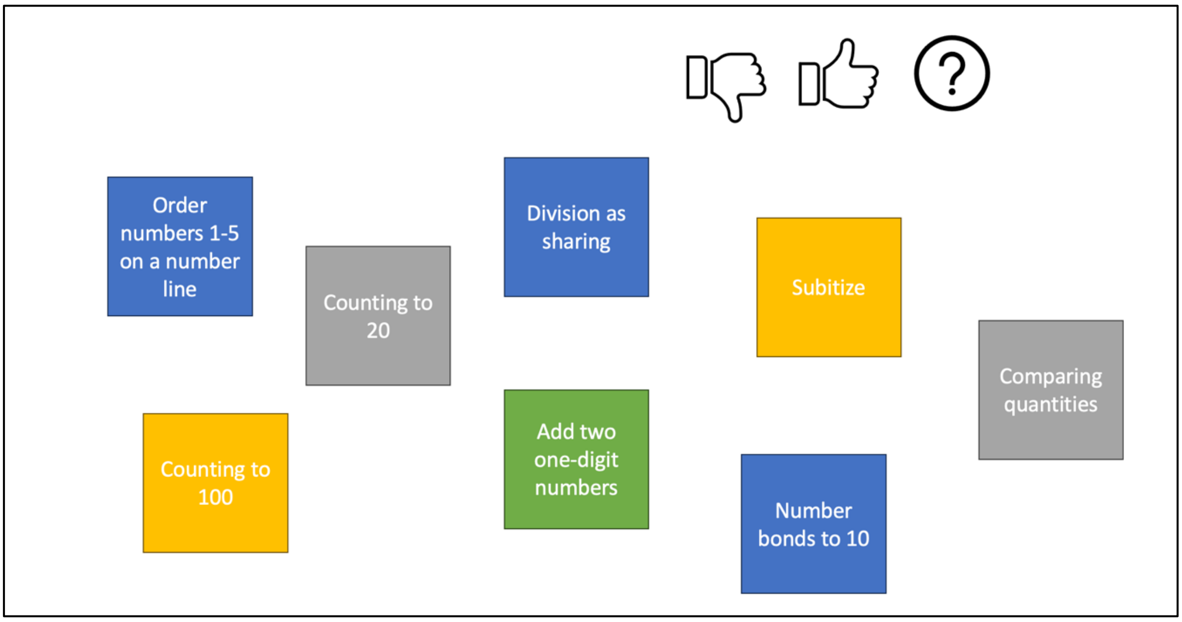
	Mon 23 May	Tue 24 May
Which resource did you use today?		
In what setting did you use it?		
How did you feel about using the resource? Were there any challenges or benefits of its implementation?		
How confident did you feel when using the resources? Choose the face that best describe your level of confidence and delete the others		

Before starting each online focus group, participants were informed of their voluntary involvement and were asked to provide oral consent before starting the recording. Ground rules were established, emphasising confidentiality and the focus of the meeting on sharing participants' experiences. Participants were encouraged to freely contribute to the discussion and to share additional thoughts by sending an email to the researcher after the session. Then, participants were provided with a brief overview of the purpose and of the structure of the session.

5.2.2.1 Focus group 1

The discussions and activities held in Session 1 focused on the students' experience of learning mathematics, the mathematical learning targets addressed, and the strategies and resources used by participants in the inclusive classroom. Participants were asked to provide information on their students (e.g., demographic information and their students' attitude towards mathematics and learning in general), on the typical setting adopted by participants to teach mathematics to the student with DS (e.g., face to face or on-line, part of the day, as part of the class or in a separate room), and the learning targets addressed during the current year. For the first activity, participants were presented with a shared digital whiteboard (Figure 5.2) listing mathematical learning targets taken from the English national curriculum (Department for Education, 2014). Due to time constraints and to the structure of the English national curriculum, the list only included the mathematical learning targets that are addressed during the first part of the school year, which focus on the maths areas of "Number and place value", "Addition and subtraction", and "Multiplication and division".

Figure 5.2. Replica of the shared digital whiteboard used to support the group activity during Session 1.



Participants were asked to work independently on the shared board and to identify all the targets relevant to the students they were supporting for the current school year. Additionally, they were asked to add any learning target that was missing on the board and to amend the learning targets provided, for example by breaking them down into smaller learning targets to better describe their students' mathematical progress. Then, participants were asked to indicate the learning targets mastered by their student (thumbs-up), the learning targets that they felt were challenging for their students (thumbs-down), and the learning targets they were unsure whether their student had a secure grasp on (question mark). Finally, participants were asked to provide some examples of resources and strategies that they were using to support specific learning objectives.

Due to recruitment issues, Session 1 was delivered twice with two subgroups of participants ($n_1 = 4$; $n_2 = 2$).

5.2.2.2 Focus group 2

During Session 2, participants engaged in two activities focused on discussing the learning resources used to support their students' mathematical skills and their teaching practices. Initially, participants shared and reflected on their responses from task 1 of the reflective journal, which asked them to describe two learning resources, one that was useful and one that they felt was not working for their student. To facilitate and structure the discussion, participants were asked to complete a real-time survey and to present the learning resources they selected to the group, highlighting how they were utilised, and why they found them effective or not for their students. The second activity involved participants reviewing a list of teaching strategies created by the researcher and based on the discussion held in the previous focus group session. The full list included: repetition, concrete-pictorial-abstract, modelling, adapting activities, use of calculator, "going back" to the basics, use of signs, collaboration with family, use of routines, collaboration with teaching team. Participants were asked to reflect on these strategies, to provide examples of their use, and to explain their perceived effectiveness. Additionally, participants were asked to identify any strategy they typically used and that was missing.

5.2.2.3 Focus group 3

During the last session participants engaged in a small-group activity where they were asked to plan a maths lesson, to comment on the revised list of teaching strategies created by the researcher, and to provide feedback on their overall experience of taking part in the research project.

For the first activity, participants were presented with a template showing a lesson plan developed by the researcher (Figure 5.3). The template presented the typical structure of a lesson plan used in English primary schools. In addition, it presented five sections, each one addressing the different areas that were identified by the researcher through reflexive thematic analysis of the data collected during the previous sessions, i.e., attention, memory, language, motor skills and motivation. At the bottom of the lesson plan, four boxes were included for participants to add notes and reflections about the session, resources to use, key vocabulary and the evaluation of the students' performance. Participants were asked to use the template and to work in small groups to plan an activity to support counting. Following the activity, participants were asked to provide their feedback on the usefulness of the template, potential applications, and improvements. Then, participants were asked to reflect on the revised list of strategies created by the researcher, to point out the strategies they were not familiar with, and to consider how they would implement them in their teaching. The researcher closed the session asking the participants to provide their feedback on their overall experience in taking part in the research study.

Figure 5.3. Lesson plan template.

<table>
<tr><td>Year Group:</td><td>Term (week):</td><td colspan="2">Time:</td></tr>
<tr><td colspan="2">Teaching Aim:</td><td colspan="2">Children Aim ("I can" statement):</td></tr>
<tr><td colspan="4">Prior Learning Recap:</td></tr>
<tr><td colspan="2" rowspan="6">Main Activity:</td><td colspan="2">How to support Attention
• Where is the child playing this activity?
• What resources will you use?
• What time of the day are you planning to play this activity?
• How long is the activity?</td></tr>
<tr><td colspan="2">How to support Memory (Short Term Memory and Working Memory)
• What numbers are you working with?
• Will you model the activity? How will you do that?
• Can you break down the task in small steps?</td></tr>
<tr><td colspan="2">How to support Language
• What verbal instructions are you using?</td></tr>
<tr><td colspan="2">How to support Motor skills
• What resources are you using?</td></tr>
<tr><td colspan="2">How to support Motivation
• How can you personalize the resources?</td></tr>
<tr><td colspan="2">How to support….</td></tr>
<tr><td colspan="4">Plenary:</td></tr>
<tr><td>Resources:</td><td>Key Vocabulary:</td><td colspan="2">Evaluation/ Assessment of Learning / Success Criteria:</td></tr>
</table>

At the end of the project, each participant received a £75 Amazon voucher as a token for their time and contribution.

All participants gave written informed consent to take part in the research and for the focus group sessions to be video recorded. Ethical approval for this study was granted by IOE, UCL's Faculty of Education and Society (data protection registration number: Z6364106/2022/02/100).

This research project was funded by a public engagement bursary from UCL Culture.

5.2.3 Data analysis

All the focus group sessions were video recorded through the Zoom platform. The interviews were transcribed verbatim and imported into NVivo 2020, a qualitative data analysis software for coding. In the interest of confidentiality, identifiable details, such as the name or the location of the school, were excluded from the transcripts, and pseudonyms were assigned to participants' and students' names before starting data analysis.

The data collected from the second task included in the reflective journal was excluded from the analysis due to incomplete participation. In fact, only three participants completed the task, and this was done inconsistently (P1: 5 weeks; P4: 3 weeks; P5: 4 weeks).

To explore the experiences of teaching maths to students with DS in the English inclusive classroom, a reflexive thematic analysis approach was employed. This involved conducting the initial analysis through an open coding method, which entailed the categorization of text from the data without an existing framework. The coding framework was developed by generating themes from these sections of text, then refined into sub-themes. According to Braun and Clarke (2019)'s guidelines, a six-step process was followed to identify the final themes. These involved first the familiarisation with the data collected. Then, the identification through initial coding of the overarching themes related to the teaching experiences of the educators. Finally, further refinement and labelling of the main themes and related sub-themes before proceeding with the write up of the analysis presented in the results section.

To test the reliability of the coding, 25% of the transcripts were coded by a second researcher. Cohen's kappa showed a substantial level of agreement between the two researchers' coding of the transcripts, with a value of Cohen's kappa of 0.73 (Landis & Koch, 1977). Where differences were found, the coding was reviewed to reach a consensual agreement.

5.3 Results

5.3.1 Description of the students from the educators' perspective

Table 5.3 describes the students who were being supported by participants. It shows sex, year group, and level of maths of the student, their attitude towards mathematics as described by the participant and the learning setting in the current year.

Participants were supporting three males and two females. One student was attending Reception (approximate CA: 4 – 5 years), two students were attending Y2 (approximate CA: 6 – 7 years), and two students were attending Y4 (approximate CA: 8 – 9 years). Two students (S1 and S3) were working at their year group level in mathematics, while the other students were working at a lower level than expected for their age.

When asked to describe their students' attitude towards learning mathematics, all participants reported that their student had an overall positive attitude and was motivated towards learning mathematics, even if some of them found it challenging.

All students received in-person instructions at school, apart from one student who attended online classes during the initial part of the year. However, this student later transitioned to in-person classes as well. When describing their student's behaviour in the class, only one participant reported encountering challenging behaviours, particularly when the student was expected to work with adults who were not the designated one-to-one support staff assigned to them.

Table 5.3. Characteristics of the students supported by the participants.

Student	Sex	Year group	Working at	Educators' view on student's attitude towards maths	Learning Setting
S1	Male	Reception (pre-KS)	pre-KS	"He's so keen and so eager and wants to please" (P1).	School
S2	Male	Y2 (KS1)	pre-KS	"[Maths] is an area that he finds quite challenging and isn't always very motivated to do" (P2).	School
S3	Female	Y2 (KS1)	KS1	"She is just so intrinsically motivated, and this is really great learning behaviour, so… I think she	School

				does enjoy it, but I think she feels like she is a stronger reader, and she enjoys English more” (P3).	
S4	Male	Y4 (KS2)	KS1	“He knows he is at school and the rest of the class is doing maths […] so, he sees the other children doing maths. And he also knows as part of a routine that he does maths and literacy, at a certain time” (P4).	School
S5	Female	Y4 (KS2)	KS1	“I think she thinks she likes it [maths] when she has got all the manipulatives” (P5).	Remote and school

Note: The numbering system employed in this table is used consistently throughout the Chapter when referring to the student (S) and to the participant (P). Pre-KS: pre-key stage standards. KS1: Key Stage 1.

Participants were asked to report the most common way their student was taught mathematics at school and to reflect on its effectiveness (Table 5.4). Although all participants acknowledged the importance and benefits of including the student with DS in their classroom, some mentioned instances where the student worked outside the classroom. This occurred specifically in situations such as noisy lessons where the student required a quiet environment for learning or when the class focused on challenging topics that proved too demanding for the student. One educator reported that they would occasionally take their student out of the classroom to run activities like role-playing. This suggests a dynamic approach to learning that involves adapting the environment and / or activities beyond the traditional classroom setting.

Table 5.4. Description of how the students were taught maths.

Student	Common way of learning maths
S1	“They normally do it [maths] in the mornings and he's taken away and does something else, and then I take him in the afternoons, and we do our own maths together. So, he's not following the same [curriculum] as the rest [of the class] and he has been left in there, you know, at the very beginning, just to see how he would cope with it and it's too much… it’s too much. Obviously, he really needs the one to one and just the quiet surrounding, so in that way it's better” (P1).
S2	“[He] does a lot of his work in class with a one-to-one support… where we can we get him involved in the group activities because he's very motivated by being, you know, working with his peers […] but if it's very noisy lesson sometimes he'll go out just to have a bit of quiet space, but he really wants to be in class” (P2).
S3	“[the TA] will only take her out if she's too distracted by the noise, because sometimes I’ll bring the children back to the carpet and we'll do a mini plenary or challenge further, and that will be a bit above of her head. So, she will just continue working with her and come to a quiet space, but for the input it's always within the class” (P3).
S4	“He is sitting at his desk in the classroom but certainly we can go out and we'll find a space and [for example] play shops” (P4).

S5	"She will be out with her one-to-one because she can't access any of what we're doing now. So, having her working with someone just isn't really an option for maths… other subjects are fine, it is great that she gets that interaction that she really needs" (P5).

Table 5.5 provides a breakdown of the mathematical learning objectives that the participants reported for the current year, separately for the three mathematical areas of the English national curriculum discussed during the workshop sessions.

Table 5.5. Mathematical learning objectives reported by participants for their student.

Number and place value	**S1**	**S2**	**S3**	**S4**	**S5**
Quantity comparison (more and less)	Green	Grey	Green	Green	Green
Subitizing	Grey	Grey	Green	Grey	Grey
Counting (cardinality principle)	Green	Red	Green	Green	Green
Verbal counting to 10	Red	Red	Green	Green	Green
Verbal counting to 20	Red	Red	Green	Green	Green
Digit recognition 1 - 10	Green	Green	Green	Green	Green
Digit recognition 1 - 20	Red	Red	Green	Green	Green
Ordering numbers 1-5	Green	Green	Green	Green	Green
Ordering numbers 1-10	Red	Green	Green	Green	Green
Ordering numbers 1-20	Red	Red	Green	Green	Green
Place value (3-digit numbers)	Red	Red	Red	Red	Green

Addition and subtraction	**S1**	**S2**	**S3**	**S4**	**S5**
Solve addition with support of concrete resources (with no adult support)	Red	Green	Green	Green	Green
Number bonds to 10	Red	Red	Red	Red	Red
Solve subtraction with support of concrete resources and modelling	Red	Red	Green	Green	Green
Solve subtraction independently	Red	Red	Red	Green	Red

Multiplication and division	**S1**	**S2**	**S3**	**S4**	**S5**
Halving	Red	Red	Red	Red	Red
Doubling	Red	Red	Green	Green	Green
Doing sharing activities	Red	Red	Green	Green	Green

Note: Secure learning targets are represented in green. Learning targets that the student has not grasped securely yet or that the student is not working on are represented in red. Learning targets for whom the participant did not report any information are reported in grey.

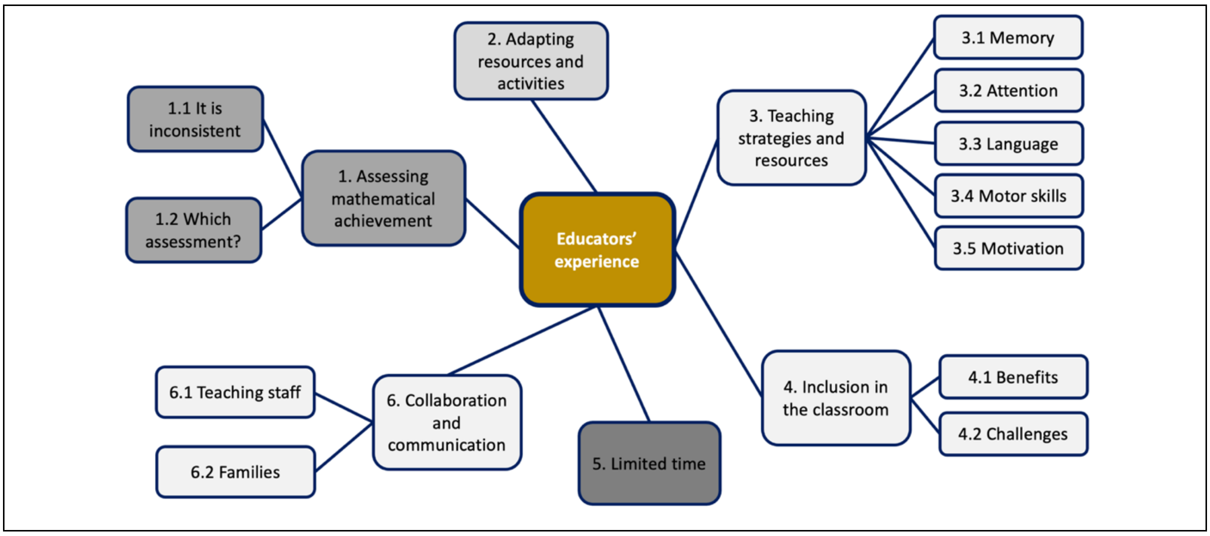
Table 5.5 shows that educators identified a higher number of learning objectives for the "Number and place value" area than for the other two areas. The objectives included in the “Number and place value” area are broken down into smaller learning targets (e.g., verbal counting to 10, and verbal counting to 20). This practice was not observed in relation to the learning targets included in the remaining areas. Most of the learning objectives within the “Number and place value” area were successfully achieved by students in the current academic year, particularly among the older ones. Only one educator commented on their student's subitizing skills (P3). Conversely, fewer learning targets were identified as successful for the other mathematics areas. Most of the mentioned targets in these areas were described as still emerging and requiring high levels of adult support.

Overall, looking at Table 5.5, a stable profile emerges for all the three mathematical areas, where all the targets reported as secure for the younger students are also reported as secure for the older ones. Moreover, Table 5.5 shows progression of mathematical learning, in that the profiles of older students are characterised by more advanced targets being secured.

5.3.2 Teaching experience of supporting mathematical learning of a student with Down syndrome

Through reflexive thematic analysis, six main themes and related sub-themes were identified under the “Educators’ experience” overarching theme (Figure 5.4). The main themes identified were: 1) assessing mathematical achievement, 2) adapting resources and activities, 3) teaching strategies and resources, 4) inclusion in the classroom, 5) limited time, and 6) collaboration and communication.

Figure 5.4. Thematic map of participants’ experience in supporting their students’ mathematical learning.



5.3.2.1 Assessing mathematical achievement

All participants agreed that one of the main challenges that they had to face was to assess mathematical achievement of the student they were supporting and “to get a clear picture of where they are”.

When sharing their experiences around this challenge, participants explained that mathematical knowledge is very complex to assess because the students are “quite inconsistent with [their] responses” (P2) and that “it's really hard to know what [the student] is secure with, or whether is just what [they] remembered” (P5), with one educator adding “he's quite a joker, so it's quite hard to tell when he’s actually confused, or when he's trying to sort of trick us” (P2). Participants explained that the difficulty in assessing mathematical knowledge is a consequence of multiple factors, which included the inconsistent display of mathematical abilities depending on the day, or time of the day the activity was run. For example, one educator explained that performance on a mathematical task “can depend on the day, how [the student is] feeling, any little thing can change” (P2). Moreover, participants highlighted the significance of discussions with families to gauge the student's knowledge gaps, with one reporting: “we have a lot of discussion with his parents as well, just to find out what he does know what he doesn't know” (P4) and another saying: “I just know from mum what she can do” (P5). Educators also emphasised that daily interaction and observation were instrumental in determining the student's knowledge and needs. One participant reported that the difficulty of assessing mathematical abilities was related to the lack of guidance on the assessment tools that teachers should use: “she is accessing the curriculum, but I have also been given branch maps. I just feel like I have got lots of forms of assessment which should give me lots of ideas of how to support her and I never know which one to prioritise” (P3).

Finally, educators agreed that one of the main consequences of this challenge was the risk of blocking or slowing down their students’ mathematical progress as “we have to keep revisiting and revisiting” (P5) and “we end up going kind of back [on] everything” (P3).

5.3.2.2 Adapting resources and activities

All participants reported that they were involved in tailoring learning activities and resources, ensuring differentiation not only in relation to the difficulty levels, but also by incorporating personalised elements to motivate and engage their student. For example, one participant explained how they adapted an activity that involved using a balance scale by having the student measure only “light objects because he's still working on counting between 1 and 10. So, it is kind of like pitching it where he can be developing his skills but still accessing what everyone else is doing” (P2). Additionally, another participant reported that the student they were supporting was “really into swashbucklers. So, I've got to print off some jewels and have physical activities outside, when we have time, where we can go jewel hunting”. Finally, a few participants mentioned that their differentiation approach included using the same worksheet used by the student’s peers but selecting only specific questions or incorporating the use of manipulatives for further support.

5.3.2.3 Teaching strategies and learning resources

During the three focus group sessions, participants engaged in various activities aimed at fostering reflection on their teaching practices and why they were using them. Table 5.6 displays a list of 20 teaching strategies discussed by the educators during these sessions. The thematic analysis identified five domain-general processes as targets of these strategies. These included memory (n = 8), motivation (n = 8), attention (n = 6), language (n = 4), and motor skills (n = 3). While some teaching strategies only targeted one cognitive process (n = 11), there were nine strategies that were used by the educators to support more than one domain-general process.

Table 5.6. Teaching strategies used by the participants.

Strategy	Domain-general process targeted
Allow for repetition	Memory
Be aware of the language that you use	Language
Break down activities in small steps and be flexible	Motivation, Attention
Collaborate with family	Memory, Language
Design the learning environment	Attention
Do not take anything for granted	Memory
Introduce a routine	Memory
Introduce pre-teaching	Memory, Motivation
Keep activities short and snappy	Attention
Make learning playful	Motivation
Make learning practical	Motivation
Personalise learning resources	Motivation
Pitch to the student’s “right level”	Motivation, Language
Plan meaningful learning interactions	Motivation

Use accessible resources	Language, Motor skills
Use modelling	Memory, Attention
Use signs and gestures	Memory
Use tablets, maths apps and the whiteboard	Attention, Motor skills
Use videos	Memory, Motivation
Use visual and concrete resources	Attention, Motor skills

During Session 2, participants were asked to report on the resources used to support the mathematical development of the student with DS they were working with. Table 5.7 presents a list of learning resources that participants found useful, along with the corresponding learning targets they were addressing while utilising these resources. The resources reported by participants included both concrete manipulatives and digital resources.

Table 5.7. Learning resources reported by participants as useful and related learning targets.

Participant	Resource	Learning target
P1	Numicon	Numbers 1-5
P2	Personalised counting cards showing digits and quantities (1-10)	Number recognition, counting and subitizing
P3	Rekenrek (concrete and virtual versions)	Number bonds to 20
P4	NumBots app	Counting
P5	Place value grid and Dienes (concrete and virtual versions)	3-digit numbers

Finally, Table 5.8 lists the resources that participants reported as not being useful and the different adjustments that they made to effectively use the resource in their classroom. Overall, educators found not useful the resources that: 1) did not promote independence of their students and required significant adult input and guidance to be used meaningfully, and 2) did not promote meaningful interactions with peers. Educators referred to issues related to the language used and mentioned that some resources used a “tricky language” (P3) or that their student “could not read or understand the [meaning of the] question” (P5). Some participants reported issues related to fine motor skills and observed that their students struggled when using resources which are “too small or tiny” (P4). Other participants highlighted issues related to the accessibility of the resources. For example, some worksheets were described as ineffective because the “writing was too small” (P3). Motivation was also mentioned as an important factor to define whether a resource was useful. For example, a participant reported that worksheets were not useful because “they don't really make [the student] want to do anything interesting” (P5). Adjustments made to use the resources effectively included adult assistance or employing an alternative resource.

Table 5.8. Learning resources used by the participants as not being useful and adjustments made.

Resource	Reasons	Adjustments made to use the resource
Worksheets	“They don't really make her want to do anything interesting” (P5). “She needs a lot of guidance” (P5). “She can't read the questions […] or she doesn't understand what the question means” (P5). “Sometimes it's too much information it's overwhelming” (P3). “It is just too small, the writings” (P3).	“We would take the questions from the worksheets but do them in a practical way” (P5). “An adult reads the question to her” (P5). “We use the “I do – we do – you do” approach” (P5). “We would focus [only] on using some elements of the worksheet, [for example] the tangible element and the pictorial element” (P3). “We would put a copy of the worksheet bigger” (P3).
Distracting and overly playful resources	“Fluffy little pop poms […] we could spend hours laughing about them, rather than doing anything” (P1).	“We've changed them. […] You just need to know that what you're working with is going to work and not just become an hour of fun” (P1).
Mathletics maths programme as post teaching retrieval practice	“It does not encourage independence even when differentiated” (P3). “Sometimes it has quite tricky language” (P3). “It does not support meaningful interactions with peers” (P3). “Parents don't find it that user friendly” (P3).	“She needs a lot of modelling” (P3).
Cuisenaire rods	“It involved a lot of adult input to make them meaningful” (P2). “Unless you see them next to each other or they've memorised the colour it is really hard to know which number is” (P3).	“We did not use it anymore” (P2).
Dienes	“They are just too small, tiny” (P4). “I think that sometimes it can be a distraction, because she is very interested in moving them, or building things more than she is actually with the maths” (P5).	Not reported
Plastic coins	“Plastic money didn’t work as he couldn’t relate to it in real life situations […] It doesn't mean anything” (P4).	“We use real money where we can” (P4).

5.3.2.4 Inclusion in the classroom

When describing how maths lessons were delivered, all participants acknowledged the importance of including the student with DS in their classroom not only for their academic growth but also for supporting social development, fostering relationships with their peers, and enhancing their overall confidence. For example, one participant reported: “you want her to be in the class and experiencing all those other kind of foundation subjects that she really… she really enjoys as well and […] she's really motivated to work with her peers… and, she's really well liked by her peers, so I really want to support that kind of self-confidence for her to stay in the classroom as much as possible” (P3). This was in line with families’ views, as highlighted by a participant who reported: “The parents wish that the child is in the classroom as much as possible because they believe he needs the social aspect” (P4).

Participants highlighted the challenges they faced in finding the right balance between providing learning opportunities within the classroom and ensuring meaningful participation to the student with DS. One participant stated: “[Working in the classroom] is really good for self-esteem, but it's a really tricky balance, I think, to make sure that he's not being overwhelmed with all of the information, some of which he can't quit’ understand yet, but also wanting him to, you know, be part of the class, because he is” (P2). Additionally, participants reflected on the challenges related to later primary years (i.e., KS2) where mathematical topics become more complex. Educators reported their concerns about the potential widening gap between the student’s achievement and the level of achievement of their peers, anticipating this might affect the amount of time spent by the student in the maths classroom, with one participant reporting: “it is difficult [to keep the student in the maths class] if you’ve got a Year 4 child who only is able to do Year 1 work” (P4).

Finally, when reflecting on the challenges faced by students with DS while working in the maths classrooms alongside their peers, participants emphasised the importance of preventing these students from comparing themselves to their peers, with one participant reporting: “he's quite self-aware at the moment of where he is with maths, where the others are. So it's about making him not compare himself too much to others, and still feel like he's having success” (P2). Additionally, educators noted that, at times, their student might refrain from participating in discussions taking place in the classroom because of the challenge with understanding the specific topic: “my conviction is [that] she doesn't like to join the discussion, sometimes I see it because she doesn't know quite what's going on either, so that makes it quite hard for us to keep her in the room” (P5).

5.3.2.5 Limited time

Participants explained how the limited time they had for planning and teaching adversely affected the quality of the support they could provide to their students with DS on different levels, e.g., in choosing the appropriate tool and assessing the mathematical abilities of their students. Moreover, participants reported that the effort to differentiate lessons and learning resources was laborious: “I’m having to create a lot of resources for him and thinking a bit outside of the box, which is good in some ways, but it's also quite time consuming […] it takes a lot of time and lots of thinking” (P2). Finally, when talking about the use of a learning platform that their school had adopted, one participant reported that they did not “have enough time to use this resource meaningfully” and that “because I am not able to look at it as much, then I am not on it as much to try and encourage parents to use it” (P3).

5.3.2.6 Collaboration and communication with teaching staff and family

All participants reported how beneficial and helpful it was to collaborate with the teaching team and with families to support their student’s learning.

When describing the collaboration between the members of the teaching team, participants reported that they learnt a lot by collaborating and observing other educators both in school and other settings (e.g., other schools, training settings). Participants reported positive examples of collaborations between class teacher and TA. For example, one participant said that their TA was “very good at judging when my input is getting to a point where [the student] is not accessing it as much, and then she might take him to the table while I’m still teaching and do some target work” (P2). Participants recognised the value of working with confident and skilled educators, with one participant reporting that their TA was “confident enough to know how they can suddenly change something or what other resources might work” (P5). Furthermore, participants acknowledged the importance of communication between members of the teaching team. For example, one participant reported: “the 1 to 1 TA [I work with] is really motivated and proactive and always […] feedback things, you know, write on the sheet what happens. Did the child understand the concept? So, then I can plan some post teaching activities” (P3).

The relationship with the families was also described as a key element to support the student’s learning. Some participants highlighted the involvement of the families in reinforcing the student’s learning at home, stating, for example: “there is a lot of consolidation [at home] and we work together” (P3). Additionally, one participant stressed the importance of sharing learning targets and resources with the family to consolidate the work done at school: “[mum] has the same book and resources at home and so they are just mirroring what happens at school” (P4). Effective communication was identified as a key component for positive collaboration with families. All participants used a communication book to engage with families and agreed that this tool was fundamental to keep a connection with them.

5.4 Discussion

Findings from this qualitative study are based on the experiences of five primary school teachers and TAs teaching mathematics to a student with DS in an English mainstream classroom. This discussion presents the findings in relation to the study RQs and covers the study limitations, implications of the findings and future directions.

5.4.1 What are the challenges related to teaching mathematics to a student with Down syndrome in the inclusive classroom?

This study identified a few challenges that educators face in the primary school inclusive classroom including a student with DS. These challenges included assessing mathematical achievement of students with DS and providing meaningful and positive learning opportunities in the inclusive classroom. Both challenges were related to the third challenge of having limited time available for planning and teaching.

One of the main challenges in assessing mathematical achievement of students with DS was the students’ inconsistent display of mathematical skills. This had already been reported in previous studies investigating primary school teachers’ experiences (Clarke & Faragher, 2015; Rietveld, 2005). Participants from the current study highlighted how this challenge impacted their students' mathematical development. Indeed, they often found themselves revisiting the same learning targets due to the observed inconsistent progress in learning. This placed educators in the dilemma of whether to prioritise mastering prior objectives or progressing to new ones, as required by the national curriculum. Additionally, educators reported having access to several assessment tools and mathematical programmes, but that this felt more like a challenge than a benefit due to their limited time to evaluate and make an informed choice on which assessment to use with their student.

In relation to the challenges related to the implementation of inclusive strategies in the maths classroom, while all participants recognised the multiple and different benefits of including students with DS in their classroom, they also acknowledged the difficulties related to the adjustments needed to ensure a meaningful inclusion. This was in line with the findings reported by Clarke and Faragher (2015). The adjustments to the maths lessons were described by participants as time consuming and laborious, especially in the case of different levels of abilities between the student with DS and their peers. Participants reported that students with DS stayed in the classroom as much as possible and were working on the same mathematical targets as their peers, but that sometimes they were taken out of the classroom, especially in the instances where the classroom environment was too busy and noisy and could negatively impact their learning. Additionally, a participant reported that because their student with DS was not accessing the same maths curriculum year as their peers, they were working outside the classroom with a small group of students who were working at the same level. A recent qualitative study conducted by Losberg and Zwozdiak-Myers (2024) with primary mainstream school teachers based in England reported an alternative solution to support the inclusion of students with special educational needs. This study reported that the students with special educational needs and disabilities who were not accessing the same curriculum year as their peers (n = 3) were receiving individual support for most of the school day and had their separate “workstation” in the classroom. However, the authors also reported that while participants exemplified in their interviews a series of inclusive practices, they also acknowledged that it was difficult to implement them on a regular basis and in all aspects of classroom life, given the complex nature and the diverse range of needs they encountered in their classrooms.

In line with the findings from Monari Martinez and Benedetti (2011), participants acknowledged the importance of the collaboration between teacher and TA, and teachers valued autonomy, initiative and proactivity in their TAs. Again, participants reflected on the limited time they had to oversee their TA’s work. When asked to provide examples of potential applications of the lesson plan template used in Session 3, participants agreed that this was a useful tool to share the planning between teacher and TA. Moreover, some of the teachers proposed that this could be used to give TAs the ownership of specific areas of the planned activities. This highlights the substantial involvement of TAs in the inclusive classroom in England and the lack of a clear definition of their role, as reported by Wren (2017).

Unlike previous studies (Clarke & Faragher, 2015; Faragher & Clarke, 2020; Rietveld, 2005), the current study did not identify any theme related to managing challenging behaviour of the student in the inclusive classroom. This divergence could be linked to multiple factors such as the age of the students or the abilities of the educators to manage the emotional regulation of their students.

5.4.2 What are the teaching strategies and resources employed in the maths inclusive classroom?

One of the outputs of this study was the creation of an infographic showing 20 strategy cards which summarise the final list of strategies reported as effective by participants when supporting mathematical skills (Figure 5.5). The infographic was shared with participants and was made publicly available on the Open Science Framework website (https://osf.io/z7m8r/) at the end of the project. Some of the strategies included in the list were reported in previous studies, such as use of routines, modelling, repetition and use of visuals (Bird, 2016; Kabashi & Kaczmarek, 2019).

The study showed that the learning objectives targeted in the primary school class seem to be more detailed for the area of “Number and place value” – which included foundation skills such as digit recognition, counting and subitizing – while they are more general and less specific for the areas that focus on calculation skills. The different level of description used by the educators might be interpreted in two ways. On one hand, it might simply indicate that calculation skills are not a learning objective that is usually targeted in primary school. This would explain why educators only use general learning objectives for this area, rather than following a clear path of development, as shown, for example, for counting. On the other hand, it may indicate that while educators are provided with enough tools to support early mathematical skills, they lack in training and resources when it comes to supporting more advanced skills of students with DS. As this study only included five participants, these findings need to be interpreted with caution and further research is needed to validate this conclusion.

Figure 5.5. Strategy cards.

Strategy cards

Disclaimer: please note that the teaching strategies presented on these cards are based on the experiences reported by primary-school educators working in a mainstream setting. While these strategies have been used by teachers and teaching assistants, the efficacy of some of these strategies has not been formally researched or validated yet.

01 Introduce a routine (Memory)

"He also knows as part of his routine that he does maths and literacy at a certain time of the day. I write his schedule on a whiteboard, so **he knows what he needs to do now and what is next**"

02 Allow for repetition (Memory)

"You have to go over and over and over the same thing until it sticks"

"You can have two or three great days and then the day after that it is almost as if we had not covered it, you just have to be **recapping** it with him **constantly.**"

03 Use videos (Motivation, Memory)

"We watch the videos and he gets excited by us **being able to do the same things that they show**. This way he gets a bit more interested"

04 Keep activities short and snappy (Attention)

"Usually, activities are not longer than **10 or 15 minutes**, but we are quite flexible with that"

05 Use modelling (Attention, Memory)

"I think it is really important for her to **see the end product**, what is expected of her"

"You have to do modelling **every time**. Even if it is something that we do regularly, you still need to go back and model it at the beginning of the session in order to get the most out of it"

06 Break down activities in small steps and be flexible (Attention, Motivation)

"If he is having a tricky day, we just focus on a **few bits of the activity** and then, if he is up for a challenge and really motivated the length and difficulty of the activity can be increased"

"If manipulatives are not working, then we might need to change the type of manipulatives that we are using or how we are doing the task"

07 Do not take anything for granted (Memory)

"We need to **go back to the beginning** for him to remember where we are and what we are doing. He will realize we have done it before, but perhaps not each stage"

"Sometimes trying something new just does not work. So, we need to come back to what we know first and **gradually bring in other things**"

08 Use signs and gestures (Memory)

"If I can, I use **Makaton** when it is relevant"

09 Collaborate with family (Language, Memory)

"I have sticky notes on my side and I write down the words that he does not understand. His mum does the same. So, between us **we find the words that he struggles with**"

"**Mum reinforces** the things we do at school and I think that it is really helpful because it allows for repetition and supports generalisation with different people"

10 Make learning playful (Motivation)

"We go **outside**, and we go jewel hunting like swashbucklers with ropes and sharks"

"He brings his **board games** from home. We got Magic Maths game at school, and we use Numicon. We build lots of towers with **Numicon** because he is counting in 10s"

Strategy cards

Disclaimer: please note that the teaching strategies presented on these cards are based on the experiences reported by primary-school educators working in a mainstream setting. While these strategies have been used by teachers and teaching assistants, the efficacy of some of these strategies has not been formally researched or validated yet.

11 Make learning practical (Motivation)

"We **take the questions from the worksheets** to do them in a practical way"

"I created some worksheets where he had to draw circles around a pair of spoons and then I got a packet of plastic spoons and I got him to put them in groups of 2 on the floor"

12 Use accessible resources (Language, Motor Skills)

"**Worksheets** must be big, bold, colourful, easily read, and he must be able to understand them"

"We use concrete resources that are **not too finicky**"

"In his exercise books I have to draw the lines for him, so that he can see them"

13 Use visual and concrete resources (Attention, Motor Skills)

"The resources that we are using are very visual, and I think that it makes a big difference for him"

"When she uses manipulatives, she really likes being able to **move things**"

14 Plan meaningful learning interactions (Motivation)

"She has a mixed-ability partner and so there is **lots of talking** there"

"A possible scenario could be that a teaching assistant has a problem and we ask the child to help them **work the problem out**"

15 Use tablets, maths apps and the whiteboard (Attention, Motor Skills)

"She quite likes to see things that are just happening on the screen which **capture her interest**"

"It is nice being able to model the activity on the **whiteboard** and then the children can use it. So, they can work on **gross motor** and **fine motor skills**"

16 Personalize learning resources (Motivation)

"I printed some **jewels** that we can use for our activities"

"He does not like stickers so now we have got a **jar with stars**, and when he is doing maths he gets a star"

17 Introduce pre-teaching (Motivation, Memory)

"The pre-teaching is modelled by the teaching assistant during the morning work and uses the **same tangibles and strategies** that I use in the lesson. This gives her the confidence to put her hand up during the lesson"

18 Be aware of the language that you use (Language)

"I think it is the **question** that sometimes he does not understand"

"When it comes to instructions, they need to be **clear** and **broken down in small steps**"

19 Design the learning environment (Attention)

"It gets a bit **overwhelming** when there is lots of things on the table"

"The worksheet can be broken down, instead of having a sheet with lots of stuff on it, that might be visually overwhelming"

"If it is a tough session, we might have **brain breaks**"

"He really needs the **quiet surrounding**"

20 Pitch to the student's "right level" (Language, Motivation)

"We were using balancing scales to weigh things and he was measuring only light objects because he is still working on counting between 1 and 10"

"**Instructions** must be on his level"

5.4.3 Are the teaching strategies used by educators aligned with the cognitive profile of students with Down syndrome?

The current study linked for the first time the inclusive practices reported by the educators to the literature outlining the cognitive profile and the needs of learners with DS. The alignment observed between the strategies and the resources used by the educators in the maths inclusive classroom together with the cognitive profile associated with students with DS, as described by the literature, validates the practices employed in the English inclusive classroom. For example, educators reported using strategies such as repetition, modelling, and the use of routine to support verbal memory and to ensure consistent language. As discussed in Chapter 1, verbal memory and language were reported to be areas of weakness for individuals with DS (Onnivello et al., 2022). Additionally, the teaching strategies reported by the participants encompassed attention and fine motor skills, which have also been identified as typical areas of weaknesses for students with DS (Silverman, 2007). Furthermore, a consistent number of the teaching strategies addressed motivation, which has been reported as a key element in the effective strategies used in the inclusive classroom in previous studies (Bird, 2016; Kabashi & Kaczmarek, 2019).

Similarly, when describing the learning resources employed in the classroom, educators regarded as ineffective those resources whose dimension (too small), use of language (too difficult) or level of engagement (too low) were not appropriate. Once again, this shows that the participants were familiar with the cognitive processes reported in the literature as areas of difficulty for individuals with DS. Notably, despite language and communication being one of the main areas of vulnerability for individuals with DS (Karmiloff-Smith et al., 2016), only four out of the 20 teaching strategies identified by educators addressed this area, while most of the strategies targeted memory (n = 8) and motivation (n = 8). A possible explanation for this discrepancy is that, given that most students with DS in the UK receive support from external Speech and Language Therapists (Hargreaves et al., 2021; Ranzato et al., 2021), educators might think that this area falls outside their sphere of responsibility.

5.4.4 Limitations

Although this study offers valuable insights into educators' experiences, it is important to acknowledge that it has certain limitations. First, the study employed qualitative methods that are subjective to the researchers' preconceived notions, experiences, and personal beliefs, which can influence the data analysis and the interpretation of the findings. Second, the sample of participants who took part in this research was small and composed of educators with a particular interest in the research topic and in their own professional development. As such, when reading these findings, it is important to consider that participants may not represent the wider population of educators that are employed in England. Third, the limited time available to participants conflicted with the multiple requests that participating in the research project involved. Although most participants attended all the online sessions (one participant withdrew after the first session), very few completed the preparatory activities. Consequently, the data from such tasks could not be used. In the future, it might be beneficial to reduce the demands on participants between sessions and integrate these activities into the focus group sessions, even if this means making the sessions slightly longer.

5.4.5 Implications and future directions

Despite these limitations, this study presents implications relevant for both research and practice. On one hand, the use of the focus group interviews proved to be a valuable tool to explore educators' experiences and to provide participants with tools to reflect on and to evaluate their practice. It is hoped that this study inspires more researchers to adopt this approach in the future. On the other hand, these findings provide additional evidence on the inclusive teaching strategies and enrich our understanding of what happens in the inclusive maths classroom. This insight can inform the development of interventions and educational programmes to support the inclusion of students with DS in the maths classroom. While this is an important piece of evidence, it is important to consider that this study did not evaluate their efficacy. Future studies should evaluate the use of these teaching strategies and measure their impact in supporting the mathematical learning of students with DS in the inclusive classroom.

Finally, as this study only focused on primary school children with DS within a specific country, further avenues for future studies would be to explore and compare experiences of educators of the inclusive maths classroom in different neurodivergent populations and educational contexts.

5.5 Conclusion

The current study used focus group interviews to explore the experiences of educators supporting a primary school student with DS in the maths inclusive classroom in terms of challenges faced and teaching strategies and resources used.

The reflexive thematic analysis identified three main challenges for the educators: 1) assessing mathematical achievement of the student with DS, 2) providing meaningful and positive learning opportunities in the classroom and, 3) limited time for planning and teaching. Moreover, findings from this study showed that the teaching strategies and the learning resources employed in the classroom target those cognitive processes that are crucial for mathematical learning and that the literature describes as areas of weaknesses for individuals with DS. These included memory, attention, language, and motor skills. Furthermore, motivation was reported to be an important factor targeted by the educators to support the learning of their students with DS.

These findings contribute to the literature on inclusive pedagogy and to the knowledge base of educators working with children with DS in English primary mainstream settings. Moreover, they provide valuable insight for the development and implementation of inclusive interventions and accommodations in the classroom, especially for teaching mathematics to students with DS.

Chapter 6: General discussion

As stated in Section 1.4, this thesis had two main aims:

1. Investigate syndrome specificity of mathematical profiles in different neurodivergent populations.
2. Investigate the mathematical learning experiences of primary school children with neurodevelopmental conditions (NDCs), with a focus on Down syndrome (DS) and Williams syndrome (WS).

The forthcoming general discussion will not involve a re-examination of the single studies presented in the previous chapters. Instead, it will assess the degree to which the overarching aims of this thesis have been addressed, discuss the wider implications and limitations of the findings, and consider future research directions.

6.1 Aim 1. Syndrome specificity of mathematical profiles

The first aim of this thesis was to investigate mathematical profiles of different neurodivergent populations by using the multilevel framework of mathematical cognition developed by Gilmore (2023) and to compare them. Because this thesis adopted the multilevel framework of mathematical cognition (Gilmore, 2023), the description of mathematical profile involved not only basic processes and specific components of mathematics (Chapter 2), but also the home learning environment, which is part of the wider component of the learning experiences (Chapter 3), and domain-general factors associated with mathematical development, such as eye movements (Chapter 2) and intellectual disability (ID; Chapter 4). Moreover, because this thesis adopted a neuroconstructivist approach, it explored how mathematical profiles change across development.

To address the question of syndrome specificity, a series of studies comparing mathematical abilities across different neurodivergent populations was presented. The comparisons were made through the use of different methods, including qualitative and quantitative experimental studies (Chapter 2 and Chapter 3), and a systematic review of literature (Chapter 4).

The study presented in Chapter 2 compared the performance on an enumeration task and the eye movements of children and adults with DS and with

WS. The results from this study showed that the performance of individuals with DS and WS in the enumeration task measured through accuracy scores and reaction times (RTs) was similar to the one observed for the typically developing (TD) control group matched for MA (RCPM). Moreover, the analyses of RT slopes showed that participants with DS and WS were using the same enumeration processes as the TD control group in the same experimental conditions, i.e., perceptual subitizing when enumerating 1-3 dots, conceptual subitizing when enumerating 4-6 dots in the dice condition and counting when enumerating 4-6 dots in the random condition. Additionally, this study compared the eye movements of participants with DS and with WS and showed significant group differences. In fact, participants with DS reported higher fixation count and shorter median fixation duration than both the WS group and the TD group, and participants with WS reported shorter median fixation duration than the TD control group. These findings highlight that the eye movements of individuals with DS and with WS differed when performing the enumeration task. However, these differences were not reflected in differences in the mathematical profiles of these populations, at least with reference to their performance level in enumeration skills.

The study presented in Chapter 3 described and compared structural and functional indicators of the home learning environment of primary school children with DS and with WS, with a specific focus on the home mathematical environment (HME), through a web-based parental questionnaire. The results from this study do not report significant differences between DS and WS groups, which only differed in the frequency of activities supporting digit recognition and counting activities, with parents of children with DS reporting a higher occurrence for these activities. Instead, the types and the frequency of home-based maths activities were shown to change on the basis of the general level of functioning of the child, with children with lower levels of general functioning scores frequently involved in activities supporting number skills, and children with higher levels of general functioning scores more frequently involved in calculation-based and literacy-based activities. Moreover, higher levels of adaptive behaviours were associated with higher parental expectations for maths- and literacy-based competences. Overall, the findings of this study suggest that the home learning environment is not syndrome specific and that the general level of functioning, rather than the clinical categories, impact the type and frequency of educational activities that occur at home.

Chapter 4 presented a systematic review of mathematical abilities in autism with the aim to define and compare the mathematical profiles of autistic individuals with ID and without ID. Due to the small number of studies investigating mathematical abilities in autistic individuals diagnosed with ID, it was not possible to identify a complete mathematical profile for this population, which reported poor arithmetic word problem skills, poor overall mathematics achievement, and mixed calculation skills. Autistic individuals without ID reported arithmetic word problem and calculation skills within the typical range, overall mathematics achievement shifted towards the lower end of the typical range, and enumeration skills, magnitude comparison skills and number line skills in line with CA and in line with MA. Because the mathematical profile identified for autistic individuals with ID was incomplete, it was not possible to make cross-syndrome comparisons and to fully address the question of syndrome specificity. For example, due to the lack of studies investigating enumeration skills in autistic individuals with ID, it was not possible to determine whether the mathematical profile of these individuals is more similar to the mathematical profile of DS and WS, as described in Chapter 2, or to the one of autistic individuals with higher IQ scores. The comparison of enumeration skills observed in DS and WS with the ones of autistic individuals without ID did not highlight any similarity. In fact, while enumeration skills are reported to be in line with CA and MA in autistic individuals without ID, the findings from Chapter 2 report enumeration skills for DS and WS populations to be in line with MA only.

The issue of syndrome specificity is clearly very complex to address, and not one that this thesis has been able to fully resolve. However, this thesis has significantly expanded the research in this under-researched area by conducting additional cross-syndrome comparison analyses on basic processes and specific components of mathematics, by synthesising existing literature, and by identifying the similarities between different NDCs as opposed to only highlighting differences between them.

6.2 Aim 2. Exploration of learning experiences in neurodivergent populations

Most studies investigating mathematical development in NDCs do not take into account the role of the environment, an important factor for development according to the neuroconstructivist approach (Mareschal, 2007). Hence, the second objective of this thesis was to investigate the mathematical learning experiences in neurodivergent populations in the form of the home mathematical environment (HME) and of the teaching strategies and resources used in the inclusive maths classroom, with a focus on primary school children with DS and with WS.

The study presented in Chapter 3 offers insights into the learning activities and resources which parents use at home to support the mathematical abilities of their primary school child with DS and with WS. Findings from this study suggest that for both NDCs literacy-based activities, which occurred once a week, occurred more frequently than mathematics-based activities, which occurred less than once a week. Furthermore, when comparing the frequency of occurrence of the three mathematics-based activities (i.e., counting and digit recognition, calculation, and broader mathematical skills), no differences were found. This shows that parents of primary school children with DS and with WS offer a rich HME characterised by activities that are not limited to counting and digit recognition but include activities which are focused on other aspects of mathematics, such as calculation, functional mathematics, and geometry. Moreover, findings from this study show that, for both groups, parents' expectations for their child's mathematical abilities were generally high, with the exception of expectations for their child's calculation skills.

The study presented in Chapter 5 described the teaching strategies and resources used by educators who support students with DS in the English mainstream primary school. The study did not include educators who support students with WS because this would have been too difficult to set up given the tight timeline to complete the project set by the funding body. As reported by Nogues and Dorneles (2021), in recent years great effort has been put into transferring results from studies in cognitive psychology into teaching practice, but this was limited to the TD population. Pedagogies, learning strategies, and the learning resources used at school are quite an unexplored area when it comes to NDCs. With that in mind, this study offered insights into what goes on in the classroom to enable researchers to support teaching staff in leveraging existing resources and improving current practices. The findings show that teachers and teaching assistants used teaching strategies aimed at supporting general cognitive processes which are described by the literature as areas of vulnerability for students with DS, as discussed in Chapter 1. These included memory, language, motor skills, attention, and motivation.

Similarly, when describing why educators found some teaching resources not useful, they supported their answers by pointing to the same cognitive areas and mentioned barriers related to the language used, the resource not being accessible because too small to be handled by the student, or the resource not being interesting enough to engage the student. These findings show that primary school educators in the mainstream class employ teaching strategies which are aligned with the cognitive profile of students with DS and that they pay attention and adapt learning material to their students' profile. The study also reports that educators identified as challenges 1) the assessment of the mathematical abilities of their students with DS, and 2) the implementation of meaningful inclusive strategies in their classroom. Finally, findings show that the learning objectives provided by educators tend to be more detailed for the area of "Number and place value" – which includes foundation skills such as digit recognition, counting and subitizing – while they become more vague and less specific for the area of calculation skills.

In summary, these studies report that mathematical learning experiences of individuals with NDCs are rich and individualised. In fact, these studies confirm that the home learning environment of primary school children with DS and with WS is rich and changes on the basis of the general level of functioning of the child, and that educators of primary school children with DS offer inclusive maths learning experiences, when possible, and that they take into account the cognitive profile of their student by adapting their teaching strategies and learning materials.

6.3 Impact of the findings

The findings from the studies presented in this thesis challenge several misconceptions concerning mathematical abilities in neurodivergent populations and provide insights into how mathematical development is supported in their homes and schools.

In relation to the DS group and the WS group, these findings reveal some similarities in the mathematical profiles of the two, particularly in relation to their performance level in enumeration. The overlap observed for the enumeration component may have implications for the development of mathematical curriculum, practices used by educators to support these students, and in general for the design of educational interventions supporting counting skills in these populations. In fact, given the fact that children and adults with DS and with WS can perform perceptual subitizing, conceptual subitizing, and counting (as shown in Chapter 2), learning objectives should be defined to specifically target these abilities. Moreover, knowing that the performance level in enumeration is similar in these two populations, even if predictors may differ, can be useful for educators who are supporting students with WS – a rare genetic condition – for the first time, but who have previously supported students with DS – a more common genetic condition. Finally, from a research point of view, this finding can have implications for the development of interventions to support counting skills in these populations. In fact, research in TD population has shown that perceptual subitizing and conceptual subitizing support the development of counting. This prompts the question of whether conceptual and perceptual subitizing support counting skills in these neurodivergent populations as well.

Regarding the autistic population, the results of the systematic review challenge the prejudice of autism being associated with exceptional mathematical abilities and highlight the existence of different mathematical profiles within autism. This may have implications for educators, policy makers, and families in relation to setting realistic expectations and providing adequate support to autistic students in relation to their mathematical development. Moreover, with respect to the impact on research, the findings from this study highlight gaps in the literature and the need to investigate mathematical development in autistic individuals with ID.

The findings from the studies investigating mathematical learning experiences in NDCs provide a snapshot of what happens in the homes and in the classrooms of children with DS and WS and provide essential information in relation to the frequency and the type of activities and resources that parents and educators employ, as well as the challenges that they face. These may inform researchers in the design of interventions to improve both the HME and the pedagogical methods used by educators to support students with NDCs in the inclusive maths classroom. For example, knowing about the frequency of the activities which occur at home and in school is relevant for researchers to design intervention programmes that are not too demanding on the adult supporting the child. Additionally, knowing about the type of activities and resources employed at home and in school gives researchers the opportunity to specifically use such resources in their interventions to reduce attrition and implementation costs. Finally, these studies provide insights into the challenges and concerns reported by parents and educators, such as the use of technology, the limited time available, the assessment of mathematical abilities of neurodivergent students, and the implementation of inclusive practices in the primary mainstream classroom. These insights may provide researchers with a rationale to develop their investigations and tools to address these specific areas of concern.

From a theoretical point of view, all the studies included in this thesis use the neuroconstructivist perspective and extended the use of the multi-level framework by Gilmore (2023) to populations with NDCs. On one hand, the use of the neuroconstructivist perspective highlights the developing nature of the mathematical profiles of these populations. On the other hand, the use of the multi-level framework of mathematical cognition (Gilmore, 2023) confirms that this framework is flexible enough to accurately describe mathematical profiles in neurodivergent populations and highlights the multi-component nature of mathematics. While this is a well-known concept to experts in the field of mathematical cognition, promoting the use of this framework to describe mathematical development in neurodivergent populations could help policy-makers and educators who are not familiar with the subject matter to understand the multi-component nature of mathematical profiles.

It is hoped that the clarification of these misconceptions, along with the insights about how mathematical abilities of neurodivergent children are supported in real-life settings, will lead to a more accurate understanding of mathematical development in neurodivergent populations. Finally, it is hoped that the implementation of the findings will promote the inclusion of individuals with NDCs in the classroom and in society. This, in turn, is anticipated to positively impact the well-being of their families and of the broader community supporting them, while making the best of the limited resources typically assigned to special needs education.

6.4 Limitations and future directions

The main limitation of this work comes from the partial description of the mathematical profile of autistic individuals with ID, which prevents its comparison against the mathematical profile of autistic individuals without ID and against the mathematical profiles of other NDCs associated with ID, such as WS and DS. Future studies should investigate mathematical abilities in autistic individuals with ID and fill this gap. This will allow researchers to better understand to which extent the use of diagnostic categories is useful to describe the mathematical profile of neurodivergent populations and if different categorizations should be used instead.

Another limitation is related to the number of components investigated in the studies included in the current thesis, which is relatively small in comparison to the whole set of components included in the multi-level framework of mathematical cognition by Gilmore (2023). In fact, the studies included in Chapter 2 and in Chapter 3 only explored enumeration skills and the HME. This limited the extent to which the mathematical profiles of the populations investigated could be compared. Future implementations of this work should investigate a wider range of components of the multi-level framework of mathematical cognition to allow for a more comprehensive comparison of the mathematical profiles of these populations. Moreover, future studies should investigate the interactions between the different components of the framework. Future developments of this work should also involve different NDCs, beyond the ones included in the current thesis. The creation of a public dataset collecting studies investigating different processes and specific components of mathematics in different neurodivergent populations is also hoped for as a future research avenue. Such dataset could be kept up to date by researchers inputting findings from their studies and using a predefined coding system. This would help researchers identify gaps in the literature, describe and compare mathematical profiles, and clarify whether and how theory can inform the design of mathematical interventions for neurodivergent populations.

Finally, the main limitation of the studies investigating mathematical experiences of individuals with NDCs is that they do not provide a measure of the mathematical abilities of the individuals whose environment is being investigated. Hence, these studies do not investigate the role of the environment on the mathematical development of neurodivergent individuals. Future studies should address this limitation. For example, with reference to the study presented in Chapter 3, the mathematical abilities of the child of the respondent could be assessed by the researcher, either face to face or online, before the completion of the survey. Depending on the specific research questions and on the characteristic of the sample (e.g., mean CA), the assessment of mathematical abilities could focus on specific components of mathematics, such as counting abilities or calculation skills, or it may focus on the individual’s overall mathematical abilities. Regression analyses investigating the relationship between the score of a child’s mathematical abilities (dependent variable) and the measures of the different indicators of the HME (independent variables) should explain and quantify the relationship between these variables. Intervention studies would instead be needed to measure the impact of formal aspects of mathematical education, such as teaching strategies and learning resources employed in school, on mathematical development. In this case, the mathematical abilities of the student supported by the educator who is taking part in the study should be measured both before the start of the study and after its completion in order to evaluate whether the adoption of a specific tool or strategy has improved the mathematical abilities of the student. In this instance, the mathematical abilities could be assessed either by the researcher or by the educator through standardised assessments and curriculum-based measures.

6.5 Conclusion

This thesis investigated the mathematical profiles of three neurodivergent populations and showed similarities in performance on the enumeration task and on the HME of DS and WS populations. A different mathematical profile emerged for autistic individuals without ID. Additionally, this thesis investigated the mathematical learning experiences of primary school children with DS and with WS in the form of their home learning environment and of the teaching strategies and resources used in the English maths inclusive classroom, and it showed that mathematical learning experiences of individuals with DS and with WS are rich and individualised.

References

Aagten-Murphy, D., Attucci, C., Daniel, N., Klaric, E., Burr, D., & Pellicano, E. (2015). Numerical estimation in children with autism. *Autism Research*, *8*(6), 668-681. https://doi.org/10.1002/aur.1482

Al Otaiba, S., Lewis, S., Whalon, K., Dyrlund, A., & McKenzie, A. R. (2009). Home literacy environments of young children with Down syndrome: Findings from a web-based survey. *Remedial and Special Education*, *30*(2), 96-107. https://doi.org/10.1177/0741932508315050

Al-Beltagi, M. (2021). Autism medical comorbidities. *World journal of clinical pediatrics*, *10*(3), 15-28. https://doi.org/10.5409/wjcp.v10.i3.15

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders*. American Psychiatric Association Publishing.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, DSM-5* (Fifth ed.). American Psychiatric Publishing. https://doi.org/doi.org/10.1176/appi.books.9780890425596

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders, DSM-5-TR* (Text revision, fifth ed.). American Psychiatric Association Publishing. https://doi.org/dsm.psychiatryonline.org/doi/book/10.1176/appi.books.9780890425787

Anneli, O. (2012). The Tobii I-VT Fixation Filter. http://www.vinis.co.kr/ivt_filter.pdf

Ansari, D. (2003). *Atypical trajectories of number development: The case of Williams syndrome* [Unpublished doctoral thesis, University College London].

Ansari, D., Donlan, C., & Karmiloff-Smith, A. (2007). Typical and atypical development of visual estimation abilities. *Cortex*, *43*(6), 758-768. https://doi.org/10.1016/S0010-9452(08)70504-5

Antonarakis, S. E., Skotko, B. G., Rafii, M. S., Strydom, A., Pape, S. E., Bianchi, D. W., Sherman, S. L., & Reeves, R. H. (2020). Down syndrome. *Nature reviews. Disease primers*, *6*(1), 9-9. https://doi.org/10.1038/s41572-019-0143-7

Ashworth, M., Palikara, O., Burchell, E., Purser, H., Nikolla, D., & Van Herwegen, J. (2021). Online and face-to-face performance on two cognitive tasks in children with Williams syndrome. *Frontiers in Psychology*, *11*. https://doi.org/doi.org/10.3389/fpsyg.2020.594465

Asperger, H. (1991). ‘Autistic psychopathy’ in childhood U. Frith, Ed. & Trans. In *Autism and Asperger Syndrome* (pp. 37–92). CUP.

Astle, D. E., Holmes, J., Kievit, R., & Gathercole, S. E. (2022). Annual Research Review: The transdiagnostic revolution in neurodevelopmental disorders. *Journal of Child Psychology and Psychiatry*, *63*(4), 397-417. https://doi.org/10.1111/jcpp.13481

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology*, *96*(4), 699-713. https://doi.org/10.1037/0022-0663.96.4.699

Bae, Y. S., Chiang, H.-M., & Hickson, L. (2015). Mathematical word problem solving ability of children with autism spectrum disorder and their typically developing peers. *Journal of Autism and Developmental Disorders*, *45*(7), 2200-2208. https://doi.org/10.1007/s10803-015-2387-8

Baedke, J. (2013). The epigenetic landscape in the course of time: Conrad Hal Waddington's methodological impact on the life sciences. *Studies in history and philosophy of science. Part C, Studies in history and philosophy of biological and biomedical sciences*, *44*(4), 756-773. https://doi.org/10.1016/j.shpsc.2013.06.001

Bashash, L., Outhred, L., & Bochner, S. (2003). Counting skills and number concepts of students with moderate intellectual disabilities. *International journal of disability, development, and education*, *50*(3), 325-345. https://doi.org/10.1080/1034912032000120480

Bird, G. (2016). *Helping children with Down's syndrome to learn*. Retrieved 4/12/2023 from https://senmagazine.co.uk/content/specific-needs/down-syndrome/1920/helping-children-with-down-s-syndrome-to-learn/

Blevins-Knabe, B., Austin, A. B., Musun, L., Eddy, A., & Jones, R. M. (2000). Family home care providers' and parents' beliefs and practices concerning mathematics with young children. *Early Child Development and Care*, *165*(1), 41-58. https://doi.org/10.1080/0300443001650104

Bradshaw, A. R., Woodhead, Z. V. J., Thompson, P. A., & Bishop, D. V. M. (2020). Profile of language abilities in a sample of adults with developmental disorders. *Dyslexia*, *27*(1), 3-28. https://doi.org/10.1002/dys.1672

Bramer, W. M., de Jonge, G. B., Rethlefsen, M. L., Mast, F., & Kleijnen, J. (2018). A systematic approach to searching: An efficient and complete method to develop literature searches. *Journal of the Medical Library Association*, *106*(4), 531-541. https://doi.org/10.5195/jmla.2018.283

Brankaer, C., Ghesquière, P., & De Smedt, B. (2011). Numerical magnitude processing in children with mild intellectual disabilities. *Research in Developmental Disabilities*, *32*(6), 2853-2859. https://doi.org/10.1016/j.ridd.2011.05.020

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health*, *11*(4), 589-597. https://doi.org/10.1080/2159676X.2019.1628806

Brigstocke, S., Hulme, C., & Nye, J. (2008). Number and arithmetic skills in children with Down syndrome. *Down Syndrome Research and Practice* 74-78. https://doi.org/doi:10.3104/reviews.2070

Bull, M. J. (2020). Down Syndrome. *The New England journal of medicine*, *382*(24), 2344-2352. https://doi.org/10.1056/NEJMra1706537

Bullen, J. C., Lerro, L. S., Zajic, M., McIntyre, N., & Mundy, P. (2020). A developmental study of mathematics in children with autism spectrum disorder, symptoms of attention deficit hyperactivity disorder, or typical development. *Journal of Autism and Developmental Disorders*, *50*(12), 4463-4476. https://doi.org/10.1007/s10803-020-04500-9

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry*, *46*(1), 3-18. https://doi.org/10.1111/j.1469-7610.2004.00374.x

Cahoon, A., Cassidy, T., & Simms, V. (2017). Parents' views and experiences of the informal and formal home numeracy environment. *Learning, culture and social interaction*, *15*, 69-79. https://doi.org/10.1016/j.lcsi.2017.08.002

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition: International Journal of Cognitive Science*, *130*(3), 278-288. https://doi.org/10.1016/j.cognition.2013.11.015

Chen, L., Abrams, D. A., Rosenberg-Lee, M., Iuculano, T., Wakeman, H. N., Prathap, S., Chen, T. W., & Menon, V. (2019). Quantitative analysis of heterogeneity in academic achievement of children with autism. *Clinical Psychological Science*, *7*(2), 362-380. https://doi.org/10.1177/2167702618809353

Cheung, S. K., & McBride, C. (2017). Effectiveness of parent-child number board game playing in promoting chinese kindergarteners' numeracy skills and mathematics interest. *Early Education and Development*, *28*(5), 572-589. https://doi.org/10.1080/10409289.2016.1258932

Chiang, H.-M., & Lin, Y.-H. (2007). Mathematical ability of students with Asperger Syndrome and High-Functioning Autism : A review of literature. *Autism: The International Journal of Research and Practice*, *11*(6), 547-556. https://doi.org/10.1177/1362361307083259

Clarke, B., & Faragher, R. (2015). Inclusive practices in the teaching of mathematics: Supporting the work of effective primary teachers. Mathematics Education Research Group of Australasia, Sunshine Coast.

Clements, D. H. (1999). Subitizing: What Is It? Why Teach It? *Teaching children mathematics*, *5*(7), 400-405. https://doi.org/10.5951/TCM.5.7.0400

Coolen, I. E. J. I., Riggs, K. J., Bugler, M., & Castronovo, J. (2022). The approximate number system and mathematics achievement: It's complicated. A thorough investigation of different ANS measures and executive functions in mathematics achievement in children. *Journal of cognitive psychology (Hove, England)*, *34*(6), 796-818. https://doi.org/10.1080/20445911.2022.2044338

Costa, H. M., Outhwaite, L. A., & Van Herwegen, J. (2023). Early years practitioners' training, beliefs and practices concerning mathematics: Implications for education and practice. *International Journal of Early Years Education*, 1-18. https://doi.org/10.1080/09669760.2023.2254822

D'Souza, H., & D'Souza, D. (2020). The Emerging Phenotype in Infants With Down Syndrome: Adaptations to Atypical Constraints. In J. A. Burack, J. O. Edgin, & L. Abbeduto (Eds.), *The Oxford Handbook of Down Syndrome and Development* (pp. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190645441.013.34

D'Souza, H., & Karmiloff-Smith, A. (2017). Neurodevelopmental disorders. *Wiley interdisciplinary reviews. Cognitive science*, *8*(1-2), e1398-n/a. https://doi.org/10.1002/wcs.1398

Daucourt, M. C., Napoli, A. R., Quinn, J. M., Wood, S. G., & Hart, S. A. (2021). The home math environment and math achievement: A meta-analysis. *Psychological Bulletin*, *147*(6), 565-596. https://doi.org/10.1037/bul0000330

De Keyser, L., Bakker, M., Rathe, S., Wijns, N., Torbeyns, J., Verschaffel, L., & De Smedt, B. (2020). No association between the home math environment and numerical and patterning skills in a large and diverse sample of 5- to 6-year-olds. *Frontiers in Psychology*, *11*, Article 547626. https://doi.org/10.3389/fpsyg.2020.547626

De Smedt, B. (2022). Individual differences in mathematical cognition: a Bert's eye view. *Current opinion in behavioral sciences*, *46*, Article 101175. https://doi.org/10.1016/j.cobeha.2022.101175

Deckers, S. R. J. M., Van Zaalen, Y., Van Balkom, H., & Verhoeven, L. (2019). Predictors of receptive and expressive vocabulary development in children with Down syndrome. *International journal of speech language pathology*, *21*(1), 10-22. https://doi.org/10.1080/17549507.2017.1363290 del Rio, M. F., Susperreguy, M. I., Strasser, K., & Salinas, V. (2017). Distinct influences of mothers and fathers on kindergartners' numeracy performance: The role of math anxiety, home numeracy practices, and numeracy expectations. *Early Education and Development*, *28*(8), 939-955. https://doi.org/10.1080/10409289.2017.1331662

Department for Education. (2014). *The national curriculum in England: Complete framework for key stages 1 to 4.* Retrieved from https://www.gov.uk/government/publications/national-curriculum-in-england-framework-for-key-stages-1-to-4

Department for Education. (2023). *Length of the school week. Non-statutory guidance*. Retrieved from https://assets.publishing.service.gov.uk/media/64a2f21fbb13dc000cb2e5e1/Minimum_School_Week_Non-Statutory_Guidance.pdf

Dietrich, J. F., Huber, S., Klein, E., Willmes, K., Pixner, S., & Moeller, K. (2016). A systematic investigation of accuracy and response time based measures used to index ANS Acuity. *PLoS ONE*, *11*(9), Article e0163076. https://doi.org/10.1371/journal.pone.0163076

Dockrell, J. E., Ricketts, J., Palikara, O., Charman, T., & Lindsay, G. A. (2019). What drives educational support for children with developmental language disorder or autism spectrum disorder: Needs, or diagnostic category? *Frontiers in Education*, *4*, Article 29. https://doi.org/10.3389/feduc.2019.00029

Dowker, A. (2020). Arithmetic in developmental cognitive disabilities. *Research in Developmental Disabilities*, *107*, Article 103778. https://doi.org/10.1016/j.ridd.2020.103778

Down’s Syndrome Association. (2020). *Screening, diagnosis and support*. Retrieved 14/11/2023 from https://www.downs-syndrome.org.uk/wp-content/uploads/2020/09/25.08.2020-Screening-diagnosis-support-final-1.pdf

Eaves, L. C., & Ho, H. H. (1997). School placement and academic achievement in children with autistic spectrum disorders. *Journal of Developmental and Physical Disabilities*, *9*(4), 277-291. https://doi.org/10.1023/A:1024944226971

Ewart, A. K., Morris, C. A., Atkinson, D., Jin, W., Sternes, K., Spallone, P., Stock, A. D., Leppert, M., & Keating, M. T. (1993). Hemizygosity at the elastin locus in a developmental disorder, Williams syndrome. *Nature genetics*, *5*(1), 11-16. https://doi.org/10.1038/ng0993-11

Facon, B., & Nuchadee, M.-L. (2010). An item analysis of Raven's Colored Progressive Matrices among participants with Down syndrome. *Research in Developmental Disabilities*, *31*(1), 243-249. https://doi.org/10.1016/j.ridd.2009.09.011

Faragher, R., Stratford, M., & Clarke, B. (2017). Teaching children with Down syndrome in inclusive primary mathematics classrooms. *Australian Primary Mathematics Classroom*, *22*(4), 13-16. https://go-gale-com.libproxy.ucl.ac.uk/ps/i.do?p=AONE&u=ucl_ttda&id=GALE|A521874913&v=2.1&it=r

Faragher, R. M., & Clarke, B. A. (2020). Inclusive practices in the teaching of mathematics : Some findings from research including children with Down syndrome. *Mathematics education research journal*, *32*(1), 121-146. https://doi.org/10.1007/s13394-019-00294-x

Farran, E. K., Purser, H. R. M., Jarrold, C., Thomas, M., Scerif, G., Stojanovik, V., & Van Herwegen, J. (2024). Cross-sectional and longitudinal assessment of

cognitive development in Williams syndrome. *Developmental Science*, *27*(1), Article e13421. https://doi.org/https://doi.org/10.1111/desc.13421
Fauzy, S. I. F. M., & Hosshan, H. (2023). Early numeracy challenges and interventions for students with Down syndrome: A scoping review. *Jurnal Pendidikan Bitara UPSI*, *16*(1), 1-10. https://doi.org/doi.org/10.37134/bitara.vol16.1.1.2023
Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307-314. https://doi.org/10.1016/j.tics.2004.05.002
Feng, J., Lazar, J., Kumin, L., & Ozok, A. (2010). Computer usage by children with Down syndrome: Challenges and future research. *ACM transactions on accessible computing*, *2*(3), 1-44. https://doi.org/10.1145/1714458.1714460
Fitzgerald, J., Roberts, J., Pierce, P., & Schuele, M. (1995). Evaluation of home literacy environment: An illustration with pre school children with Down syndrome. *Reading & Writing Quarterly*, *11*(4), 311-334. https://doi.org/10.1080/1057356950110401
Fletcher-Watson, S., & Happé, F. (2019). *Autism: A New Introduction to Psychological Theory and Current Debate* (1 ed.). Routledge. https://doi.org/10.4324/9781315101699
Fodstad, J. C., Elias, R., & Sarawgi, S. (2020). Intellectual Disability in Autism Spectrum Disorder. In S. W. White, B. B. Maddox, & C. A. Mazefsky (Eds.), *The Oxford handbook of autism and co-occurring psychiatric conditions / edited by* Oxford University Press.
Fritzsche, B., & Köpfer, A. (2022). (Para-)professionalism in dealing with structures of uncertainty - a cultural comparative study of teaching assistants in inclusion-oriented classrooms. *Disability & society*, *37*(6), 972-992. https://doi.org/10.1080/09687599.2020.1867068
Gagnon, L., Mottron, L., Bherer, L., & Joanette, Y. (2004). Quantification judgement in high functioning autism: Superior or different? *Journal of Autism and Developmental Disorders*, *34*(6), 679-689. https://doi.org/10.1007/s10803-004-5288-9
Gautier, M., & Harper, P. S. (2009). Fiftieth anniversary of trisomy 21: Returning to a discovery. *Human genetics*, *126*(2), 317-324. https://doi.org/10.1007/s00439-009-0690-1
Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*.
Gilmore, C. (2023). Understanding the complexities of mathematical cognition: A multi-level framework. *Quarterly Journal of Experimental Psychology*, *76*(9), 1953-1972. https://doi.org/10.1177/17470218231175325
Gilmore, C., Goebel, S. M., & Inglis, M. (2018). *An introduction to mathematical cognition* (First ed.). Routledge.
Goldstein, G., Minshew, N. J., & Siegel, D. J. (1994). Age differences in academic achievement in high-functioning autistic individuals. *Journal of Clinical and Experimental Neuropsychology*, *16*(5), 671-680. https://doi.org/10.1080/01688639408402680
Grieco, J., Pulsifer, M., Seligsohn, K., Skotko, B., & Schwartz, A. (2015). Down syndrome: Cognitive and behavioral functioning across the lifespan. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, *169*(2), 135-149. https://doi.org/10.1002/ajmg.c.31439
Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*, 665-668. https://doi.org/10.1038/nature07246

Hargreaves, S., Holton, S., Baxter, R., & Burgoyne, K. (2021). Educational experiences of pupils with Down syndrome in the UK. *Research in Developmental Disabilities*, *119*, Article 104115. https://doi.org/10.1016/j.ridd.2021.104115
Hermelin, B., & O'Connor, N. (1990). Factors and primes: A specific numerical ability. *Psychological Medicine*, *20*(1), 163-169. https://doi.org/http://dx.doi.org/10.1017/S0033291700013349
Hessels, R. S., Niehorster, D. C., Nyström, M., Andersson, R., Hooge, I. T. C., Leerstoel, K., Helmholtz, I., Experimental, P., Developmental, P., Social, personality development, A. t. a., & Leerstoel, P. (2018). Is the eye-movement field confused about fixations and saccades? A survey among 124 researchers. *Royal Society open science*, *5*(8), Article 180502. https://doi.org/10.1098/rsos.180502
Hiniker, A., Rosenberg-Lee, M., & Menon, V. (2016). Distinctive role of symbolic number sense in mediating the mathematical abilities of children with autism. *Journal of Autism and Developmental Disorders*, *46*(4), 1268-1281. https://doi.org/10.1007/s10803-015-2666-4
Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2009). Savant skills in autism: Psychometric approaches and parental reports. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1522), 1359-1367. https://doi.org/10.1098/rstb.2008.0328
Huttenlocher, J., Jordan, N. C., & Levine, S. C. (1994). A mental model for early arithmetic. *Journal of experimental psychology*, *123*(3), 284-296. https://doi.org/10.1037/0096-3445.123.3.284
Iuculano, T., Padmanabhan, A., Chen, L., Nicholas, J., Mitsven, S., de los Angeles, C., & Menon, V. (2020). Neural correlates of cognitive variability in childhood autism and relation to heterogeneity in decision-making dynamics. *Developmental Cognitive Neuroscience*, *42*, 12, Article 100754. https://doi.org/10.1016/j.dcn.2020.100754
Iuculano, T., Rosenberg-Lee, M., Supekar, K., Lynch, C. J., Khouzam, A., Phillips, J., Uddin, L. Q., & Menon, V. (2014). Brain organization underlying superior mathematical abilities in children with autism. *Biological Psychiatry*, *75*(3), 223-230. https://doi.org/10.1016/j.biopsych.2013.06.018
Jansen, B. R. J., Hofman, A. D., Straatemeier, M., Bers, B. M. C. W., Raijmakers, M. E. J., & Maas, H. L. J. (2014). The role of pattern recognition in children's exact enumeration of small numbers. *British Journal of Developmental Psychology*, *32*(2), 178-194. https://doi.org/10.1111/bjdp.12032
Jarrold, C., & Brock, J. (2004). To match or not to match? Methodological issues in autism-related research. *Journal of Autism and Developmental Disorders*, *34*(1), 81-86. https://doi.org/10.1023/B:JADD.0000018078.82542.ab
Jarrold, C., & Russell, J. (1997). Counting abilities in autism: Possible implications for central coherence theory. *Journal of Autism and Developmental Disorders*, *27*(1), 25-37. https://doi.org/10.1023/A:1025817121137
Jevons, W. S. (1871). The power of numerical discrimination. *Nature (London)*, *3*(71), 367-367. https://doi.org/10.1038/003367b0
Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research and Practice*, *22*(1), 36-46. https://doi.org/10.1111/j.1540-5826.2007.00229.x

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*(4), 441-480. https://doi.org/10.1016/0010-0285(76)90015-3

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329-354. https://doi.org/10.1037/0033-295X.87.4.329

Kabashi, L., & Kaczmarek, L. (2019). Educating a child with Down syndrome in an inclusive kindergarten classroom. *Journal of Childhood and Developmental Disorders*, *5*(4), 1-9. https://doi.org/10.36648/2472-1786.5.2.82

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2*, 217–250.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, *2*(10), 389-398. https://doi.org/10.1016/S1364-6613(98)01230-3

Karmiloff-Smith, A., Al-Janabi, T., D'Souza, H., Groet, J., Massand, E., Mok, K., Startin, C., Fisher, E., Hardy, J., Nizetic, D., Tybulewicz, V., & Strydom, A. (2016). The importance of understanding individual differences in Down syndrome. https://doi.org/10.12688/f1000research.7506.1

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *The American Journal of Psychology*, *62*(4), 498-525. https://doi.org/10.2307/1418556

Keen, D., Webster, A., & Ridley, G. (2016). How well are children with autism spectrum disorder doing academically at school? An overview of the literature. *Autism*, *20*(3), 276-294. https://doi.org/10.1177/1362361315580962

Kelly, S. J., Macaruso, P., & Sokol, S. M. (1997). Mental calculation in an autistic savant: A case study. *Journal of Clinical and Experimental Neuropsychology*, *19*(2), 172-184. https://doi.org/doi.org/10.1080/01688639708403848

Kennedy, A. (2016). Eye tracking: A comprehensive guide to methods and measures. In (Vol. 69, pp. 607-609). ABINGDON: Routledge.

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, *20*(4), 442-462. https://doi.org/10.1177/1362361315588200

Kim, S. H., Bal, V. H., & Lord, C. (2018). Longitudinal follow-up of academic achievement in children with autism from age 2 to 18. *Journal of Child Psychology and Psychiatry*, *59*(3), 258-267. https://doi.org/10.1111/jcpp.12808

King, S. A., Powell, S. R., Lemons, C. J., & Davidson, K. A. (2017). Comparison of mathematics performance of children and adolescents with and without Down syndrome. *Education and Training in Autism and Developmental Disabilities*, *52*(2), 208-222. www.jstor.org/stable/26420391

Kirk, H. E., Gray, K., Riby, D. M., Taffe, J., & Cornish, K. M. (2017). Visual attention and academic performance in children with developmental disabilities and behavioural attention deficits. *Developmental Science*, *20*(6), 12, Article e12468. https://doi.org/10.1111/desc.12468

Kljajevic, V. (2023). Literacy and numeracy in children on autism spectrum disorder. *Advances in Neurodevelopmental Disorders*, *7*(1), 123-129. https://doi.org/10.1007/s41252-022-00291-5

Kozel, B. A., Barak, B., Kim, C. A., Mervis, C. B., Osborne, L. R., Porter, M., & Pober, B. R. (2021). Williams syndrome. *Nature reviews. Disease primers*, *7*(1), 42-42. https://doi.org/10.1038/s41572-021-00276-z

Krajcsi, A., Szabo, E., & Morocz, I. A. (2013). Subitizing is sensitive to the arrangement of objects. *Experimental Psychology*, *60*(4), 227-234. https://doi.org/10.1027/1618-3169/a000191

Krajewski, K. (2009). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. *Learning and instruction*, *19*(6), 513-526. https://doi.org/10.1016/j.learninstruc.2008.10.002

Kroesbergen, E. H., Van Luit, J. E. H., Van Lieshout, E. C. D. M., Van Loosbroek, E., & Van de Rijt, B. A. M. (2009). Individual differences in early numeracy: The role of executive functions and subitizing. *Journal of Psychoeducational Assessment*, *27*(3), 226-236. https://doi.org/10.1177/0734282908330586

Kurth, J., & Mastergeorge, A. M. (2010). Academic and cognitive profiles of students with autism: Implications for classroom practice and placement. *International journal of special education*, *25*(2), 8-14.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174. https://doi.org/10.2307/2529310

Lashkari, A., Smith, A. K., & Graham, J. M. (1999). Williams-Beuren Syndrome: An update and review for the primary physician. *Clinical Pediatrics*, *38*(4), 189-208. https://doi.org/10.1177/000992289903800401

LeFevre, J.-A., Skwarchuk, S.-L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian journal of behavioural science*, *41*(2), 55-66. https://doi.org/10.1037/a0014532

Lettington, J., & Van Herwegen, J. (2024). Home literacy environment and literacy outcomes in individuals with Williams syndrome and Down syndrome. *Journal of Intellectual & Developmental Disability 49*(1), 67-83. https://doi.org/doi.org/10.3109/13668250.2023.2226917

Levine, S. C., Jordan, N. C., & Huttenlocher, J. (1992). Development of calculation abilities in young children. *Journal of Experimental Child Psychology*, *53*(1), 72-103. https://doi.org/10.1016/S0022-0965(05)80005-0

Levine, S. C., Suriyakham, L. W., Rowe, M. L., Huttenlocher, J., & Gunderson, E. A. (2010). What counts in the development of young children's number knowledge? *Developmental Psychology*, *46*(5), 1309-1319. https://doi.org/10.1037/a0019671

Li, X., Li, J., Zhao, S., Liao, Y., Zhu, L., & Mou, Y. (2023). Magnitude representation of preschool children with autism spectrum condition. *Autism*, 13623613231185408-13623613231185408. https://doi.org/10.1177/13623613231185408

Lincoln, A. J., Courchesne, E., Kilman, B. A., Elmasian, R., & Allen, M. (1988). A study of intellectual abilities in high-functioning people with autism. *Journal of Autism and Developmental Disorders*, *18*(4), 505-524. https://doi.org/10.1007/BF02211870

Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., Jones, E. J. H., Jones, R. M., Pickles, A., State, M. W., Taylor, J. L., & Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature reviews. Disease primers*, *6*(1), 5-5. https://doi.org/10.1038/s41572-019-0138-4

Losberg, J., & Zwozdiak-Myers, P. (2024). Inclusive pedagogy through the lens of primary teachers and teaching assistants in England. *International Journal of Inclusive Education*, *28*(4), 402-422. https://doi.org/10.1080/13603116.2021.1946722

Lusby, S., & Heinz, M. (2020). Shared reading practices between parents and young children with Down syndrome in Ireland. *Irish educational studies*, *39*(1), 19-38. https://doi.org/10.1080/03323315.2020.1729836

Marcelino, L., de Sousa, Ó., & Lopes, A. (2017). Predictive relation between early numerical competencies and mathematics achievement in first grade portuguese children. *Frontiers in Psychology*, *8*, 1103-1103. https://doi.org/10.3389/fpsyg.2017.01103

Mareschal, D. (2007). *Neuroconstructivism* (D. Mareschal, Ed.). Oxford University Press.

Martens, M. A., Wilson, S. J., & Reutens, D. C. (2008). Research review: Williams syndrome: A critical review of the cognitive, behavioral, and neuroanatomical phenotype. *J Child Psychol Psychiatry*, *49*(6), 576-608. https://doi.org/10.1111/j.1469-7610.2008.01887.x

Masi, A., DeMayo, M. M., Glozier, N., & Guastella, A. J. (2017). An overview of Autism Spectrum Disorder, heterogeneity and treatment options. *Neuroscience bulletin*, *33*(2), 183-193. https://doi.org/10.1007/s12264-017-0100-y

May, T., Rinehart, N., Wilding, J., & Cornish, K. (2013). The role of attention in the academic attainment of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *43*(9), 2147-2158. https://doi.org/10.1007/s10803-013-1766-2

May, T., Rinehart, N. J., Wilding, J., & Cornish, K. (2015). Attention and basic literacy and numeracy in children with autism spectrum disorder: A one-year follow-up study. *Research in Autism Spectrum Disorders*, *9*, 193-201. https://doi.org/10.1016/j.rasd.2014.10.010

Mayes, S. D., & Calhoun, S. L. (2003a). Ability profiles in children with autism: Influence of age and IQ. *Autism*, *7*(1), 65-80. https://doi.org/10.1177/1362361303007001014

Mayes, S. D., & Calhoun, S. L. (2003b). Analysis of WISC-III, Stanford-Binet : IV, and academic achievement test scores in children with autism. *Journal of Autism and Developmental Disorders*, *33*(3), 329-341. https://doi.org/10.1023/a:1024462719081

Mayes, S. D., & Calhoun, S. L. (2007). Learning, attention, writing, and processing speed in typical children and children with ADHD, autism, anxiety, depression, and oppositional-defiant disorder. *Child neuropsychology*, *13*(6), 469-493. https://doi.org/10.1080/09297040601112773

Mayes, S. D., & Calhoun, S. L. (2008). WISC-IV and WIAT-II profiles in children with high-functioning autism. *Journal of Autism and Developmental Disorders*, *38*(3), 428-439. https://doi.org/10.1007/s10803-007-0410-4

McCauley, J. B., Zajic, M. C., Oswald, T. M., Swain-Lerro, L. E., McIntyre, N. C., Harris, M. A., Trzesniewski, K., Mundy, P. C., & Solomon, M. (2018). Brief report: Investigating relations between self-concept and performance in reading and math for school-aged children and adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *48*(5), 1825-1832. https://doi.org/10.1007/s10803-017-3403-y

McDougal, E., Riby, D. M., & Hanley, M. (2020). Profiles of academic achievement and attention in children with and without autism spectrum disorder. *Research in Developmental Disabilities*, *106*, Article 103749. https://doi.org/10.1016/j.ridd.2020.103749
Mervis, C. B., & John, A. E. (2010). Cognitive and behavioral characteristics of children with Williams syndrome: Implications for intervention approaches. *American journal of medical genetics*, *154C*(2), 229-248. https://doi.org/10.1002/ajmg.c.30263
Mervis, C. B., Robinson, B. F., Bertrand, J., Morris, C. A., Klein-Tasman, B. P., & Armstrong, S. C. (2000). The Williams Syndrome cognitive profile. *Brain and Cognition*, *44*(3), 604-628. https://doi.org/10.1006/brcg.2000.1232
Miller, L. E., Burke, J. D., Troyb, E., Knoch, K., Herlihy, L. E., & Fein, D. A. (2017). Preschool predictors of school-age academic achievement in autism spectrum disorder. *Clinical Neuropsychologist*, *31*(2), 382-403. https://doi.org/10.1080/13854046.2016.1225665
Minshew, N. J., Goldstein, G., Taylor, H. G., & Siegel, D. J. (1994). Academic achievement in high functioning autistic individuals. *Journal of Clinical and Experimental Neuropsychology*, *16*(2), 261-270. https://doi.org/10.1080/01688639408402637
Minshew, N. J., Muenz, L. R., Goldstein, G., & Payton, J. B. (1992). Neuropsychological functioning in nonmentally-retarded autistic individuals. *Journal of Clinical and Experimental Neuropsychology*, *14*(5), 749-761. https://doi.org/10.1080/01688639208402860
Mock, J., Huber, S., Klein, E., & Moeller, K. (2016). Insights into numerical cognition: Considering eye-fixations in number processing and arithmetic. *Psychological Research*, *80*(3), 334-359. https://doi.org/10.1007/s00426-015-0739-9
Moeller, K., Neuburger, S., Kaufmann, L., Landerl, K., & Nuerk, H. C. (2009). Basic number processing deficits in developmental dyscalculia: Evidence from eye tracking. *Cognitive Development*, *24*(4), 371-386. https://doi.org/10.1016/j.cogdev.2009.09.007
Monari Martinez, E., & Benedetti, N. (2011). Learning mathematics in mainstream secondary schools: Experiences of students with Down's syndrome. *European Journal of Special Needs Education*, *26*(4), 531-540. https://doi.org/10.1080/08856257.2011.597179
Morris, C. A., Braddock, S. R., Chen, E., Trotter, T. L., Berry, S. A., Burke, L. W., Geleske, T. A., Hamid, R., Hopkin, R. J., Introne, W. J., Lyons, M. J., Scheuerle, A. E., Stoler, J. M., Freedenberg, D. L., Jones, M. C., Saul, R. A., & Tarini, B. A. (2020). Health care supervision for children with Williams syndrome. *Pediatrics*, *145*(2), 1. https://doi.org/10.1542/peds.2019-3761
Muhle, R. A., Reed, H. E., Stratigos, K. A., & Veenstra-VanderWeele, J. (2018). The emerging clinical neuroscience of autism spectrum disorder: A review. *JAMA Psychiatry*, *75*(5), 514-523. https://doi.org/10.1001/jamapsychiatry.2017.4685
Mutaf-Yildiz, B., Sasanguie, D., De Smedt, B., & Reynvoet, B. (2020). Probing the relationship between home numeracy and children's mathematical skills: A systematic review. *Frontiers in Psychology*, *11*, Article 2074. https://doi.org/10.3389/fpsyg.2020.02074
Myles, B. S., Barnhill, G. P., Hagiwara, T., Griswold, D. E., & Simpson, R. L. (2001). A synthesis of studies on the intellectual, academic, social/ emotional and sensory characteristics of children and youth with Asperger syndrome.

Education and training in mental retardation and developmental disabilities, *36*(3), 304-311. https://psycnet.apa.org/record/2001-11156-007
Myles, B. S., Simpson, R. L., & Becker, J. (1994). An analysis of characteristics of students diagnosed with higher-functioning autistic disorder. *Exceptionality*, *5*(1), 19-30. https://doi.org/10.1207/s15327035ex0501_2
National Organization for Rare Disorders. (2006). *Williams Syndrome*. https://rarediseases.org/rare-diseases/williams-syndrome/
NHS England. (2022). *National congenital anomaly and rare disease registration service: Congenital anomaly statistics report 2020*. https://digital.nhs.uk/data-and-information/publications/statistical/ncardrs-congenital-anomaly-statistics-annual-data/ncardrs-congenital-anomaly-statistics-report-2020
Nogues, C. P., & Dorneles, B. V. (2021). Systematic review on the precursors of initial mathematical performance. *International Journal of Educational Research Open*, *2*, Article 100035. https://doi.org/10.1016/j.ijedro.2021.100035
Nordahl-Hansen, A., Tondevold, M., & Fletcher-Watson, S. (2018). Mental health on screen: A DSM-5 dissection of portrayals of autism spectrum disorders in film and TV. *Psychiatry research*, *262*, 351-353. https://doi.org/10.1016/j.psychres.2017.08.050
O'Hearn, K., Franconeri, S., Wright, C., Minshew, N., & Luna, B. (2013). The development of individuation in autism. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(2), 494-509. https://doi.org/10.1037/a0029400
O'Hearn, K., Hoffman, J. E., & Landau, B. (2011). Small subitizing range in people with Williams syndrome. *Visual Cognition*, *19*(3), 289-312. https://doi.org/10.1080/13506285.2010.535994
O'Hearn, K., & Luna, B. (2009). Mathematical skills in Williams syndrome: Insight into the importance of underlying representations. *Developmental disabilities research reviews*, *15*(1), 11-20. https://doi.org/10.1002/ddrr.47
Ohta, M. (1987). Cognitive disorders of infantile autism: A study employing the WISC, spatial relationship conceptualization, and gesture imitations. *Journal of Autism and Developmental Disorders*, *17*(1), 45-62. https://doi.org/10.1007/BF01487259
Onnivello, S., Lanfranchi, S., & Zorzi, M. (2019). Mathematical abilities in Down syndrome. In S. Lanfranchi (Ed.), *International Review of Research in Developmental Disabilities* (Vol. 56, pp. 257-291). Academic Press.
Onnivello, S., Pulina, F., Locatelli, C., Marcolin, C., Ramacieri, G., Antonaros, F., Vione, B., Caracausi, M., & Lanfranchi, S. (2022). Cognitive profiles in children and adolescents with Down syndrome. *Scientific reports*, *12*(1), 1936-1936. https://doi.org/10.1038/s41598-022-05825-4
Oswald, T. M., Beck, J. S., Iosif, A. M., McCauley, J. B., Gilhooly, L. J., Matter, J. C., & Solomon, M. (2016). Clinical and cognitive characteristics associated with mathematics problem solving in adolescents with autism spectrum disorder. *Autism Research*, *9*(4), 480-490. https://doi.org/10.1002/aur.1524
Özdem, Ş., & Olkun, S. (2021). Improving mathematics achievement via conceptual subitizing skill training. *International Journal of Mathematical Education in Science and Technology*, *52*(4), 565-579. https://doi.org/10.1080/0020739X.2019.1694710

Ozonoff, S., South, M., & Miller, J. N. (2000). DSM-IV-defined Asperger syndrome: Cognitive, behavioral and early history differentiation from high-functioning autism. *Autism*, *4*(1), 29-46. https://doi.org/10.1177/1362361300041003

Paterson, S. J. (2000). *The development of language and number understanding in Williams Syndrome and Down's Syndrome: Evidence from the infant and mature phenotypes* [Unpublished doctoral thesis, University College London].

Paterson, S. J., Girelli, L., Butterworth, B., & Karmiloff-Smith, A. (2006). Are numerical impairments syndrome specific? Evidence from Williams syndrome and Down's syndrome. *Journal of Child Psychology and Psychiatry*, *47*(2), 190-204. https://doi.org/10.1111/j.1469-7610.2005.01460.x

Paul, J., Reeve, R., & Forte, J. D. (2017). Taking a(c)count of eye movements: Multiple mechanisms underlie fixations during enumeration. *Journal Of Vision*, *17*(3). https://doi.org/10.1167/17.3.16

Payne, A. C., Whitehurst, G. J., & Angell, A. L. (1994). The role of home literacy environment in the development of language ability in preschool children from low-income families. *Early childhood research quarterly*, *9*(3), 427-440. https://doi.org/10.1016/0885-2006(94)90018-3

Polo-Blanco, I., Suárez-Pinilla, P., Goñi-Cervera, J., Suárez-Pinilla, M., & Payá, B. (2024). Comparison of mathematics problem-solving abilities in autistic and non-autistic children: The influence of cognitive profile. *Journal of Autism and Developmental Disorders*, *54*(1), 353-365. https://doi.org/10.1007/s10803-022-05802-w

Porter, J. (2019). Discriminating quantity: New points for teaching children with Down syndrome about number? *International Journal of Disability Development and Education*, *66*(2), 133-150. https://doi.org/10.1080/1034912X.2019.1569208

Purser, H. (2015). Experiemntal difficulties in neurodevelopmental disorders. In J. Van Herwegen & D. Riby (Eds.), *Neurodevelopmental disorders: Research challenges and solutions*. Psychology Press.

Ranzato, E., Tolmie, A., & Van Herwegen, J. (2020). Perceptual subitizing and conceptual subitizing in Williams syndrome and Down syndrome: Insights from eye movements. *Research in Developmental Disabilities*, *106*, Article 103746. https://doi.org/10.1016/j.ridd.2020.103746

Ranzato, E., Tolmie, A., & Van Herwegen, J. (2021). The home learning environment of primary school children with Down syndrome and those with Williams syndrome. *Brain Sciences*, *11*(6). https://doi.org/doi.org/10.3390/brainsci11060733

Raven, J. C. (1993). *Coloured progressive matrices*. Information Press Ltd.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422. https://doi.org/10.1037/0033-2909.124.3.372

Reigosa-Crespo, V., Gonzalez-Alemany, E., Leon, T., Torres, R., Mosquera, R., & Valdes-Sosa, M. (2013). Numerical capacities as domain-specific predictors beyond early mathematics learning: A longitudinal study. *PLoS ONE*, *8*(11), Article e79711. https://doi.org/10.1371/journal.pone.0079711

Ricci, L. (2011). Home literacy environments, interest in reading and emergent literacy skills of children with Down syndrome versus typical children. *Journal of Intellectual Disability Research*, *55*(6), 596. https://doi.org/10.1111/j.1365-2788.2011.01415.x

Ricci, L., & Osipova, A. (2012). Visions for literacy: Parents' aspirations for reading in children with Down syndrome. *British Journal of Special Education*, *39*(3), 123-129. https://doi.org/10.1111/j.1467-8578.2012.00547.x
Rietveld, C. M. (2005). Classroom learning experiences of mathematics by new entrant children with Down syndrome. *Journal of intellectual & developmental disability*, *30*(3), 127-138. https://doi.org/10.1080/13668250500204117
Root, J. R., Henning, B., & Boccumini, E. (2018). Teaching students with autism and intellectual disability to solve algebraic word problems [Article]. *Education and Training in Autism and Developmental Disabilities*, *53*(3), 325-338. https://doi.org/10.1177/00400599221116821
Ruel, E. E., Wagner, W. E., & Gillespie, B. J. (2016). *The practice of survey research: Theory and applications*. SAGE Publications.
Rumsey, J. M., & Hamburger, S. D. (1988). Neuropsychological findings in high-functioning men with infantile autism, residual state. *Journal of Clinical and Experimental Neuropsychology*, *10*(2), 201-221. https://doi.org/10.1080/01688638808408236
Russell, G., Mandy, W., Elliott, D., White, R., Pittwood, T., & Ford, T. (2019). Selection bias on intellectual ability in autism research: A cross-sectional review and meta-analysis. *Molecular autism*, *10*(9). https://doi.org/doi.org/10.1186/s13229-019-0260-x
Sacks, O. (1986). The Twins. In *The man who mistook his wife for a hat*. Picador.
Santos, F. H., Ribeiro, F. S., Dias-Piovezana, A. L., Primi, C., Dowker, A., & von Aster, M. (2022). Discerning developmental dyscalculia and neurodevelopmental models of numerical cognition in a disadvantaged educational context. *Brain Sciences*, *12*(5), 653. https://doi.org/10.3390/brainsci12050653
Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. Routledge.
Satsangi, R., & Bofferding, L. (2017). Improving the numerical knowledge of children with autism spectrum disorder: The benefits of linear board games. *Journal of Research in Special Educational Needs*, *17*(3), 218-226. https://doi.org/10.1111/1471-3802.12380
Schindler, M., Schovenberg, V., & Schabmann, A. (2020). Enumeration processes of children with mathematical difficulties: An explorative eye-tracking study on subitizing, groupitizing, counting, and pattern recognition. *Learning Disabilities*, *18*(2), 193-211. https://eric.ed.gov/?id=EJ1281060
Schleifer, P., & Landerl, K. (2011). Subitizing and counting in typical and atypical development. *Developmental Science*, *14*(2), 280-291. https://doi.org/10.1111/j.1467-7687.2010.00976.x
Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3), Article e12372-n/a. https://doi.org/10.1111/desc.12372
Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta-analysis. *Child Development*, *89*(5), 1467-1484. https://doi.org/10.1111/cdev.13068

Sella, F., Lanfranchi, S., & Zorzi, M. (2013). Enumeration skills in Down syndrome. *Research in Developmental Disabilities*, *34*(11), 3798-3806. https://doi.org/10.1016/j.ridd.2013.07.038
Senechal, M., & LeFevre, J.-A. (2014). Continuity and change in the home literacy environment as predictors of growth in vocabulary and reading. *Child Development*, *85*(4), 1552-1568. https://doi.org/10.1111/cdev.12222
Senechal, M., LeFevre, J.-A., Thomas, E., M., & Daley, K., E. . (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly*, *33*(1), 96-116. https://doi.org/10.1598/RRQ.33.1.5
Sharafi, Z., Sharif, B., Guéhéneuc, Y.-G., Begel, A., Bednarik, R., & Crosby, M. (2020). A practical guide on conducting eye tracking studies in software engineering. *Empirical software engineering*, *25*(5), 3128-3174. https://doi.org/10.1007/s10664-020-09829-4
Siegler, R., S., & Opfer, J., E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*(3), 237-243. https://doi.org/10.1111/1467-9280.02438
Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*(2), 428-444. https://doi.org/10.1111/j.1467-8624.2004.00684.x
Silverman, W. (2007). Down syndrome: Cognitive phenotype. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*(3), 228-236. https://doi.org/10.1002/mrdd.20156
Simms, V., Clayton, S., Cragg, L., Gilmore, C., & Johnson, S. (2016). Explaining the relationship between number line estimation and mathematical achievement: The role of visuomotor integration and visuospatial skills. *Journal of Experimental Child Psychology*, *145*, 22-33. https://doi.org/10.1016/j.jecp.2015.12.004
Simms, V., Karmiloff-Smith, A., Ranzato, E., & Van Herwegen, J. (2020). Understanding number line estimation in Williams syndrome and Down syndrome. *Journal of Autism and Developmental Disorders*, *50*(2), 583-591. https://doi.org/10.1007/s10803-019-04268-7
Soulieres, I., Hubert, B., Rouleau, N., Gagnon, L., Tremblay, P., Seron, X., & Mottron, L. (2010). Superior estimation abilities in two autistic spectrum children. *Cognitive Neuropsychology*, *27*(3), 261-276. https://doi.org/10.1080/02643294.2010.519228
Spaniol, M., & Danielsson, H. (2022). A meta-analysis of the executive function components inhibition, shifting, and attention in intellectual disabilities. *Journal of Intellectual Disability Research*, *66*(1-2), 9-31. https://doi.org/10.1111/jir.12878
Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive behavior scales.* (Second ed.). PsychCorp.
Spek, A. A., Scholte, E. M., & van Berckelaer-Onnes, I. A. (2008). Brief report: The use of WAIS-III in adults with HFA and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *38*(4), 782-787. https://doi.org/10.1007/s10803-007-0446-5
Starkey, G. S., & McCandliss, B. D. (2014). The emergence of "groupitizing" in children's numerical cognition. *Journal of Experimental Child Psychology*, *126*, 120-137. https://doi.org/10.1016/j.jecp.2014.03.006

Steel, J., Gorman, R., & Flexman, J. E. (1984). Neuropsychiatric testing in an autistic mathematical idiot-savant: Evidence for nonverbal abstract capacity [Clinical Case Study; Empirical Study]. *Journal of the American Academy of Child Psychiatry*, *23*(6), 704-707. https://doi.org/10.1016/S0002-7138%2809%2960540-9

Strohmaier, A. R., MacKay, K. J., Obersteiner, A., & Reiss, K. M. (2020). Eye-tracking methodology in mathematics education research: A systematic literature review. *Educational studies in mathematics*, *104*(2), 147-200. https://doi.org/10.1007/s10649-020-09948-1

Stromme, P., Bjomstad, P. G., & Ramstad, K. (2002). Prevalence estimation of Williams syndrome. *Journal of child neurology*, *17*(4), 269-271. https://doi.org/10.1177/088307380201700406

Thioux, M., Stark, D. E., Klaiman, C., & Schultz, R. T. (2006). The day of the week when you were born in 700 ms: Calendar computation in an autistic savant. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1155-1168. https://doi.org/http://dx.doi.org/10.1037/0096-1523.32.5.1155

Thomas, M. S. C., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using developmental trajectories to understand developmental disorders. *Journal of speech, language, and hearing research*, *52*(2), 336-358. https://doi.org/10.1044/1092-4388(2009/07-0144)

Thomas, M. S. C., Ojinaga Alfageme, O., D'Souza, H., Patkee, P. A., Rutherford, M. A., Mok, K. Y., Hardy, J., Karmiloff-Smith, A., & LonDown, S. C. (2020). A multi-level developmental approach to exploring individual differences in Down syndrome: Genes, brain, behaviour, and environment. *Research in Developmental Disabilities*(104), Article 103638. https://doi.org/10.1016/j.ridd.2020.103638

Titeca, D., Roeyers, H., Ceulemans, A., & Desoete, A. (2015). Early numerical competencies in 5- and 6-year-old children with autism spectrum disorder. *Early Education and Development*, *26*(7), 1012-1034. https://doi.org/10.1080/10409289.2015.1004515

Titeca, D., Roeyers, H., & Desoete, A. (2015). Early numerical competencies in 4- and 5-year-old children with autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities*, *32*(4), 279-292. https://doi.org/10.1177/1088357615588523

Titeca, D., Roeyers, H., Josephy, H., Ceulemans, A., & Desoete, A. (2014). Preschool predictors of mathematics in first grade children with autism spectrum disorder. *Research in Developmental Disabilities*, *35*(11), 2714-2727. https://doi.org/10.1016/j.ridd.2014.07.012

Titeca, D., Roeyers, H., Loeys, T., Ceulemans, A., & Desoete, A. (2015). Mathematical abilities in elementary school children with autism spectrum disorder. *Infant and Child Development*, *24*(6), 606-623. https://doi.org/10.1002/icd.1909

Tobii AB. (2023). *Calibration*. https://developer.tobiipro.com/commonconcepts/calibration.html

Tonizzi, I., & Usai, M. C. (2023). Math abilities in autism spectrum disorder: A meta-analysis. *Research in Developmental Disabilities*, *139*, Article 104559. https://doi.org/10.1016/j.ridd.2023.104559

Trenholm, B., & Mirenda, P. (2006). Home and community literacy experiences of individuals with Down syndrome. *Down Syndrome Research and Practice*, *10*(1), 30-40. https://doi.org/doi:10.3104/reports.303
Troyb, E., Orinstein, A., Tyson, K., Helt, M., Eigsti, I. M., Stevens, M., & Fein, D. (2014). Academic abilities in children and adolescents with a history of autism spectrum disorders who have achieved optimal outcomes. *Autism*, *18*(3), 233-243. https://doi.org/10.1177/1362361312473519
van Bysterveldt, A., Gillon, G., & Foster-Cohen, S. (2010). Literacy environments for children with Down syndrome: What's happening at home? *Down Syndrome Research and Practice*, *12*(2), 98-102. https://doi.org/doi:10.3104/reports.2111
Van Herwegen, J. (2015). Williams syndrome and its cognitive profile: The importance of eye movements. *Psychology Research and Behavior Management*, *8*, 143-151. https://doi.org/10.2147/PRBM.S63474
Van Herwegen, J., Ansari, D., Xu, F., & Karmiloff-Smith, A. (2008). Small and large number processing in infants and toddlers with Williams syndrome. *Developmental Science*, *11*(5), 637-643. https://doi.org/10.1111/j.1467-7687.2008.00711.x
Van Herwegen, J., Ashworth, M., & Palikara, O. (2018). Parental views on special educational needs provision: Cross-syndrome comparisons in Williams syndrome, Down syndrome, and sutism spectrum disorders. *Research in Developmental Disabilities*, *80*, 102-111. https://doi.org/10.1016/j.ridd.2018.06.014
Van Herwegen, J., Farran, E., & Annaz, D. (2011). Item and error analysis on Raven's Coloured Progressive Matrices in Williams syndrome. *Research in Developmental Disabilities*, *32*(1), 93-99. https://doi.org/10.1016/j.ridd.2010.09.005
Van Herwegen, J., & Karmiloff-Smith, A. (2015). Genetic developmental disorders and numerical competence across the lifespan. In R. C. Kadosh & A. Dowker (Eds.), *Numerical cognition*. Oxford University Press.
Van Herwegen, J., Ranzato, E., Karmiloff-Smith, A., & Simms, V. (2019). Eye movement patterns and approximate number sense task performance in Williams syndrome and Down syndrome: A developmental perspective. *Journal of Autism and Developmental Disorders*, *49*, 4030–4038. https://doi.org/10.1007/s10803-019-04110-0
Van Herwegen, J., Ranzato, E., Karmiloff-Smith, A., & Simms, V. (2020). The foundations of mathematical development in Williams syndrome and Down syndrome. *Journal of Applied Research in Intellectual Disabilities*, *33*(5), 1080-1089. https://doi.org/10.1111/jar.12730
Van Herwegen, J., & Simms, V. (2020). Mathematical development in Williams syndrome: A systematic review. *Research in Developmental Disabilities*, *100*, Article 103609. https://doi.org/10.1016/j.ridd.2020.103609
Viñuela-Navarro, V., Erichsen, J., Little, J., Saunders, K. J., & Woodhouse, M. L. (2017, June). Evidence for fixation instability in children with Down syndrome. Child Vision Research Society XVI conference, Ulster University, Coleraine.
Vostanis, A., Padden, C., McTiernan, A., & Langdon, P. E. (2023). Comparing the Minimum Celeration Line and the Beat Your Personal Best Goal-Setting Approaches During the Mathematical Practice of Students Diagnosed with Autism. *Journal of Behavioral Education*, *32*(1), 21-50. https://doi.org/10.1007/s10864-021-09432-7

Wang, L., Liang, X., Jiang, B., Wu, Q., & Jiang, L. (2023). What ability can predict mathematics performance in typically developing preschoolers and those with autism spectrum disorder? *Journal of Autism and Developmental Disorders*, *53*(5), 2062-2077. https://doi.org/10.1007/s10803-022-05454-w

Watson, D. G., Maylor, E. A., & Bruce, L. A. M. (2007). The role of eye movements in subitizing and counting. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(6), 1389-1399. https://doi.org/10.1037/0096-1523.33.6.1389

Wechsler, D. (2005). *Wechsler Individual Achievement Test* (2nd ed.). Psychological Corporation.

Whitby, P. J. S., & Mancil, G. R. (2009). Academic achievement profiles of children with high functioning autism and asperger syndrome: A review of the literature. *Education and Training in Developmental Disabilities*, *44*(4), 551-560. https://www.jstor.org/stable/24234262

Wilkey, E. D. (2023). The domain-specificity of domain-generality: Attention, executive function, and academic skills. *Mind, brain and education*, *17*(4), 349-361. https://doi.org/10.1111/mbe.12373

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders*, *9*(1), 11-29. https://doi.org/10.1007/BF01531288

Wren, A. (2017). Understanding the role of the Teaching Assistant: Comparing the views of pupils with SEN and TAs within mainstream primary schools. *Support for Learning*, *32*(1), 4-19. https://doi.org/10.1111/1467-9604.12151

Wu, M.-J., Zhao, K., & Fils-Aime, F. (2022). Response rates of online surveys in published research: A meta-analysis. *Computers in human behavior reports*, *7*, Article 100206. https://doi.org/10.1016/j.chbr.2022.100206

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition: International Journal of Cognitive Science*, *89*(1), B15-B25. https://doi.org/10.1016/S0010-0277(03)00050-7

Zielińska, M., Sterczyński, R., & Bagińska, A. (2014). The profile of WISC-R scores in children with high-functioning autism. *Psychiatria Polska*, *48*(4), 667-676. https://pubmed.ncbi.nlm.nih.gov/25314795/#:~:text=Children%20with%20autism%20scored%20lower,and%20%22Block%20design%22%20scales.

Zimpel, A. F., & Rieckmann, T. (2022). The influence of Trisomy 21 on subitising imit. *International journal of disability, development, and education*, *69*(3), 987-1005. https://doi.org/10.1080/1034912X.2020.1737317

Zippert, E. L., & Rittle-Johnson, B. (2020). The home math environment: More than numeracy. *Early childhood research quarterly*, *50*, 4-15. https://doi.org/10.1016/j.ecresq.2018.07.009

Appendix A

Copy of the Maths at Home survey

Start of Block: Introduction

Dear Parent,

My name is Erica, and I am a Researcher working in collaboration with Dr Jo Van Herwegen at Kingston University London.

Thank you for your interest in the "Maths at Home" project. Many studies highlight that Home Learning Environment plays an important role in children's development. In this study we focus on maths and Home Numeracy Environment (HNE). Your answers will help us support parents in providing a positive and effective HNE for their children.
We kindly ask you to complete the following questionnaire, ticking the choice that best describes your family. There are no right or wrong answers. You are free to withdraw from the survey up until 30/06/2019 and you are free to omit any question. However, missing answers will make it more difficult to analyze the data and to publish the study.

This questionnaire will take approximately 30 minutes to complete. Please be assured that all personal details and data related to the questionnaire will be treated strictly confidentially and that your answers will be used for this research only. All data will be collected by password protected Qualtrics account that only the research team has access to. A copy of the collected data will be stored in an anonymised database on an encrypted hard drive. All data will be deleted after 5 years. This project has been reviewed according to procedures specified by Kingston University London and has been allowed to proceed (1718CHA12).

Many thanks in advance!

Should you have any further questions, please do not hesitate to contact us:

Principal Investigator	**Researcher**
Dr Jo Van Herwegen	Erica Ranzato
[redacted]	[redacted]

End of Block: Introduction

Start of Block: Consent (compulsory section)

Please check the boxes below to confirm that:

- ☐ I read and understood the information about what this questionnaire entails
- ☐ I consent to take part in this study
- ☐ I understand my participation in this study is voluntary and that I may withdraw up until 30/06/2019
- ☐ I understand all personal information will be treated strictly confidentially
- ☐ I understand I have the opportunity to ask questions

End of Block: Consent

Start of Block: Questionnaire Description

The following questionnaire has 8 sections.

If you do stop the questionnaire at any point, the responses you have already given will be saved and you can pick up from where you left off at another time by **using the same link to the questionnaire**. Please ensure that you use the **same computer and web browser** and complete the entire questionnaire **within 2 weeks** or your answers will not be saved.

If there are any questions you feel require a more detailed answer, we would be happy to hear from you. Please use the contact details provided to get in touch

Erica Ranzato [redacted]

End of Block: Questionnaire Description

Start of Block: SECTION 1. About your child

SECTION ONE

In this section we will ask some questions about your child.

Q1. [compulsory] What is your child's date of birth? (DD/MM/YYYY)

____

Q2. What is your child's gender?

○ Male

○ Female

Q3. Has your child been officially diagnosed with a developmental disorder?

○ My child has not been diagnosed with a developmental disorder

○ Down Syndrome

○ Williams Syndrome

○ Other, please specify ____

Skip To: Q6 If Has your child been officially diagnosed with a developmental disorder? = My child has not been diagnosed with a developmental disorder

Q4. What type of schooling is your child attending at the moment?

- ◯ Mainstream school
- ◯ Mainstream school with special educational unit on site
- ◯ Special educational needs school
- ◯ Home educated
- ◯ Other, please specify

____

Q5. What additional support has your child received in the last academic year? Tick all that apply

- ☐ Speech and Language Therapy (SLT or SALT)
- ☐ Physiotherapy
- ☐ Occupational Therapy (OT)
- ☐ Special Educational Needs (SEN) support
- ☐ Music therapy/music lessons
- ☐ Visual supports
- ☐ Life skills teaching
- ☐ Extra reading help/phonics
- ☐ Other, please specify

____

Q6. What level of **maths** is your child functioning at in school? E.g., if the child is working towards the national curriculum: a) working towards the expected standard, b) working at expectations, c) working at greater depth within the expected standard. If the child is working below the standard of the national curriculum: P-scales.

[short paragraph]

Q7. Compared to your child's overall abilities, how do you consider your child's **mathematical abilities**?

○ My child's mathematical abilities are better than his/her overall abilities

○ My child's mathematical abilities are in line with his/her overall abilities

○ My child's mathematical abilities are worse than his/her overall abilities

Q8. If you have any further comments about your child's mathematical abilities, please use the box below.

[short paragraph]

Q9. What level of **reading** is your child functioning at in school? E.g., if the child is working towards the national curriculum: a) working towards the expected standard, b) working at expectations, c) working at greater depth within the expected standard. If the child is working below the standard of the national curriculum: P-scales.

[short paragraph]

Q10. Compared to your child's overall abilities, how do you consider your child's **reading abilities**?

- ○ My child's reading abilities are better than his/her overall abilities
- ○ My child's reading abilities are in line with his/her overall abilities
- ○ My child's reading abilities are worse than his/her overall abilities

Q11. If you have any further comments about your child's reading abilities, please state them in the box below

[short paragraph]

Q12. What level of **writing** is your child functioning at in school? E.g., if the child is working towards the national curriculum: a) working towards the expected standard, b) working at expectations, c) working at greater depth within the expected standard. If the child is working below the standard of the national curriculum: P-scales.

[short paragraph]

Q13. Compared to your child’s overall abilities, how do you consider your child’s **writing abilities**?

- ◯ My child's writing abilities are better than his/her overall abilities
- ◯ My child's writing abilities are in line with his/her overall abilities
- ◯ My child's writing abilities are worse than his/her overall abilities

---

Q14. If you have any further comments about your child's writing abilities, please state them in the box below.

[short paragraph]

End of Block: SECTION 1. About your child

Start of Block: SECTION 2. Frequency of household activities

SECTION TWO

These questions ask about the type and the **frequency** of household learning activities you are involved in with your child.

Q15. **How often** do you engage in the following activities together with your child? Rate your answer on a scale from "Never" to "Every day". If you feel that the activity is not appropriate for the stage of development of your child, select the option "Not appropriate".

	Never	Less than once a week	Once a week	More than once a week	Every day	Not appropriate
Drawing	○	○	○	○	○	○
Worksheets on addition and subtraction	○	○	○	○	○	○
Paying attention to letters and/or words during daily activities (e.g., cooking)	○	○	○	○	○	○
Writing letters and/or words (e.g., writing birthday cards)	○	○	○	○	○	○
Playing sports	○	○	○	○	○	○
Using number activity books	○	○	○	○	○	○
Playing mental orientation games (e.g., games that include walking through virtual worlds)	○	○	○	○	○	○

Q16. How often do you engage in the following activities together with your child?

	Never	Less than once a week	Once a week	More than once a week	Every day	Not appropriate
Doing shopping	○	○	○	○	○	○
Using Numicon resources	○	○	○	○	○	○
Playing games that include writing and/or reading (e.g., Fishbowl game)	○	○	○	○	○	○
Handling and naming common 2-D or 3-D shapes	○	○	○	○	○	○
Doing maths homework	○	○	○	○	○	○
Playing dominoes	○	○	○	○	○	○
Playing estimation games (e.g., "Guess which one is more?")	○	○	○	○	○	○

Q17. How often do you engage in the following activities together with your child?

	Never	Less than once a week	Once a week	More than once a week	Every day	Not appropriate
Practicing ordering (e.g., first, second, third)	○	○	○	○	○	○
Playing memory games (e.g., Shopping List)	○	○	○	○	○	○
Talking about money when shopping	○	○	○	○	○	○
Playing jigsaw puzzles	○	○	○	○	○	○
Elementary calculations during daily activities (e.g., "There are five apples in the fruit bowl. If I take one, how many apples are left?")	○	○	○	○	○	○
Singing number songs together (e.g., five little monkeys)	○	○	○	○	○	○
Writing/typing your child’s name	○	○	○	○	○	○

Q18. How often do you engage in the following activities together with your child?

	Never	Less than once a week	Once a week	More than once a week	Every day	Not appropriate
Doing “connect the dots” activities	○	○	○	○	○	○
Using public transports together	○	○	○	○	○	○
Reading books	○	○	○	○	○	○
Creating patterns with concrete materials (e.g., creating a necklace alternating red and blue beads)	○	○	○	○	○	○
Watching TV	○	○	○	○	○	○
Playing with building blocks – such as Lego	○	○	○	○	○	○
Telling the time	○	○	○	○	○	○

Q19. How often do you engage in the following activities together with your child?

	Never	Less than once a week	Once a week	More than once a week	Every day	Not appropriate
Using number flashcards	○	○	○	○	○	○
Playing board games that require elementary computations (e.g., with two dice)	○	○	○	○	○	○
Listening to music	○	○	○	○	○	○
Playing with toys/video games together	○	○	○	○	○	○
Using sticker reward charts	○	○	○	○	○	○
Watching literacy education programs and TV shows	○	○	○	○	○	○
Recognising and finding half of a quantity, length, set of objects or shape (e.g., “can I have half of your sweets?”)	○	○	○	○	○	○

Q20. How often do you engage in the following activities together with your child?

	Never	Less than once a week	Once a week	More than once a week	Every day	Not appropriate
Cooking together	○	○	○	○	○	○
Using measuring tools such as a ruler when drawing or a scale when cooking	○	○	○	○	○	○
Counting during daily activities (e.g., counting the number of apples when cooking)	○	○	○	○	○	○
Learning new words during daily activities	○	○	○	○	○	○
Watching mathematical educational programs and TV shows	○	○	○	○	○	○
Playing pairs games	○	○	○	○	○	○
Reading number story books that include numbers or counting	○	○	○	○	○	○

Q21. If you have any further comments on your household activities, please write them in the space provided below.

[short paragraph]

End of Block: SECTION 2. Frequency of household activities

Start of Block: SECTION 3. Technology

SECTION THREE

This section of the questionnaire is about how your child uses new technology in relation to numeracy activities.

Q22. Does your child use technology **at home** - e.g., tablets, computers, smartphones, video game consoles?

○ Yes

○ No

Skip To: End of Block If Does your child use technology at home - e.g., tablets, computers, smartphones, video game consoles? = No

---

Q23. Does your child own his/her own tablet/iPad?

○ Yes

○ No

Q24. How often does your child engage in the following activities **at home**? Rate your answer on a scale from "Never" to "Daily".

	Never	Once a month	Once a week	2-3 times a week	Daily
Has access to iPads, tablets, smartphones, or computers	○	○	○	○	○
Watches videos, television programs and cartoons on YouTube	○	○	○	○	○
Uses drawing apps	○	○	○	○	○
Plays video games	○	○	○	○	○
Uses apps such as Skype and Facetime	○	○	○	○	○
Reads e-books	○	○	○	○	○

Q25. In the **past month**, how often did your child engage in the following? Rate your answer on a scale from "Activity did not occur" to "Almost daily". If you feel that the activity is not appropriate for the stage of development of your child, select the option "Not appropriate".

	Activity did not occur	Few times a month	Few times a week	Almost daily	Not appropriate
Counting apps	◯	◯	◯	◯	◯
Racing games	◯	◯	◯	◯	◯
Size/matching apps	◯	◯	◯	◯	◯
Additions and subtractions games	◯	◯	◯	◯	◯
Digital puzzle games	◯	◯	◯	◯	◯
'Filling in the gap' number games (e.g., what is the next in the sequence?)	◯	◯	◯	◯	◯
Number recognition apps	◯	◯	◯	◯	◯
Maths related websites	◯	◯	◯	◯	◯

Q26. Please, list the name of 3 of **your child's favourite** apps at the moment.

[short paragraph]

Q27. Please, list the name of 3 of **your favourite** educational apps.

[short paragraph]

Q28. Do you have any **concern** about your child's use of technology?

- ☐ I am concerned about undesirable and/or not age-appropriate content
- ☐ I am concerned about the time my child spends on the screen
- ☐ I am concerned about accidental in-app purchases
- ☐ I am concerned about the effectiveness of these apps
- ☐ Other, please specify ________________
- ☐ I don't have any concern

Q29. If you have any further comments or concerns on the use of technology in relation to your child's learning, please write them in the space provided below.

[short paragraph]

End of Block: SECTION 3. Technology

Start of Block: SECTION 4. Expectations

SECTION FOUR

This section is about your **expectations as a parent**.

Q30. How well do you expect your child to master the following abilities **by the end of the primary school**? Rate your answer on a scale from 0 to 10, where 0 is "Not at all" and 10 is "Very well".

Not at all

Very well

0 1 2 3 4 5 6 7 8 9 10

Putting on a coat	
Understanding and following spatial directions	
Boiling an egg	
Recognising odd and even numbers	
Getting dressed in the morning independently	
Opening locks	
Reading books	
Solving problems involving the calculation of percentages	
Counting up to 100	

Q31. How well do you expect your child to master the following abilities **by the end of the primary school**?

Not at all

Very well

0 1 2 3 4 5 6 7 8 9 10

Writing full sentences with different forms (e.g., statement, question)	
Sorting objects by size, colour or shape	
Reading the time	
Planning and preparing for a task (e.g., collecting equipment needed for school)	
Tying shoelaces	
Stringing beads	
Reciting the alphabet	
Converting between miles and kilometres	
Rounding a number to the nearest 10, 100 or 1000	

Q32. How well do you expect your child to master the following abilities **by the end of the primary school**?

Not at all

Very well

0 1 2 3 4 5 6 7 8 9 10

Adding and subtracting numbers mentally	
Using microwave for heating or cooking (e.g., setting the time and the power)	
Using money to buy something	
Solving two single digit number sums (e.g. 3 + 6)	
Reading letter symbols (e.g., “a” for apple)	
Writing numbers from 1 to 20 in words	
Using measurements (e.g., baking, height)	
Consistent use of present tense and past tense	
Retelling past events	

Q33. How well do you expect your child to master the following abilities **by the end of the primary school**?

Not at all

Very well

0 1 2 3 4 5 6 7 8 9 1
0

Counting backwards from 20	
Placing numbers on a number line	
Remembering what has occurred recently (e.g., who has phoned)	
Calculating the area of a rectangle	
Knowing and using number bonds within 20	
Counting up to 20 objects (e.g., counting 13 candies)	
Writing or typing letter symbols (e.g., k for key)	
Toilet trained	
Reading written number symbols up to 20	

Q34. How well do you expect your child to master the following abilities **by the end of the primary school**?

Not at all

Very well

0 1 2 3 4 5 6 7 8 9 10

Recognising dice patterns	
Accurate use of pronouns in sentences	
Learning multiplication tables by heart	
Writing number symbols up to 10 (e.g., 2)	
Eating using knife and fork	
Doing shopping independently	
Buttoning and unbuttoning	
Solving simple puzzles	
Writing his/her own name	

Q35. How well do you expect your child to master the following abilities **by the end of the primary school**?

Not at all

Very well

0 1 2 3 4 5 6 7 8 9 10

Comparing quantities (e.g., “which one is more?”)	
Using public transports independently	
Reciting the number sequence up to 10	
Given a number up to 20, identifying one more and one less	
Using a toaster	
Comparing the size of objects (e.g., "which one is bigger?")	
Using efficient calculation strategies (e.g., 25 - 8 = 25 - 5 - 3)	
Remembering information about personal data (e.g., age)	
Interpreting and constructing pie charts	

Q36. How well do you expect your child to master the following abilities **by the end of the primary school**?

Not at all

Very well

	0	1	2	3	4	5	6	7	8	9	10

Counting up and down in tens (e.g., 10, 20, 30)	
Extending a pattern or a sequence (e.g., red, blue, green, red, blue, green, etc.)	
Having a best friend	
Recognising and naming common 2-D (e.g., circles) and 3-D shapes (e.g., spheres)	
Following multiple-step instructions to complete a task (e.g., emptying the dishwasher and loading it with dirty dishes)	
Riding a bicycle	
Solving simple subtractions (e.g., 9 - 2)	

Q37. If you have any comments regarding your expectations about your child's abilities, please write them in the space provided below.

[short paragraph]

End of Block: SECTION 4. Expectations

Start of Block: SECTION 5. Importance

SECTION FIVE

We previously asked about your expectations as a parent. This section asks **how important** you consider those competences.

Q38. How **important** is it for you that your child masters the following competencies **at the end of the primary school**? Rate your answer on a scale from 0 to 10, where 0 is "Not important at all" and 10 is "Very important".

Not important at all										Very important
0	1	2	3	4	5	6	7	8	9	10

Putting on a coat	
Understanding and following spatial directions	
Boiling an egg	
Recognising odd and even numbers	
Getting dressed in the morning independently	
Opening locks	
Reading books	
Solving problems involving the calculation of percentages	
Counting up to 100	

Q39. How **important** is it for you that your child masters the following competencies **at the end of the primary school**?

	Not important at all										Very important
	0	1	2	3	4	5	6	7	8	9	10
Writing full sentences with different forms (e.g., statement, question)											
Sorting objects by size, colour or shape											
Reading the time											
Planning and preparing for a task (e.g., collecting equipment needed for school)											
Tying shoelaces											
Stringing beads											
Reciting the alphabet											
Converting between miles and kilometres											
Rounding a number to the nearest 10, 100 or 1000											

Q40. How **important** is it for you that your child masters the following competencies **at the end of the primary school**?

Not important at all — Very important

0 1 2 3 4 5 6 7 8 9 10

Adding and subtracting numbers mentally	
Using microwave for heating or cooking (e.g., setting the time and the power)	
Using money to buy something	
Solving two single digit number sums (e.g., 3 + 6)	
Reading letter symbols (e.g., “a” for apple)	
Writing numbers from 1 to 20 in words	
Using measurements (e.g., baking, height)	
Consistent use of present tense and past tens	
Retelling past events	

Q41. How **important** is it for you that your child masters the following competencies **at the end of the primary school**?

Not important at all — Very important

0 1 2 3 4 5 6 7 8 9 10

Counting backwards from 20	
Placing numbers on a number line	
Remembering what has occurred recently (e.g., who has phoned)	
Calculating the area of a rectangle	
Knowing and using number bonds within 20	
Counting up to 20 objects (e.g., counting 13 candies)	
Writing or typing letter symbols (e.g., k for key)	
Toilet trained	
Reading written number symbols up to 20	

Q42. How **important** is it for you that your child masters the following competencies **at the end of the primary school**?

Not important at all — Very important

0 1 2 3 4 5 6 7 8 9 10

Recognising dice patterns	
Accurate use of pronouns in sentences	
Learning multiplication tables by heart	
Writing number symbols up to 10	
Eating using knife and fork	
Doing shopping independently	
Buttoning and unbuttoning	
Solving simple puzzles	
Writing his/her own name	

Q43. How **important** is it for you that your child masters the following competencies **at the end of the primary school**?

Not important at all — Very important

0 1 2 3 4 5 6 7 8 9 10

Comparing quantities (e.g., “which one is more?”)	
Using public transports independently	
Reciting the number sequence up to 10	
Given a number up to 20, identifying one more and one less	
Using a toaster	
Comparing the size of objects (e.g., "which one is bigger?")	
Using efficient calculation strategies (e.g., 25 - 8 = 25 - 5 - 3)	
Remembering information about personal data (e.g., age)	
Interpreting and constructing pie charts	

Q44. How **important** is it for you that your child masters the following competencies **at the end of the primary school**?

Not important at all — Very important

0 1 2 3 4 5 6 7 8 9 10

Counting up and down in tens (e.g., 10, 20, 30)	
Extending a pattern or a sequence (e.g., red, blue, green, red, blue, green, etc.)	
Having a best friend	
Recognising and naming common 2-D (e.g., circles) and 3-D shapes (e.g., spheres)	
Following multiple-step instructions to complete a task (e.g., emptying the dishwasher and loading it with dirty dishes)	
Riding a bicycle	
Solving simple subtractions (e.g., 9 - 2)	

Q45. If you have any further comments, please write them in the space provided below.

[short paragraph]

End of Block: SECTION 5. Importance

Start of Block: SECTION 6. Attitudes and beliefs

SECTION SIX

This section of the questionnaire is about **your** attitudes and beliefs.

Q46. What is **your attitude** towards mathematics and reading? Select the option that correctly indicates your answer.
Rate your answer on a scale from "Strongly disagree" to "Strongly agree".

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I like mathematics	○	○	○	○	○
Mathematics is important	○	○	○	○	○
I am competent in mathematics	○	○	○	○	○
I like to read	○	○	○	○	○
Reading is important	○	○	○	○	○
I am competent in reading	○	○	○	○	○

Q47. Select the option that correctly indicates your answer. Rate your answer on a scale from "Strongly disagree" to "Strongly agree". If you feel that the item is not appropriate for the stage of development of your child, select the option "Not appropriate".

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree	Not appropriate
My child is not good at maths	○	○	○	○	○	○
My child is good at calculations	○	○	○	○	○	○
My child enjoys maths	○	○	○	○	○	○
My child doesn't like doing hard maths problems	○	○	○	○	○	○
My child avoids doing his/her maths homework	○	○	○	○	○	○

Q48. If you have any further comments about your attitude and beliefs toward mathematics, please use the space below.

[short paragraph]

End of Block: SECTION 6. Attitudes and beliefs

Start of Block: SECTION 7. VABS-II

SECTION SEVEN

The next section asks about your child's behaviours.

Read each item and mark the response that best describes your child's behaviour:

- **Usually**: if your child usually performs the behaviour independently (that is, without physical help or reminder)
- **Sometimes**: if your child sometimes performs the behaviour independently or partially performs the behaviour independently)
- **Never**: if your child never performs the behaviour or never performs it independently
- **I don't know**: if you have never seen your child performing the behaviour or if have no knowledge about your child's performance on the given behaviour.

When rating your child's behaviour, please consider that some of the behaviours included in this section may be too hard for younger children and some may be too easy for older children.

We know that some children with additional needs use *Makaton* or *other Sign Languages* to communicate, please consider these abilities as their talking skills when completing the following part of the survey.

Likewise, we know that some children with additional needs use assistive technology to write, so when completing the items that ask about your child's writing skills, consider their *typing* skills as their writing abilities.

Q49. For each item, please select the answer the best describes your child's behaviour.

	Usually	Sometimes	Never	I don't know
Points to common objects in a book or magazine as they are named (for example, dog, cup, car, key, etc.)	○	○	○	○
Listens to instructions	○	○	○	○
Follows instructions with one action and one objects (for example, “Bring me the book”; “Close the door”; etc.)	○	○	○	○
Points to at least five minor body parts when asked (for example, fingers, elbows, teeth, toes, etc.)	○	○	○	○
Follows instructions with two actions or an action and two objects (for example, “Bring me the crayon and the paper”; “Sit down and eat your lunch”; etc.)	○	○	○	○
Follows instruction in “if-then” form (for example, “If you want to play outside, then put your things away”; etc.)	○	○	○	○
Listens to a story for at least 15 minutes	○	○	○	○

Listens to a story for at least 30 minutes	○	○	○	○
Follows three-part instructions (for example, “Brush your teeth, get dressed and make your bed”; etc.)	○	○	○	○
Follows instructions or directions heard 5 minutes before	○	○	○	○

Q50. For each item, please select the answer the best describes your child's behaviour.

	Usually	Sometimes	Never	I don't know
Understands saying that are not meant to be taken word for word (for example, “Button your lip”; “Hit the road”; etc.)	○	○	○	○
Listens to an informal talk for at least 15 mins	○	○	○	○
Listens to an informal talk for at least 30 mins	○	○	○	○
Says correct age when asked	○	○	○	○
Says at least 100 recognizable words	○	○	○	○
Uses “in”, “on” or “under” in phrases or sentences (for example, "Ball go under chair"; "Put it on the table"; etc.)	○	○	○	○
Uses “and” in phrases or sentences (for example, "Mom and Dad"; "I want ice cream and cake"; etc.)	○	○	○	○
Says first and last name when asked	○	○	○	○

Identifies and says most common colours (that is red, blue, green, yellow, orange, purple, brown and black) - Mark "Always" if your child names 6 to 8 colours, mark "Sometimes" if your child names 2 to 5 colours, mark "Never" if your child names 0 or 1 colour	○	○	○	○
Asks questions beginning with "who" or "why" (for example, "Who's that?"; "Why do I have to go?"; etc.)	○	○	○	○

Q51. For each item, please select the answer the best describes your child's behaviour.

	Usually	Sometimes	Never	I don't know
Uses present tense verbs ending in -ing (for example, "Is singing"; "Is playing"; etc.)	○	○	○	○
Uses possessives in phrases or sentences (for example, "That's her book"; "This is Carlo's ball"; etc.)	○	○	○	○
Uses pronouns in phrases or sentences; must use correct gender and form of the pronoun, but sentences need not be grammatically correct (for example, "He done it; "They went"; etc.	○	○	○	○
Asks questions beginning with "when" (for example, "When is dinner?"; "When can we go home?"; etc.)	○	○	○	○
Uses regular past tense verbs (for example, walked, baked etc.); may use irregular past tense verb ungrammatically (for example, "I runned away"; etc.)	○	○	○	○

Uses “behind” or “in front of” in phrases or sentences (for example, "I walked in front of her"; "Terrell is behind you"; etc.)	○	○	○	○
Pronounces words clearly without sound substitutions (for example, does not say "wabbit" for "rabbit", "Thally" for "Sally"; etc.)	○	○	○	○
Tells basic parts of a story, fairy tale, or television show plot; does not need to include great detail or recount in perfect order	○	○	○	○
Identifies one or more alphabet letters as letters and distinguishes them from numbers	○	○	○	○
Recognizes own name in printed form	○	○	○	○

Q52. For each item, please select the answer the best describes your child's behaviour.

	Usually	Sometimes	Never	I don't know
Identifies at least 10 printed letters of the alphabet	○	○	○	○
Prints or writes using the correct orientation (for example, in English from left to right; in some other languages from right to left or top to bottom)	○	○	○	○
Copies own first name	○	○	○	○
Identifies all printed letters of the alphabet, upper- and lowercase	○	○	○	○
Prints at least three simple words from example (for example, cat, see, bee, etc.)	○	○	○	○
Prints or writes own first and last name from memory	○	○	○	○
Reads at least 10 words aloud	○	○	○	○
Prints at least 10 simple words from memory (for example, hat, ball, the, etc.)	○	○	○	○
Reads simple stories aloud (that is, stories with sentences of three to five words	○	○	○	○

Prints simple sentences of three or four words; may make small errors in spelling or sentence structure	○	○	○	○

Q53. For each item, please select the answer the best describes your child's behaviour.

	Usually	Sometimes	Never	I don't know
Prints more than 20 words from memory; may make small spelling errors	○	○	○	○
Demonstrates understanding of function of telephone (for example, pretend to talk on phone, etc.)	○	○	○	○
Talks to familiar person on telephone	○	○	○	○
Uses TV or radio without help (for example, turns equipment on, accesses channel or station, selects program, etc.)	○	○	○	○
Counts at least 10 objects one by one	○	○	○	○
Is aware of and demonstrates appropriate behaviour while riding in car (for example, keeps seat belt on, refrain from distracting driver, etc.)	○	○	○	○
Demonstrate understanding of the function of money (for example, says, "Money is what you need to buy things at a store", etc.)	○	○	○	○

Uses sidewalk (where available) or shoulder of road when walking or using wheeled equipment (skates, scooter, tricycle, etc.)	○	○	○	○
Demonstrates understanding of function of clock (for example, says, "Clocks tell time"; "What time can we go?" etc.)	○	○	○	○
Follows household rules (for example, no running in the house, no jumping on the furniture, etc.)	○	○	○	○

Q54. For each item, please select the answer the best describes your child's behaviour.

	Usually	Sometimes	Never	I don't know
Demonstrates computer skills necessary to play games or start programs with computer turned on, does not need to turn computer on by self	○	○	○	○
Summons to the telephone the person receiving a call or indicates that the person is not available	○	○	○	○
Identifies penny, pence and pounds by name when asked; does not need to know the value of coins	○	○	○	○
Looks both ways when crossing streets or roads	○	○	○	○
Says current day of the week when asked	○	○	○	○
Demonstrates understanding of right to personal privacy for self and others (for example, while using restroom or changing clothes etc.)	○	○	○	○
Demonstrates knowledge of what phone number to call in an emergency when asked	○	○	○	○

Tells time using a digital clock or watch	○	○	○	○
States value of penny, pence and pounds	○	○	○	○
Discriminates between bills of different denominations (for example, refers to £1 bills, £5 bills, etc., in conversation; etc.)	○	○	○	○

Q55. For each item, please select the answer the best describes your child's behaviour.

	Usually	Sometimes	Never	I don't know
Obeys traffic lights and “Walk” and “Don't walk” signs	○	○	○	○
Points to current or other date on calendar when asked	○	○	○	○
Demonstrate understanding that some items cost more than others (for example, says, "I have enough money to buy gum but not a candy bar"; "Which pencil costs less?" etc.)	○	○	○	○
Unwraps small objects (for example, gum or candy)	○	○	○	○
Completes simple puzzle of at least two pieces or shapes	○	○	○	○
Turns book or magazine pages one by one	○	○	○	○
Uses twisted hand-wrist motion (for example, winds up toy, screws/unscrews lid of jar; etc.)	○	○	○	○
Holds pencil in proper position (not with fist) for writing or drawing	○	○	○	○

Colours simple shapes; may colours inside lines	○	○	○	○
Builds three-dimensional structures (for example, a house, bridge, vehicle, etc.) with at least five small blocks	○	○	○	○

Q56. For each item, please select the answer the best describes your child's behaviour.

	Usually	Sometimes	Never	I don't know
Opens and closes scissors with one hand	○	○	○	○
Glues or pastes two or more pieces together (for example, for art or science projects; etc.)	○	○	○	○
Uses tape to hold things together (for example, torn page, art project, etc.)	○	○	○	○
Draws more than one recognizable form (for example, person, house, tree, etc.) - Mark "Always" if your child draws two or more recognizable forms; mark "Sometimes" if your child draws one recognizable form; mark "Never" if your child does not draw any recognizable forms	○	○	○	○
Makes recognizable letters or numbers	○	○	○	○
Draws circles freehand while looking at example	○	○	○	○

Uses scissors to cut across paper along a straight line	○	○	○	○
Colours simple shapes; colours inside the lines	○	○	○	○
Cuts out simple shapes (for example, circles, squares, rectangles, etc.)	○	○	○	○
Uses eraser without tearing paper	○	○	○	○

Q57. For each item, please select the answer the best describes your child's behaviour.

	Usually	Sometimes	Never	I don't know
Draws square freehand while looking at example	○	○	○	○
Draws triangle freehand while looking at example	○	○	○	○
Ties knot	○	○	○	○
Draws straight lines using a ruler or straightedge	○	○	○	○
Unlocks dead-bolt, key, or combination locks that require twisting	○	○	○	○
Cuts out complex shapes (for example, stars, animals, alphabet letters, etc.)	○	○	○	○
Uses keyboard or touch screen to type name or short words; may look at keys	○	○	○	○
Ties secure bow	○	○	○	○
Uses a keyboard to type up to 10 lines; may look at keys	○	○	○	○

Q58. If you have any other comments, please state them in the box below.

[short paragraph]

End of Block: SECTION 7. VABS-II

Start of Block: SECTION 8. About you

SECTION EIGHT

This final section will ask some information **about you**.

Q59. What is your gender?

- ○ Male
- ○ Female

Q60. What is your ethnic origin?

- ○ Asian
- ○ Black or African American
- ○ White
- ○ Chinese
- ○ Mixed
- ○ Other, please specify

____

Q61. What country do you live in?

[……]

Q62. What is your relationship to the participating child?

- ○ Mother
- ○ Father
- ○ If different, please specify

____

Q63. What is the first language you speak with your participating child?

[……]

Q64. What is your highest educational qualification?

- ○ No formal qualification
- ○ O-level/GCSEs or equivalent
- ○ A-levels or equivalent
- ○ Graduate degree
- ○ Post-graduate degree
- ○ Vocational training
- ○ Other, please specify

____

Q65. What is the highest mathematical level that you achieved? Please indicate the course grade

- ○ O-level/GCSEs or equivalent [...]
- ○ A levels [...]
- ○ Other, please specify [...]
- ○ Does not apply

Q66. Are you happy to be contacted by the research team for the second phase of this study?

- ○ Yes
- ○ No

Skip To: End of Block If Are you happy to be contacted by the research team for the second phase of this study? = No

Q67. Please provide your email address, so that the research team can contact you.

[......]

End of Block: SECTION 8. About you

Start of Block: Thank you

Thank you for taking the time to fill out the survey today.

Should you have any further questions or concern about the research project or want to withdraw your responses (by quoting your email address), please contact the research team using the email address [redacted].

If there are any questions you feel require a more detailed answer, we would be happy to hear from you. Please use the contact details provided to get in touch:

Principal Investigator
Dr Jo Van Herwegen
[redacted]

Researcher
Erica Ranzato
[redacted]

Once again, we would like to thank you for your valuable contribution to this research.

Your participation is greatly appreciated!

End of Block: Thank you