Perceptual Subitizing and Conceptual Subitizing in Williams syndrome and Down syndrome: Insights from eye movements

Authors: Erica Ranzato; Andrew Tolmie; Jo Van Herwegen

Key words: Williams Syndrome; Down Syndrome; Perceptual Subitizing; Conceptual Subitizing; Eye tracking

Published in Research in Developmental Disabilities
Accepted 17 July 2020, Available online 20 August 2020

Highlights

- This study examined perceptual and conceptual subitizing in individuals with Down syndrome and Williams syndrome, by means of eye tracking.
- No significant differences in accuracy, RT and fixation count suggested that all participants performed perceptual and conceptual subitizing.
- Participants with DS showed significantly shorter fixations in all experimental conditions, when compared to the control group.
- When counting, both participants with DS and participants with WS used inefficient scanning strategies.

Abstract

Background and aims

Mathematical difficulties in individuals with Williams Syndrome (WS) and in individuals with Down Syndrome (DS) are well-established. Perceptual subitizing and conceptual subitizing are domain-specific precursors of mathematical achievement in typically developing (TD) population. This study employed, for the first time, eye-tracking methodology to investigate subitizing abilities in WS and DS.

Methods and procedures

Twenty-five participants with WS and 24 participants with DS were compared to a younger group of TD children (n = 25) matched for mental age. Participants were asked to enumerate one to six dots arranged either in a dice or a random pattern.

Outcomes and Results

Accuracy rates and analyses of reaction time showed no significant differences between the clinical groups (WS and DS) and the control group, suggesting that all participants used the same processes to perform the enumeration task in the different experimental conditions. Analyses of the eye movements showed that both individuals with WS and individuals with DS were using inefficient scanning strategies when counting. Moreover, analyses of the eye movements showed significantly shorter fixation duration in participants with DS compared to the control group in all the experimental conditions.

Conclusions and implications

The current study provides evidence that individuals with WS and individuals with DS perform both perceptual subitizing and conceptual subitizing. Moreover, our results suggest a fixation instability in DS group that does not affect their performance when subitizing but might explain their low accuracy rates when counting. Findings are discussed in relation to previous studies and the impact for intervention programmes to improve counting and symbolic mathematical abilities in these populations.

1. Introduction

1.1. Enumeration abilities

Perceptual subitizing is defined as the ability to enumerate quickly and effortlessly up to five items without having to count (Kaufman, Lord, Reese, & Volkmann, 1949). There is wide consensus about the assumption that this ability is innate and develops before verbal counting (Feigenson, Dehaene, & Spelke, 2004). In particular, the subitizing range increases with age during early childhood from 1 to 3 to 1–5 (P. Starkey & Cooper, 1995). Although the concept of perceptual subitizing has been present in the literature for almost 70 years, there is continued debate on the cognitive mechanisms behind this skill and whether they differ from counting processes or not (see Schleifer & Landler, 2011 for a discussion). A characteristic pattern appears to be that while the time for subitizing up to three or four items increases at most 40–100 ms per item, for larger numbers the cost of any additional item is 250–350 ms (Trick & Pylyshyn, 1993). Therefore, a difference between subitizing and counting can be seen by a sharp increase in the response times (RT) when individuals engage in counting, which is called the point of discontinuity[1] (Reeve, Reynolds, Humberstone, & Butterworth, 2012), and decrease in accuracy.

Yet, people can still accurately and quickly estimate quantities above the subitizing range and without counting when these items are arranged in a familiar and recognizable pattern (e.g. dice patterns or show of fingers). This ability is named conceptual subitizing and involves a pattern recognition system (Jansen et al., 2014; Krajcsi, Szabo, & Morocz, 2013; Piazza, Mechelli, Butterworth, & Price, 2002). For instance, Krajcsi et al. (2013) found that participants' RT to canonical patterns up to six dots were similar to enumerating dots within the perceptual subitizing range (i.e. 1 to 3 dots). Jansen et al. (2014) reported that when children were presented with different dot configurations, they were more accurate in the dice pattern condition than in the random displays. 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 and follows a hypothetical developmental trajectory that originates in the innate perceptual subitizing

---

[1] "Point of discontinuity" is not the correct mathematical term to describe this function, the use of "piecewise function" would have been more appropriate. Nevertheless, we decided to keep using this inaccurate term in this paper because this is how the subitizing / counting function has been consistently described in the literature.

ability (Sarama & Clements, 2009). Despite the wide acceptance of conceptual subitizing within the field of education, empirical exploration of this concept is still limited (Goukon, 2016).

Research in the typically developing (TD) population has shown that perceptual subitizing and conceptual subitizing may serve as a cognitive scaffold for the development of counting and, in general, for the development of arithmetic skills (Benoit, Lehalle, & Jouen, 2004; Kroesbergen, Van Luit, Van Lieshout, Van Loosbroek, & Van de Rijt, 2009; Ozdem & Olkun, 2019). In alignment with such views, researchers have found that children's perceptual subitizing strongly predicts their mathematics abilities in the early years (Reigosa-Crespo et al., 2013). Moreover, perceptual subitizing has been found to be a core deficit in individuals with Developmental Dyscalculia (DD). The prediction that individuals with DD compensate for an impaired subitizing mechanism by applying the sequential counting process has been supported by the increasing evidence that individuals with dyscalculia show steeper RT slopes in the subitizing range compared to typically developing peers, indicating that such individuals tend to adopt serial counting to determine the numerosity of small sets (Landerl, Bevan, & Butterworth, 2004; Moeller, Neuburger, Kaufmann, Landerl, & Nuerk, 2009; Schleifer & Landerl, 2011). Furthermore, Ashkenazi, Mark-Zigdon, and Henik (2013) found that children diagnosed with DD failed to benefit from the canonical arrangement of dots up to nine, as they showed no enumeration speed difference between random and canonical pattern stimuli. This suggests that not only perceptual subitizing but also conceptual subitizing may be a challenge for this population and that this may impact the development of their mathematical abilities.

1.2. Enumeration abilities in Williams syndrome and Down syndrome

Williams syndrome (WS) and Down syndrome (DS) are two genetic developmental disorders that have similar cognitive impairments within the mild-to moderate learning difficulties range and great mathematical difficulties (Brigstocke, Hulme, & Nye, 2008; Van Herwegen & Simms, 2020). However, they have different uneven cognitive profiles with those with WS having better language and short-term memory abilities compared to their visuo-spatial difficulties (Martens, Wilson, & Reutens, 2008), and those with DS showing poorer language abilities and short-term memory abilities compared to their visuospatial skills (Silverman, 2007). There are very few studies that have specifically investigated perceptual subitizing in individuals with WS or DS. To the best of our knowledge, there are no studies investigating conceptual subitizing in these populations. In their systematic review of mathematical abilities in WS, Van Herwegen and Simms (2020) reported that 27 studies show that mathematical abilities are delayed in this population and likely to follow an atypical developmental pathway. As for subitizing abilities, only two studies assessed perceptual subitizing both in children and in adults (Ansari, Donlan, & Karmiloff-Smith, 2007; O'Hearn, Hoffman, & Landau, 2011). Both of these studies reported a high level of accuracy in the subitizing range and lower accuracy starting at four items. Moreover, the analyses of RT reported by O'Hearn et al. (2011) replicated what is usually observed in the typical population, that is a relatively flat RT function for the subitizing range (reduced to numerosities 1 to 3) and greater slopes for higher numerosities. Analyses of the RT slopes showed no differences between the WS group and the TD control groups in the subitizing range.

Only one study investigated enumeration skills in children with DS (Sella, Lanfranchi, & Zorzi, 2013). In this study participants were presented with a sample image on a screen – that could be either a set of dots or an Arabic digit from 1 to 9 – and with a target image – a set of dots that could either match the number of dots or the digit in the sample image or differ for one dot. Participants with DS were asked to compare the two stimuli. Note that the experimental paradigm used by Sella et al. (2013) differs from the standard dot enumeration paradigm in which participants typically explicitly state the number of dots presented. The performance of children with DS showed a pattern that is consistent with the use of a serial counting process even in the subitizing range, as their RT increased systematically with the number of dots presented, without presenting any point of discontinuity.

In summary, past evidence indicates that perceptual subitizing in individuals with WS shows similarities to TD population with a subitizing range reduced to numerosity 1-3 not only in children but also in adults, whereas individuals with DS seem to use counting processes by default even for low numerosities, as observed in individuals with DD.

1.3. The use of eye tracking to investigate enumeration abilities

The use of eye tracking to investigate numerical cognition has important methodological implications. First, eye tracking allows for a number of informative and sensitive measures, beyond accuracy and RT. Second, there is strong evidence that eye movements can provide insight into how individuals process numerical information

(Sullivan, Juhasz, Slattery, & Barth, 2011). Third, Van Herwegen and Karmiloff-Smith (2015) argued that general basic-level abilities, such as how people move their eyes across the display, may impact the development of number abilities or lead to different developmental pathways for mathematical abilities. In other words, similar performance in terms of accuracy and RT may be driven by different cognitive skills. Therefore, the examination of eye movements not only allows for the collection of more informative measurements, but also for a better insight into the cognitive mechanisms behind enumeration skills.

The recent use of eye-tracking measures to study enumeration abilities has allowed the investigation into how eye movements relate to the pattern described for the RT during an enumeration task. Findings in typically developing children and typically developing adults (Schleifer & Landerl, 2011; Watson, Maylor, & Bruce, 2007) showed that participants use few or no eye movements for enumerating 1 to 4 objects and report a monotonic increase in the number of saccades and fixations for enumerating 5 or more objects. On the other hand, in these populations, average fixation duration has not been found to vary systematically with the number of the items displayed (Schleifer & Landerl, 2011). To the best of our knowledge, the paper by Moeller et al. (2009) is the first and only study that has employed eye-tracking methodology to investigate enumeration abilities in 2 children with DD, compared to a control group of 8 TD children matched for chronological age. Their findings show that, when enumerating 1 to 3 dots, the RT and the fixation count slopes of both children with DD were larger than those of the control group, showing that children with DD were impaired in subitizing and that they were possibly counting at least on some proportion of the trials. However, DD participants did not differ from the control group in terms of average fixation duration, showing that they did not present impairments in the access to numerical magnitude representation when subitizing.

Very few studies have investigated numerical cognition in neurodevelopmental disorders by means of the eye-tracking methodology (e.g. Van Herwegen, Ranzato, Karmiloff-Smith, & Simms, 2019). Eye-tracking studies that investigated the scanning behaviours of individuals with WS and DS have reported that these populations show atypical but different eye movement patterns. Studies that have investigated eye movements in WS have shown that these are impaired from infancy onwards (Van Herwegen, 2015). In particular, it has been reported that individuals with WS struggle to plan or execute eye movements and thus to disengage from a previously fixed target (Brown et al., 2003). This impairment results in longer, or "sticky", fixations, and this could impact the development of domain specific abilities, in that the inability to plan eye movements might affect their fixation duration and thus the visual processing of numerical information (Van Herwegen, 2015). As for the DS population, research by Viñuela-Navarro, Erichsen, Little, Saunders, and Woodhouse (2017) on scanning strategies in children with DS has suggested a fixation stability deficit, with significantly shorter fixation durations in participants with DS compared to a control group of TD children. The same patterns have been reported in the early stage of development of these clinical populations in a study by Brown et al. (2003) that found that toddlers with WS displayed evidence of deficits in saccade planning while toddlers with DS presented shorter and fewer periods of sustained attention. Given the different looking behaviours, a comparison study will allow the investigation of whether and how domain-general attention abilities, such as eye movements, affect enumeration abilities and will give further insight into the mechanisms behind these skills.

1.4. The current study

The current study is the first to use eye movements to examine enumeration abilities in children and adults with WS and DS in comparison to a group of typically developing children matched for mental age. In addition, the current study not only focuses on perceptual subitizing, but also examines for the first time conceptual subitizing by using canonical and random displays of 1 to 6 dots. It also examines the development of enumeration skills within these clinical populations using a cross-sectional sample, as well as considering when participants default to counting the items instead of subitizing.

Based on the existing literature, it was predicted that:

1. WS participants would use different enumeration processes to perform the task depending on the number of dots presented – i.e. perceptual subitizing for numerosity 1 to 3 and counting for numerosity 4 to 6 – while DS participants would not use subitizing, but would rather use counting to perform the task, regardless of the number of dots presented.
2. Overall, during the enumeration task, there would be differences in the mean fixation durations of both clinical groups, with WS presenting longer fixation durations and DS presenting shorter fixation durations than the control group.
3. Enumeration abilities of both clinical groups would be aligned with their mental age rather than their chronological age.

Due to the lack of literature on conceptual subitizing, no predictions were made on the outcomes regarding this ability for the WS and DS groups.

2. Methods

2.1 Participants

Seventy-eight participants took part in the study.

Twenty-six participants with WS aged 8;00 to 51;08 (16 females) were recruited through the Williams Syndrome Foundation UK across the UK. All had a genetic fluorescent in situ hybridisation (FISH) test confirming the genetic deletion implicated in WS, in addition to a clinical diagnosis for WS.

Twenty-four participants with DS aged 8;08 to 49;02 (12 females) were recruited via Down syndrome support groups across the South-East of the UK. They all had a genetic diagnosis for Down syndrome.

Twenty-eight TD children (17 females) were recruited via local schools and online recruitment through social media. One participant was excluded because they did not complete the entire assessment. Of the twenty-seven TD children that completed the whole assessment, only those whose scores fell within the range of the Raven's Coloured Progressive Matrices (RCPM) scores of the two neurodevelopmental groups (min = 4, max = 25) were included in the control group (n = 25) (see Van Herwegen et al., 2019 for a similar approach). Therefore, the included TD children were much younger than the WS and DS groups (aged between 3;11 and 6;07) but had similar non-verbal intelligence scores as measured by the RCPM. A mental age matched group was used as studies have shown that participants with DS and WS rarely perform at a level for their chronological age (see Van Herwegen et al., 2019) and the current study examined performance strategies during the enumeration task rather than performance level itself.

All of the participants had English as a first language and none of the TD participants had a diagnosis for a learning difficulty. In addition, all participants came from white middle class family backgrounds. See Table 1 for full descriptive characteristics of the three groups.

Table 1

Mean (M) and standard deviation (SD) for the three groups for Chronological Age (CA) and raw scores on the Raven's Coloured Progressive Matrices Test (RCPM)

Group	N (F)	CA			RCPM	
	Count	*M*	*SD*	*Range*	*M*	*SD*
TD	25 (16)	5;02	0;08	3;11 – 6;07	15.48	4.87
WS	26 (16)	19;06	13;07	8;00 – 51;08	15.69	4.44
DS	24 (12)	21;07	11;00	8;08 – 49;02	15.42	6.32

2.2 Background measures

Overall intelligence and reasoning. The Raven's Coloured Progressive Matrices (RCPM; Raven, 2008) was used to assess the participants' overall non-verbal intelligence. A total of 36 items in 3 sets (A, Ab, B), with 12 items per set were presented to each participant. Each item contained a picture of an abstract figure with a missing part. For each item the participant had to choose the correct part that completed the abstract figure from 6 options. Participants received feedback only for the first two items for both correct and incorrect answers provided (e.g. "yes, that is correct. That is the right shape and the right pattern"). Each item was presented for an unlimited amount of time until the participant reported their response. If the participant didn't know the answer or struggled with a particular item they were allowed to guess. A score of 1 was given for every trial performed correctly. The minimum score was 0 and the maximum score was 36.

Numerical and arithmetical knowledge. The Numerical Operations sub-test from the Wechsler Individual Achievement Test Second UK Edition (WIAT-II; Wechsler, 2005) was administered to assess basic numerical knowledge and arithmetical knowledge. Numerical Operations assessed the ability to identify and recognize numbers, count using 1:1 correspondence and solve written calculation problems and simple equations involving the four operations. All participants started from item 1 and a score of 1 point was given for each correct response. The assessment was terminated after 6 consecutive failed items. The minimum raw score was 0 and the maximum was 54. All participants were able to correctly discriminate and recognize digits (items 1 to 5), to count (items 6 and 7) and most of them were able to solve one-digits additions and subtractions (items 8 to 12). Fewer participants were able to solve two-digits additions and subtractions, multiplications, and divisions (items 13 and above). See Table 2 for full descriptive of the scores of the three groups.

Table 2

Mean (M), Standard Deviation (SD), Median, Minimum (Min) and Maximum (Max) for the three groups for raw scores on Numerical Operation subtest WIAT-II

	N	*M*	*SD*	*Median*	*Min*	*Max*
TD	25	7.60	2.02	8.00	3	11
WS	26	8.85	2.99	8.50	3	17
DS	24	7.67	2.82	7.00	4	16

Number name knowledge: Participants were also asked to count from 1 to 20. A score of 20 was given for a correct performance. If the participant said they did not know, made a mistake, or stopped counting, a score equal to the last correct number reported by the participant was given. This task was included to assess that participants could accurately count set sizes of 1 to 6. Table 3 shows that all participants but one individual with DS correctly counted up to 6 and that most of the participants were able to count up to 10. Approximately one in three participants with DS failed to use “teen” words.

Table 3

Percentage of participants that correctly counted up to 6, 10 and 20 by group

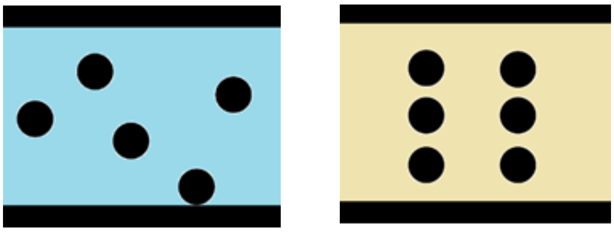
	N	*Score ≥ 6*	*Score ≥ 10*	*Score = 20*
TD	25	100%	96.0%	88.0%
WS	26	100%	100%	92.3%
DS	24	95.8%	95.8%	62.5%

2.3. Experimental measure

Enumeration Task. Subitizing abilities were assessed on a response-terminated computer based test. The participant was seated in an adjustable chair, 60 cm away from the monitor screen. Stimuli contained between one and six black dots of the same size that were arranged either in a canonical (dice condition, i.e. the dots were centred in the middle of the screen) or non-canonical (random condition, i.e. the dots were positioned in random locations on the screen) pattern for a total of 12 trials (6 per condition) – see Figure 1. Thus, stimuli were presented in four combinations of numerosity and arrangement: subitizing range in dice arrangement (i.e. D1-D3), subitizing range in random arrangement (i.e. R1-R3), counting range in dice arrangement (i.e. D4-D6) and counting range in dice arrangement (i.e. R4-R6). These 12 stimuli were arranged in 2 predefined orders, both alternating between dice condition and random condition (list A and list B). Each participant was assigned a list at random. Stimuli were presented for an unlimited amount of time on a 17" monitor. Participants were told that they would see a number of black dots on the computer screen and that they had to say how many dots they saw as quickly and accurately as possible. Before the beginning of each trial, a fixation cross was displayed in the centre of the screen in order to capture the participant’s attention. The experimenter initiated each trial when the participant appeared to be attentive and looked at the fixation cross. Each display was presented until the participant reported how many dots there were, and the response was recorded by the experimenter. The participant could report their response verbally or by using signing. Participants did not receive any feedback for their response. Eye movements were recorded using a Tobii T120 screen based eye tracker. Eye movement recordings were controlled with Tobii’s Studio software (version 2.06) at 120 Hz and fixations were defined using the Tobii I-VT fixation filter. The task started with a five-point calibration that was subject-paced.

Figure 1

Sample stimuli for set sizes 5 and 6 in both arrangements. Stimulus R5 (5 dots in random pattern) and stimulus D6 (6 dots in dice pattern). AoI is in colour. Note that the two black bands on the top and bottom of the screen were not visible to the participant.



2.4. Procedure

Adult participants and parents of all children provided written informed consent prior to their participation. Children provided verbal assent. Participants completed the tasks in a quiet room at the university. The session started with assessment of mental age, followed by numerical and arithmetical knowledge tasks and the enumeration task. The entire session lasted approximately 50 min and breaks were taken between different assessments when needed. This study was assessed by the Kingston University’s Ethics Committee and was fully approved.

2.5. Data analysis process

In order to extract and calculate RTs and eye movements using the Tobii's Studio software, the Area of Interests (AoI) had to be set. One AoI was set to cover the entire screen, for each stimulus (see Figure 1). . Response time (RT) and eye movements were recorded for each participant. RT was measured as the duration of all visits within the AoI. Fixation count was measured as the number of times the participant fixated the AoI and median fixation duration was defined as the median of the duration of each individual fixation within the AoI. Median fixation durations were used instead of mean fixation durations scores, as they are less strongly influenced by outliers (see Schleifer, 2011).

Participants for whom the eye-tracking 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[2]. This is because an excessive number of missing responses may denote poor attention, which could undermine the validity and reliability of the administered task (Sella et al., 2013). Moreover, trials were discarded from analyses if a response was made in less than 0.2 s or more than 10 s from trial onset (n = 6 trials, 2 trials in each of the groups were excluded), as it was deemed such RT were too short or beyond a fixed threshold denoting poor attention (see Paul, Reeve, & Forte, 2017 for a similar approach). Finally, statistical analyses were conducted only on those trials for which participants produced a correct response (Table 4).

To increase the power of statistical analyses, analyses were run on a 2 (dice pattern vs random pattern) x 2 (subitizing range vs counting range) design rather than on a 2 (pattern) x 6 (number of dots) design. In previous studies the subitizing range has been variably defined, sometimes ranging to three, sometimes to four. As the current study involved young children and participants with neurodevelopmental disorders who might have a restricted subitizing range, we decided to be conservative and define the subitizing range for numerosities one to three and the counting range for numerosities four to six. This led to the definition of four experimental conditions: dice pattern in subitizing range (1 - 3 dots displayed in dice pattern), dice pattern in counting range (4 - 6 dots displayed in dice pattern), random pattern in subitizing range (1 - 3 dots displayed in random pattern), and random pattern in counting range (4 - 6 dots displayed in random pattern). Nonparametric analyses were conducted,

---

[2] This is not unusual when assessing individuals with WS, and it is due to their strabismus rather than to experimental errors (Kapp, Von Noorden, & Jenkins, 1995).

because of violation of the normality assumption. Welch ANOVAs were run in case of violation of the assumption of equality of variance.

In order to examine mean accuracy rates between groups and within the same group in different experimental conditions, we ran a series of Kruskal-Wallis and Wilcoxon Signed Ranks Tests. Then, we ran Spearman's correlations between RT and the two eye movement measures to determine whether there was an association between these variables.

In order to examine whether participants were using different enumeration processes in the different experimental conditions, we conducted an analysis focused on the data slopes for RT and fixation count (Schleifer et al., 2011). First, regression lines for each experimental condition were individually computed for the RT and fixation count data respectively. Then, the average slope score for each group was computed and submitted to four separate one-way Welch ANOVAs to determine if the three groups were using different enumeration processes in the same experimental conditions. Finally, a Wilcoxon Signed Ranks Test was conducted on the averaged slopes for RT and fixation count to determine whether the pattern and/or the enumeration range had an effect on the enumeration process used by the participants.

In order to examine differences in the gazing behaviour between the groups, a one-way Welch ANOVA was conducted on the median fixation duration for the four experimental conditions. Finally, correlations were run to determine the relationship between accuracy in the experimental conditions, chronological age, mental age and numerical and arithmetical knowledge of the participants.

3. Results

3. 1. Background measures

There were no significant differences in age between the WS group and the DS group $t(47) = -1.017$, $p = .981$, but the TD group was significantly younger than DS group ($t(47) = -7.434$, $p < .0005$) and WS group ($t(48) = -5.419$, $p < .0005$). There were no significant difference between the TD, WS, and DS groups for RCPM scores; $F(2,46.002) = .001$, $p = .999$ or for numerical and arithmetical knowledge $F(2,45.429) = 1.627$, $p = .208$. There was a significant difference between the three groups for number name knowledge $F(2,36.518) = 3.666$, $p = .035$. Post hoc analysis showed that DS group had a statistically significant lower score than the WS group (-2.685, 95% CI [-5.27, -.10], $p = .040$), but not other differences were statistically significant.

3.2. Enumeration accuracy rates

The percent accuracy rate for each numerosity and display condition is shown in Table 4.

Table 4

Percent accuracy rate for each display condition for each group

	Dice Pattern			Random Pattern		
Dots presented	TD (n= 25)	WS (n= 25)	DS (n =24)	TD (n= 25)	WS (n= 25)	DS (n= 24)
1	100%	100%	100%	100%	100%	100%
2	100%	96.0%	100%	100%	100%	100%
3	96.0%	100%	95.8%	96.0%	100%	91.7%
4	88.0%	96.0%	83.3%	96.0%	92.0%	87.5%
5	92.0%	88.0%	87.5%	80.0%	80.0%	79.2%
6	84.0%	88.0%	87.5%	70.8%	76.0%	58.3%

A Kruskal-Wallis 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. Median accuracy rates were not statistically significantly different between the three groups in all the experimental conditions (D1-D3: $\chi2(2) = .587$, $p = .745$; D4-D6: $\chi2(2) = 2.157$, $p = .340$; R1-R3: $\chi2(2) = .580$, $p = .784$; R4-R6: $\chi2(2) = .846$, $p = .655$).

A series of Wilcoxon Signed Ranks tests was used to determine whether there was a statistically significant mean difference for each group between the accuracy rate when participants had to enumerate dots in the subitizing range compared to the counting range, separately for the two pattern conditions. Accuracy rates decreased for all groups as the numerosity increased. The difference in the accuracy rates between low and high numerosity was always found to be statistically significant, except for the dice pattern in the WS group (Table 5).

Table 5

Wilcoxon Signed Ranks test on the accuracy rate for the subitizing and the counting range over the two pattern conditions for each group

	Dice Subitizing vs Dice Counting	Random Subitizing Vs Random Counting
TD (n= 25)	$z = -2.271, p = .023$ (*)	$z = -3.125, p = .002$ (*)
WS (n= 25)	$z = -1.857, p = .063$	$z = -2.565, p = .010$ (*)
DS (n= 24)	$z = -2.264, p = .024$ (*)	$z = -2.801, p = .005$ (*)

Note: * $p < 0.05$.

A series of Wilcoxon Signed Ranks tests was used to analyse the differences between accuracy rate in the random and in the dice conditions, separately for the two numerosity ranges. When a difference in the accuracy score was apparent, the accuracy was always lower in the random condition than in the respective dice condition, but the difference was never significant, except for the DS group, for whom there was a statistically significant lower rate of accuracy when estimating 4 to 6 dots in the random condition compared to the accuracy score when estimating 4 to 6 dots arranged in the dice pattern (Table 6).

Table 6

Wilcoxon Signed Ranks test on the accuracy rate for the random and the dice conditions over the two numerosity ranges for each group

	Dice Subitizing vs Random Subitizing	Dice Counting vs Random Counting
TD (n= 25)	$z = .000, p = 1.000$	$z = -.975, p = .329$
WS (n= 25)	$z = 1.000, p = .317$	$z = -1.035, p = .301$
DS (n= 24)	$z = -\ .577, p = .564$	$z = -2.333, p = .020$ (*)

Note: * $p < 0.05$.

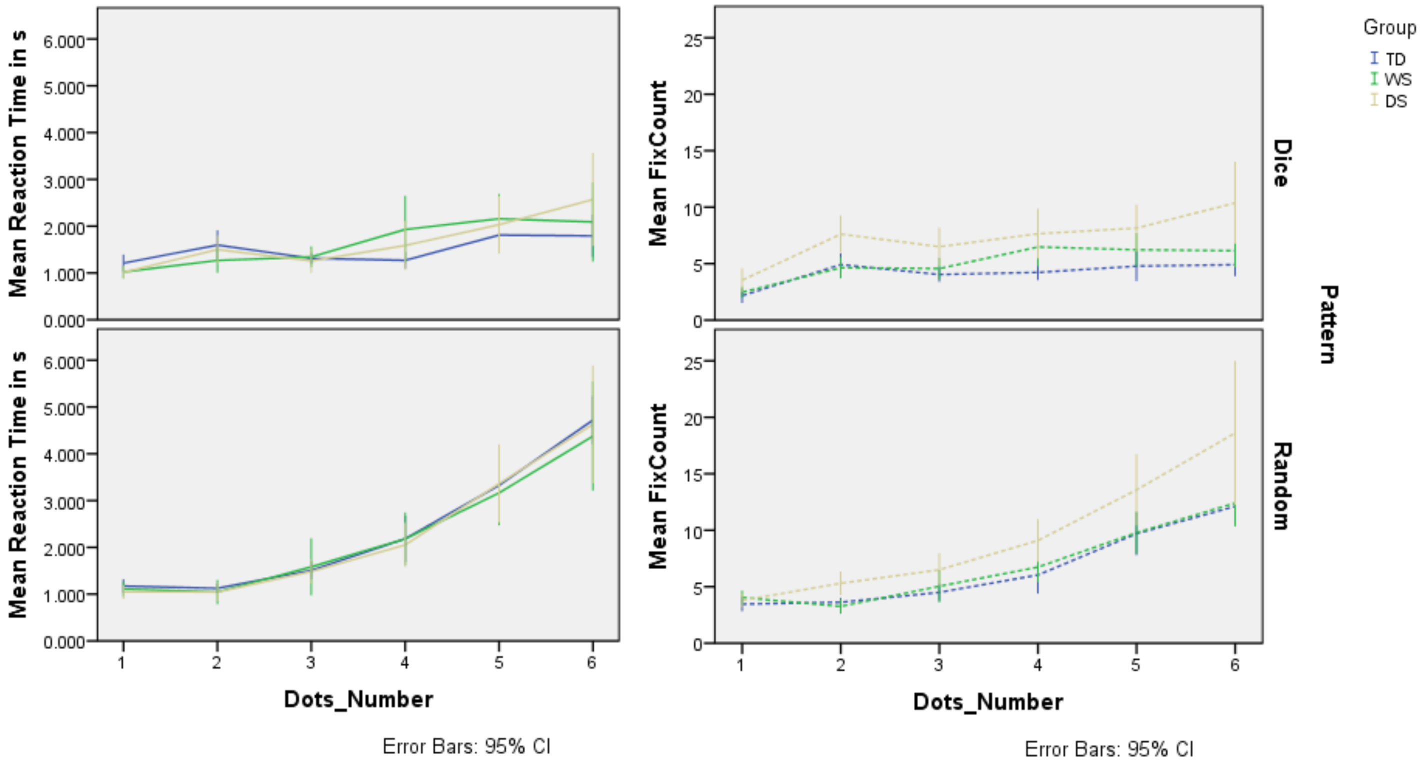
3. 3. Enumeration processes

The Spearman's correlation between RT and median fixation duration showed no statistically significant association, except for the experimental condition R2 (*rs*(72) = .302, *p* = .018). The Spearman's correlation between RT and fixation count for all experimental conditions except D1 showed a statistically significant, positive association (averaged *rs* over experimental conditions was >.55). Because of the tight coupling between RT and fixation count, statistical analyses on these measures are reported together.

Figure 2 shows mean RT (solid lines) and mean fixation count (dashed lines) for each group for all experimental conditions. Visual analysis of the graphs shows a somewhat flat function for both the dice and random conditions for the subitizing range for all groups. In contrast, there is a steeper increase for the counting range but only in the random condition, again in all groups. Importantly, this same pattern is observed for RT and fixation count. The increase observed in RT and fixation count with the number of items suggests that the participants were perceptually subitizing up to three dots and conceptually subitizing four to six dots in the dice condition. For the random patterns they were engaged in counting when enumerating 4 to 6 dots.

Figure 2

Mean RT (solid lines) and mean fixation count (dashed lines) for the three groups for each display condition



Statistical analysis of the slopes for RT and fixation count confirmed the results of the visual analysis of the graphs. Table 7 shows the averaged slopes and intercepts of individual regressions lines for RT and fixation count for each group. RT and fixation count slopes were submitted to separate one-way Welch ANOVAs that showed no statistically significant differences between the three groups in any of the experimental conditions (RT: F values between .30 and 2.13, all $p > .13$; fixation count: F values between .22 and 1.96, all $p > .16$). This suggests that the same enumeration processes were employed by all the participants in each experimental condition, regardless of their group. Moreover, the Wilcoxon Signed Ranks test on the averaged slopes for RT (Table 8) showed that there was a statistically significant increase in the slope when participants were enumerating 4 to 6 dots in the random condition. This confirms that when participants were shown 4 to 6 randomly displayed dots, they were using a different enumeration process than in the other experimental conditions. In other words, participants were subitizing in all the conditions, but they were counting when 4 to 6 randomly displayed dots were shown. As such, the typically observed discontinuity between subitizing and counting was confirmed in our experiment by the RT data observed in all the three groups. This was not the case for the fixation count slopes, as a significant difference was found for an experimental condition where we expected to observe the same enumeration process – D1-3 vs D4-6 for DS participants ($z = -2.112$, $p = .035$), and D1-D3 vs R1-R3 for WS participants ($z = -2.521$, $p = .012$). In line with the expectation, and with the RT data, a statistically significant increase in the fixation count slope was observed when participants were enumerating 4 to 6 dots in the random condition.

Table 8

Wilcoxon Signed Ranks test on mean RT slopes (top row) and mean fixation count slopes (bottom row) for all the experimental conditions

	Dice Counting	Random Subitizing	Random Counting
Dice Subitizing	$z = .250$, $p = .802$	$z = 1.725$, $p = .085$	$z = 5.558$, $p < .0005$ (*)
	$z = -1.802$, $p = .001$ (*DS)	$z = -2.797$, $p =.005$ (*WS)	$z = 4.018$, $p < .0005$ (*)
Dice Counting		$z = -.044$, $p = .965$	$z = 5.304$, $p < .0005$ (*)
		$z = 1.522$, $p = .128$	$z = 5.408$, $p < .0005$ (*)
Random Subitizing			$z = 5.858$, $p < .0005$ (*)
			$z = 4.733$, $p < .0005$ (*)

Note: *: $p < 0.05$ for all groups, *DS: $p < 0.05$ in the DS group only, *WS: $p < 0.05$ in the WS group only.

3.4. Overall gazing behaviour

A one-way Welch ANOVA was conducted to determine if there were any differences in the median fixation duration between the three groups. The analysis showed that median fixation duration was significantly higher for the TD group (M = 0.37, SD = 0.24) when compared to both the clinical groups (WS: M = .28, SD = 0.18; DS: M = .23, SD = 0.19); Welch's *F*(2, 506.822) = 27.902, $p < .001$. Games-Howell post hoc analysis revealed that the median fixation duration of the TD group was significantly higher than the WS group (.08, 95% CI [.04, .13], $p < .001$) as well as the DS group (.14, 95% CI [.10, .19], $p < .001$). Furthermore, median fixation duration of the WS group was significantly higher than the DS group (.06, 95% CI [.02, .09], $p = .001$). Then, separate one-way Welch ANOVAs for each experimental condition were run. Games-Howell post hoc analysis revealed that the DS group showed significantly shorter fixation duration than the TD group in all the experimental conditions, while the WS showed significantly shorter fixation duration only in the D1-D3 condition and in the R4-R6 condition (Table 9).

Table 9

Games-Howell post-hoc comparisons on median fixation duration between TD group and WS group and DS group

	TD (n= 25) vs WS (n= 25)	TD (n= 25) vs DS (n= 24)
Dice Subitizing	.14, 95% CI [.03, 2.49], p = .011 (*)	.18, 95% CI [.06, .30], p = .001 (*)
Random Subitizing	.04, 95% CI [-.03, .10], p = .335	.11, 95% CI [.05, .17], p < .0005 (**)
Dice Counting	.09, 95% CI [-.02, .20], p = .155	.17, 95% CI [.07, .28], p = .001 (*)
Random Counting	.07, 95% CI [.04, .10], p < .0005 (**)	.12, 95% CI [.08, .15], p < .0005 (**)

Note: * $p < 0.05$, ** $p < 0.01$.

A Spearman's correlation analyses was run in order to highlight the relation between mean median fixation duration (i.e. the averaged value of the median fixation duration over the experimental trials) and chronological age and mental age for each group (Table 10). No statistically significant correlations were found.

Table 10

Spearman's correlation between mean median fixation duration in the enumeration task and Chronological Age (CA) and Raven's Coloured Progressive Matrices Test score (RCPM) separately computed for the three groups

Measures	TD (n= 25)	WS (n= 25)	DS (n= 24)
CA	.077	-.079	.225
RCPM	.016	.054	-.119

Note: * $p < 0.05$, ** $p < 0.01$.

3.5. Correlations

We ran Spearman's correlation analyses in order to highlight the relation between the performance in each experimental condition and chronological age, mental age and numerical competence for each group (Table 11). We found a moderate to strong positive correlation between accuracy in D4-D6 and numerical competence in all groups (TD: *rs*(25) = .614, *p* = .001; WS: *rs*(25) = .555, *p* = .004; DS *rs*(24) = .418, *p* = .042). We found a strong positive correlation between D4-D6 accuracy and chronological age for the WS group (*rs*(25) = .608, *p* = .001). Moreover, we found a moderate positive correlation between the D4-D6 accuracy and the D1-D3 accuracy for both the TD group (*rs*(25) = .455, *p* = .022) and the WS group (*rs*(25) = .489, *p* = .013). This correlation was not significant for the DS group (*rs*(24) = .201, *p* = .347).

We found, in all groups, a moderate to strong positive correlation between R4-R6 accuracy and both numerical competence (TD: *rs*(25) = .414, *p* = .040; WS: *rs*(25) = .403, *p* = .046; DS *rs*(24) = .553, *p* = .005) and mental age (TD: *rs*(25) = .439, *p* = .028; WS: *rs*(25) = .452, *p* = .023; DS *rs*(24) = .484, *p* = .017). The correlation between R4-R6 accuracy and chronological age was moderate positive for both the TD group (*rs*(25) = .464, *p* = .019) and the DS group (*rs*(24) = .469, *p* = .021). This correlation was not significant for the WS group (*rs*(25) = .348, *p* = .088). Moreover, only in the DS group, accuracy in counting was correlated with accuracy in conceptual subitizing (DS: *rs*(24) = .686, *p* < .0005; WS: *rs*(25) = .231, *p* = .266; TD *rs*(25) = .280, *p* = .175).

4. Discussion

In this study we employed eye-tracking methodology to investigate enumeration processes in children and adults with WS and DS, compared to a TD group matched for mental age. Participants were asked to enumerate visual sets with 1 to 6 dots arranged either in a dice pattern or a random pattern. This task allowed us to evaluate the enumeration process used by each participant (i.e. perceptual subitizing, conceptual subitizing and counting) in different experimental conditions.

Our first aim was to establish whether individuals with WS and DS were using the same enumeration processes as the control group. In particular, we were interested in investigating whether participants used perceptual subitizing and conceptual subitizing. Our second aim was to investigate the participants' eye movements when performing the task, in order to assess whether their scanning behaviour affected their subitizing and counting abilities, as previous studies argued that such general basic-level abilities could impact the development of domain-specific skills, such as mathematical abilities (Van Herwegen & Karmiloff-Smith, 2015).

Accuracy rates (Table 4) showed that performance was influenced by the spatial arrangement of the dots, for all groups. In line with previous studies (Jansen et al., 2014), all participants were more accurate when enumerating 4-6 dots arranged in the dice pattern than in the random pattern. This is the first study that investigates conceptual subitizing in individuals with WS and showed that despite their visuo-spatial difficulties, individuals with WS benefit from pattern like presentations when needing to enumerate objects. Gordon, Smith-Spark, Newton, and Henry (2020) recently investigated the relationship between working memory and high-level cognition in TD children aged 7 and 8 years old and found that children were faster and more accurate on the counting span compared to the other complex span tasks (i.e. Listening span, and Odd One Out span). Their interpretation of this unexpected finding was that children were subitizing the dots presented on the screen, rather than counting, and that this reduced the cognitive load of the task processing component compared to other complex span tasks. Further research should investigate whether subitizing processes would also reduce cognitive load compared to counting processes in WS and whether presenting individuals with WS with subitizing presentations would benefit their symbolic mathematical learning and counting.

In line with previous studies on TD population (Reeve et al., 2012), the accuracy rates for all groups decreased as the number to enumerate increased. A statistically significant difference in the accuracy rates between enumerating 1-3 dots and 4-6 dots was found for all groups, with the exception of the WS group in the dice condition (Table 5). This was due to the surprisingly high accuracy rate of the WS group in conceptual subitizing. A plausible explanation for this might be found in the statistically significant correlation between accuracy rate in conceptual subitizing and chronological age, suggesting that, when performing conceptual subitizing, individuals with WS – unlike individuals with DS –

compensate their visuospatial difficulties with experience. The absence of such significant correlation in the TD group might be explained by the fact that participants in the control group were much younger and had less experience with subitizing.

Accuracy scores also showed that there were no statistically significant differences between the three groups in any of the conditions, despite the fact that those with DS had lower performance on the number name knowledge task. When looking at the counting range condition, this unexpected finding might be explained by the relatively low upper bound of the range assessed (i.e. 6 dots) and by the fact that almost all participants with DS were able to correctly name numbers up to 6. When evaluating accuracy within the same group, only individuals with DS presented a significantly different, lower level of accuracy when enumerating 4 to 6 dots in the random condition compared to their level of accuracy in the dice condition (Table 6). This is consistent with previous literature reporting difficulties with counting in the DS population (Nye, Fluck, & Buckley, 2001).

In order to examine whether individuals with WS and individuals with DS were performing either subitizing or counting in different experimental conditions, we examined RT and eye movements. In line with previous literature in TD and DD populations (Moeller et al., 2009; Schleifer & Landerl, 2011), we found a weak correlation between median fixations duration and RT, showing that, on average, neither the number nor the spatial arrangement of the dots presented on the screen influenced visual processing time. On the other hand, we found that for all groups the different enumeration processes were well characterized, not only by the RT, but also by the number of fixations applied when scanning the display, with a close-to-constant number of fixations and RT both for numerosities up to three and for the dice pattern condition, and a monotonic increase in the number of fixations and RT for higher numerosities (i.e. 4-6), in the random condition only (Figure 2). Hence, a close correspondence between RT and fixation count was found for all groups, meaning that enumeration processes are well described by both of these measures.

The analysis of RT slopes and fixation count slopes showed that slopes were significantly larger when counting compared to when subitizing for all groups. Moreover, in line with findings from individuals with DD reported by Moeller et al. (2009), RT and fixation count slopes were larger among individuals with WS and individuals with DS compared to the control group. Generally, when comparing the fixation count slopes with the RT slopes for each experimental condition (Table 8), we found a tight coupling. However, we found some discrepancies in that eye movements showed different patterns than the ones showed by the RT in two specific cases. In particular, 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. 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. Our findings highlight that the dice pattern seems to influence the fixation count but not the RT, at least in the clinical population. Further research is needed, but these findings support the methodological choice of combining RT and eye-movement behaviour discussed by Schleifer and Landerl (2011) to get a deeper understanding of the processes underlying enumeration skills, as in the current study eye-tracking measures were more sensitive than RT and highlighted qualitative differences in how individuals approach and perform a dot-counting task.

Moreover, our results on RT and fixation count slopes confirmed the characteristic pattern of an almost flat function with a point of discontinuity where subitizing gives way to counting, for all groups. Thus, an important finding was that the DS group and the WS group used the same enumeration processes used by the TD group, in all experimental conditions. This led to another important finding: individuals with WS and individuals with DS in this study were able to perform both perceptual subitizing and conceptual subitizing. This result was supported by three sources of evidence: accuracy, RT and fixation count. This is in contrast with the conclusions of the study by Sella et al. (2013), where children with DS showed a pattern consistent with the use of a serial counting routine, 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, that could be either a digit or a set of dots, with another set of dots that could either match or differ for one dot from the target image. The task set-up by Sella et al. (2013), therefore, did not only measure enumeration skills but relied on the participant to compare different stimuli as well and thus, relied heavily on the participants' working memory abilities, in contrast to the current study.

In order to investigate whether there were different scanning behaviours between the groups and whether this affected the enumeration processes, we analysed the eye movements. In line with previous research in TD and DD populations that showed that saccadic movements take place in the subitizing range (Schleifer & Landerl, 2011; Watson et al., 2007), we found that, on average, all groups showed more than one fixation when subitizing.

As for the experimental conditions where participants were counting, we found that, on average, for all groups the 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. A greater number of fixations than the number of dots to be enumerated was also reported on the restricted 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.

In line with our predictions and with research by Viñuela-Navarro et al. (2017), we were able to show that individuals with DS presented an overall gazing behaviour characterized by significantly shorter fixation duration in all the experimental conditions. Interestingly, their fixation instability did not have a significant effect on their performance in the subitizing condition, but it seemed to have a negative impact on their accuracy rates when counting. This result supports the argument that general basic-level abilities may impact the development of number abilities (Van Herwegen & Karmiloff-Smith, 2015). In the current study, the frequency and the average length of participants' eye movements influenced their counting skills. This result also provides new evidence in the debate regarding counting skills in individuals with DS (Abdelahmeed, 2007). For a long time, it has been debated whether individuals with DS have a superficial or a deep understanding of the concept of cardinality and, accordingly, of counting. The current findings highlight that counting performance in DS might not be explained by a domain-specific skill, but rather by their scanning patterns. This result also provides further insight into the underlying mechanisms of subitizing and counting in that the same impairment (i.e. fixation instability) affected subitizing and counting differently. Thus, different mechanisms underlie these enumeration abilities (see Schleifer & Landler, 2011 for a similar explanation). Further research is required, especially regarding the investigation of the nature and the development of fixation instability in DS and the relation between eye movements and chronological age.

Our results showed that, overall, individuals with WS presented shorter fixations than the TD group. In particular, individuals with WS showed significantly shorter fixation duration in the D1-D3 condition and in the R4-R6 condition. This is contrary to our predictions, but can be explained by the age range of our sample. In fact, "sticky fixation" has been reported in several studies on mathematical development, face recognition, visuo-spatial abilities and language learning, but only in toddlers and children with WS younger than 5 years old (see Van Herwegen, 2105 for a review). Given that the youngest participant with

WS in our sample was 8 years old, this may explain the absence of a “sticky fixation” pattern in the current study.

Finally, examination of the correlations (Table 11) showed that accuracy in conceptual subitizing correlated significantly with numerical competence in all the groups. Moreover, accuracy in conceptual subitizing showed statistically significant correlations with D1-D3 accuracy both in the TD group and in the WS group. These findings fit well with the hypothesis that conceptual subitizing develops with number knowledge and follows a trajectory that originates in the perceptual subitizing ability (Sarama & Clements, 2009), at least in the TD and WS populations. Moreover, correlation analyses suggested that accuracy in counting was related to numerical competence and mental age in all the groups. Interestingly, accuracy in counting was strongly correlated with accuracy in conceptual subitizing in the DS group. Because correlations do not provide any insight into cause and effect, these results should be interpreted with caution. Nonetheless combining this finding with the fact that individuals with DS were more accurate in conceptual subitizing than they were in counting, might suggest that conceptual subitizing could be leveraged in educational training programmes as an alternative strategy to counting to support enumeration abilities in particular, and symbolic mathematical abilities in general, in individuals with DS. A recent study on TD children by Paliwal and Baroody (2020) has shown how developing perceptual subitizing abilities might facilitate the understanding of the cardinality principle and it has suggested that the development of the 1-4 subitizing range is critical for achieving general cardinality principle knowledge and its understanding. Moreover, 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 (G. S. Starkey & McCandliss, 2014). However, the use of conceptual subitizing and groupitizing requires conceptual development in number knowledge as it emerges and develops as children’s conceptual knowledge of numbers is being progressively enhanced (G. S. Starkey & McCandliss, 2014).

There are some important limitations to consider when interpreting the results. First, only one trial per experimental condition was assessed. Having more trials per experimental condition could lead to a more robust estimation of the slopes for RT and fixation count. In addition, the current study included participants from a wide age range. This was to reach a reasonable sample size with the WS group, given the difficulties related to eye-tracking 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. Finally, although the lack of a control group matched for chronological age was justified by the scope of the present study, it did limit the conclusion regarding whether all processes were age appropriate.

To conclude, the current study examined both perceptual subitizing and conceptual subitizing in individuals with WS and DS for the first time and showed that both groups use both types of subitizing. Moreover, our findings show that individuals in our sample presented shorter fixation duration than those observed in the control group when performing an enumeration task. In particular, we showed that individuals with DS presented significantly shorter fixation durations than the control group. However, this fixation instability did not affect their performance when subitizing but could explain their low accuracy rates when counting. Thus, this appears to suggest that domain-general attentional processes may contribute to delay in mathematical abilities in individuals with DS. Whether these domain-general processes are related to specific mathematical underachievement in individuals with DS should be further investigated. Moreover, these findings together with our previous studies investigating difficulties in mathematical domain-specific skills in the same group of participants (Simms, Karmiloff-Smith, Ranzato, & Van Herwegen, 2020; Van Herwegen, Ranzato, Karmiloff-Smith, & Simms, 2019, 2020) seem to suggest that the reported difficulties in participants with DS could stem from weakness in general-domain attentional processes, rather than from basic domain-specific skills. Findings from this study also suggest that, for those with WS experience with subitizing could overcome their visuo-spatial difficulties when counting. These findings might have implications for interventions and educational programmes with regards to how to support counting abilities in individuals with DS and WS, and subsequently their symbolic mathematical abilities.

Table 7 –

Averaged slopes, intercepts and R2s of regression lines and Standard Deviations (SD) for RT and fixation count, separately computed for the three groups for each condition aggregated on pattern arrangement (dice, random) and numerosity range (subitizing, counting)

	Dice Pattern						Random Pattern					
	Subitizing Range			*Counting Range*			*Subitizing Range*			*Counting Range*		
	TD	WS	DS	TD	WS	DS	TD	WS	DS	TD	WS	DS
RT												
Slope	.02 (.23)	.12 (.28)	.18 (.33)	.24 (.47)	.37 (.91)	.37 (.85)	.14 (.24)	.25 (.67)	.19 (.31)	1.06 (.86)	1.18 (1.08)	1.56 (1.76)
Intercept	1.33 (.67)	.95 (.49)	.92 (.49)	.48 (2.15)	.14 (3.84)	.07 (3.66)	.99 (.50)	.71 (.99)	.79 (.54)	-2.02 (4.22)	-2.68 (4.67)	-4.19(7.53)
R^2	0.45	0.61	0.54	0.53	0.62	0.62	0.55	0.60	0.60	0.80	0.80	0.94
N	25	23	24	23	22	21	24	24	24	22	21	20
Fixation Count												
Slope	.84 (1.37)	.80 (1.59)	1.73 (2.03)	.52 (1.52)	.52 (1.70)	1.14 (4.07)	.55 (1.06)	.43 (1.66)	1.27 (1.86)	2.90 (2.86)	3.21 (2.28)	5.35 (6.14)
Intercept	2.13 (3.49)	2.36 (3.51)	2.51 (4.10)	2.00 (7.46)	3.57 (7.98)	2.90 (18.41)	2.73 (2.30)	3.24 (2.67)	2.53 (2.94)	-5.84 (13.93)	-6.10 (10.21)	-12.14 (25.83)
R^2	0.59	0.63	0.63	0.44	0.56	0.50	0.48	0.51	0.67	0.80	0.79	0.91
N	25	22	24	23	22	21	21	23	22	20	21	20

Table 11

Spearman's correlation between accuracy scores in the four experimental conditions, Chronological Age (CA), Raven's Coloured Progressive Matrices Test score (RCPM) and WIAT-II score separately computed for the three groups

	Dice Subitizing Accuracy			Random Subitizing Accuracy			Dice Counting Accuracy			Random Counting Accuracy		
Measures	TD (n = 25)	WS (n = 25)	DS (n = 24)	TD (n = 25)	WS (n = 25)	DS (n = 24)	TD (n = 25)	WS (n = 25)	DS (n = 24)	TD (n = 25)	WS (n = 25)	DS (n = 24)
CA	.171	.326	.305	.242	n/a	.359	.239	.608 (**)	.396	.464 (*)	.348	.469 (*)
RCPM	-.157	.156	.109	.342	n/a	.207	.140	.380	.376	.439 (*)	.452 (*)	.484 (*)
WIAT-II	.346	.315	.165	.000	n/a	.055	.614 (**)	.555 (**)	.418 (*)	.414 (*)	.403 (*)	.553 (**)
D1-D3 Accuracy												
R1-R3 Accuracy	-.042	n/a	-.091									
D4-D6 Accuracy	.455 (*)	.489 (*)	.201	-.114	n/a	.244						
R4-R6 Accuracy	.350	-.137	.313	.350	n/a	.121	.280	.231	.686 (**)			

Note: * $p < 0.05$, ** $p < 0.01$, n/a: correlations were not calculated due to variance = 0.

Acknowledgements

We would like to thank all participants and their families for taking part in this project. Further thanks go to Lorenzo Barasti for his assistance with the data analysis. This project was supported by a Grant from the British Academy (SG151035).

References

Abdelahmeed, H. (2007). Do children with Down syndrome have difficulty in counting and why. *International Journal of Special Education, 22*, 1–11.

Ansari, D., Donlan, C., & Karmiloff-Smith, A. (2007). Typical and atypical development of visual estimation abilities. *Cortex, 43*(6), 758-768. doi:10.1016/S0010-9452(08)70504-5

Ashkenazi, S., Mark-Zigdon, N., & Henik, A. (2013). Do subitizing deficits in developmental dyscalculia involve pattern recognition weakness? *Developmental science, 16*(1), 35-46. doi:10.1111/j.1467-7687.2012.01190.x

Benoit, L., Lehalle, H., & Jouen, F. (2004). Do young children acquire number words through subitizing or counting? *Cognitive Development, 19*(3), 291-307. doi:10.1016/j.cogdev.2004.03.005

Brigstocke, S., Hulme, C., & Nye, J. (2008). Number and arithmetic skills in children with Down syndrome. *Down Syndrome Research and Practice*, 74-78. doi:doi:10.3104/reviews.2070

Brown, J. H., Johnson, M. H., Paterson, S. J., Gilmore, R., Longhi, E., & Karmiloff-Smith, A. (2003). Spatial representation and attention in toddlers with Williams syndrome and Down syndrome. *Neuropsychologia, 41*(8), 1037-1046. doi:10.1016/S0028-3932(02)00299-3

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307-314. doi:10.1016/j.tics.2004.05.002

Gordon, R., Smith-Spark, J. H., Newton, E. J., & Henry, L. A. (2020). Working memory and high-level cognition in children: An analysis of timing and accuracy in complex span tasks. *Journal of Experimental Child Psychology, 191*, 1-20. doi:10.1016/j.jecp.2019.104736

Goukon, R. (2016). *Examining Children's Conceptual Subitizing Skill and its Role in Supporting Math Achievement.* (Master's thesis). University of Calgary, Alberta.

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. doi:10.1111/bjdp.12032

Kapp, M. E., Von Noorden, G. K., & Jenkins, R. (1995). Strabismus in Williams Syndrome. *American Journal of Ophthalmology, 119*(3), 355-360. doi:10.1016/S0002-9394(14)71180-8

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. doi:10.2307/1418556

Krajcsi, A., Szabo, E., & Morocz, I. A. (2013). Subitizing is sensitive to the arrangement of objects. *Experimental Psychology, 60*(4), 227-234. doi:10.1027/1618-3169/a000191

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. doi:10.1177/0734282908330586

Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8–9-year-old students. *Cognition, 93*, 99–125.

Martens, M. A., Wilson, S. J., & Reutens, D. C. (2008). Research Review: Williams syndrome: A critical review of the cognitive, behavioral, and neuroanatomical phenotype. In (Vol. 49, pp. 576-608). Oxford, UK.

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. doi:10.1016/j.cogdev.2009.09.007

Nye, J., Fluck, M., & Buckley, S. (2001). Counting and cardinal understanding in children with Down syndrome and typically developing children. *Down Syndrome Research and Practice, 7*(2), 68-78.

O'Hearn, K., Hoffman, J. E., & Landau, B. (2011). Small subitizing range in people with Williams syndrome. *Visual Cognition, 19*(3), 289-312. doi:10.1080/13506285.2010.535994

Ozdem, S., & Olkun, S. (2019). Improving mathematics achievement via conceptual subitizing skill training. *International Journal of Mathematical Education in Science and Technology*. doi:10.1080/0020739X.2019.1694710

Paliwal, V., & Baroody, A. J. (2020). Cardinality principle understanding: the role of focusing on the subitizing ability.

Piazza, M., Mechelli, A., Butterworth, B., & Price, C. J. (2002). Are subitizing and counting implemented as separate or functionally overlapping processes? *NeuroImage, 15*(2), 435-446. doi:10.1006/nimg.2001.0980

Reeve, R., Reynolds, F., Humberstone, J., & Butterworth, B. (2012). Stability and change in markers of core numerical competencies. *Journal of Experimental Psychology: General, 141*(4), 649-666. doi:10.1037/a0027520

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), e79711. doi:10.1371/journal.pone.0079711

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. Abingdon: Routledge.

Schleifer, P., & Landerl, K. (2011). Subitizing and counting in typical and atypical development. *Developmental science, 14*(2), 280. doi:10.1111/j.1467-7687.2010.00976.x

Sella, F., Lanfranchi, S., & Zorzi, M. (2013). Enumeration skills in Down syndrome. *Research in Developmental Disabilities, 34*(11), 3798-3806. doi:10.1016/j.ridd.2013.07.038

Silverman, W. (2007). Down syndrome: Cognitive phenotype. *Mental Retardation and Developmental Disabilities Research Reviews, 13*(3), 228-236. doi:10.1002/mrdd.20156

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. doi:10.1007/s10803-019-04268-7

Starkey, G. S., & McCandliss, B. D. (2014). The emergence of "groupitizing" in children's numerical cognition. *Journal of Experimental Child Psychology, 126*, 120-137. doi:10.1016/j.jecp.2014.03.006

Starkey, P., & Cooper, R. G. (1995). The development of subitizing in young children. *British Journal of Developmental Psychology, 13*, 399-420.

Sullivan, J. L., Juhasz, B. J., Slattery, T. J., & Barth, H. C. (2011). Adults' number-line estimation strategies: Evidence from eye movements. *Psychonomic bulletin & review, 18*(3), 557. doi:10.3758/s13423-011-0081-1

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 331-351. doi:10.1037/0096-1523.19.2.331

Van Herwegen, J. (2015). Williams syndrome and its cognitive profile: The importance of eye movements. *Psychology Research and Behavior Management, 8*, 143-151. doi:10.2147/PRBM.S63474

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, United Kingdom: 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. doi: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 : JARID*. doi:10.1111/jar.12730

Van Herwegen, J., & Simms, V. (2020). Mathematical development in Williams syndrome: A systematic review. *Research in Developmental Disabilities, 100*, 103609. doi: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.* Paper presented at the Child Vision Research Society XVI conference, Ulster University, Coleraine.

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. doi:10.1037/0096-1523.33.6.1389

Wechsler, D. (2005). Wechsler individual achievement test, 2nd edition (WIAT-II). London: Psychological Corporation.