Journal of Child Psychology and Psychiatry 47:2 (2006), pp 190–204 doi:10.1111/j.1469-7610.2005.01460.x

Are numerical impairments syndrome specific? Evidence from Williams syndrome and Down's syndrome

Sarah J. Paterson,[1,2] Luisa Girelli,[3] Brian Butterworth,[3] and Annette Karmiloff-Smith[1]

[1]Neurocognitive Development Unit, Institute of Child Health, UCL London, UK; [2]Infancy Studies Laboratory, Center for Molecular and Behavioral Neuroscience, Rutgers University, USA; [3]Institute of Cognitive Neuroscience, University College London, UK

Background: Several theorists maintain that exact number abilities rely on language-relevant processes whereas approximate number calls on visuo-spatial skills. We chose two genetic disorders, Williams syndrome and Down's syndrome, which differ in their relative abilities in verbal versus spatial skills, to examine this hypothesis. Five experiments assessed number skills in these two genetic syndromes and in their mental age (MA) and chronological age (CA) matched controls. **Methods:** Experiment 1 used a preferential looking paradigm with infants and toddlers to measure sensitivity to changes in numerosity. Experiment 2 measured reaction times in older children and adults in a numerosity comparison task with dots in a random pattern. Experiment 3 comprised a number battery that measured various forms of counting and simple arithmetic. **Results:** The WS infants displayed a level of performance equal to that of their CA-controls, whereas the DS infants failed to reach even the level of their MA-controls. By contrast, the older DS children and adults outstripped the older WS group in their numerosity abilities, with different patterns of errors in the two clinical groups. **Conclusions:** Differences in the infant and adult number phenotypes between these two genetic disorders are discussed with reference to the processing styles used by each group and how these might impact on their developmental trajectories. Theoretically, we highlight our contention that one cannot infer the infant starting state from the adult end state. Rather, the development process itself must be taken into account. **Keywords:** Williams syndrome, Down's syndrome, number development, adults, infants. **Abbreviations:** WS: Williams syndrome; DS: Down's syndrome; SDE: symbolic distance effect.

In everyday life, we are bombarded with number-relevant information. From infancy onwards, numbers are part of our environmental stimuli. We must learn to deal at least with: (1) approximate magnitude representation on the basis of which we can easily discriminate between large and small numerosities; (2) the Arabic system, e.g., 7 or 56; and (3) the verbal number system, e.g.,/seven/or/fifty-six/. These skills are vital in a variety of settings, from making judgements about whether something in a shop costs a reasonable amount, through to formal arithmetic. Because of the importance of numeracy for everyday functioning, the impairment of numerical skills in developmental disorders is a cause for serious concern. The investigation of how general cognitive level and/or more specialised number systems are involved in numeracy impairments is therefore crucial.

The extent to which numerical processing is specialised and independent of general cognitive ability is a subject of much debate. The existence of an independent module for number, such as that suggested by Butterworth (1999), seems to be supported by evidence from adult neuropsychological patients and brain imaging research. These studies provide data suggesting that the brain areas subserving number tasks differ from those used in language and reasoning tasks. For example, one patient with semantic dementia and damage to the left temporal lobe was found to be unable to name objects, but could read and write numbers and compare numerosities (Cappelletti, Butterworth, & Kopelman, 2001). In contrast, other patients exist who are unable to perform number tasks such as calculation and subitizing, but have no difficulty with spoken number language (Cipolotti, Butterworth, & Denes, 1991). However, it should be noted that neuropsychological patients have mature brains that had developed normally until brain insult and selective impairment. The brains of children with genetic disorders, by contrast, *develop differently* from embryogenesis onwards, frequently resulting in atypical brain anatomy, brain biochemistry and brain electrophysiology. It is likely therefore that the mechanisms underlying numerically relevant behaviour in such disorders are very different from impaired number computations in adult neuropsychological patients (Karmiloff-Smith, 1998; Paterson, Brown, Gsödl, Johnson, & Karmiloff-Smith, 1999; Karmiloff-Smith, Brown, Grice, & Paterson, 2002; Karmiloff-Smith, Scerif, & Ansari, 2003).

Even if number were to some degree to be independent of other aspects of cognition, there is

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

mounting evidence that language plays an important role in the *development* of number skills and in the storage and retrieval of number facts. For example, data from a recent brain imaging study suggest that within the number domain, different functions are subserved by different brain areas. One study revealed that when participants had to perform exact arithmetic, language areas were activated, whereas for approximate calculation visuo-spatial networks in the parietal lobes were activated (Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999). In addition, for bilingual participants the language in which the problems were taught and tested had an effect on performance. For exact calculation, adults performed better in the language in which they were taught. By contrast, results for numerical approximation were language independent. As far as children are concerned, as development proceeds they will need to use language to map the exact system of sequential numbers to their representations of numerosity, in order to enumerate sets of objects and not merely reproduce the counting sequence.

Despite the importance of language in some types of number tasks, such as retrieval of number facts from multiplication tables as well as counting, or reading/writing numbers, other kinds of numerical ability turn out to be less dependent on language. Current models postulate that magnitude representation, such as that used in approximation and number comparison, is mediated by a pre-verbal system which is available to rats and pigeons as well as to human infants (e.g., Antell & Keating, 1983; Washburn & Rumbaugh, 1991; Wynn, 1990). Indeed, well before the onset of language, very young infants are able to distinguish between different arrays of small numbers of objects.

In older typically developing children and adults, numerosity comparisons are also likely to rely on a non-verbal mechanism. In such tasks, participants are asked which is the larger of two numerosities. This taps the mapping between non-verbal magnitude representations and Arabic numerals or dot arrays. Such numerical comparison tasks give rise to a special effect, known as the symbolic distance effect – SDE (Moyer & Landauer, 1973). Participants take longer to discriminate between numerosities that are close together, e.g., 3 and 4, than those that are far apart, e.g., 3 and 7. It is likely that the SDE results from variability in the magnitude representation. If the representation of the mental number line in the brain is relatively fuzzy (Dehaene & Cohen, 1994), then it will be easier to distinguish the larger of two numbers if they are far apart because in such cases overlapping activation is reduced.

The SDE is a robust finding in the performance of young typically developing children from at least 6 years of age (Sekuler & Mierkiewicz, 1977; Duncan & McFarland, 1980). Moreover, research with adults has shown that the SDE is also present in studies which use two languages with very different number symbols, e.g., English and Persian (Dehaene, Bossini, & Giraux, 1993). The SDE is therefore likely to stem from a non-verbal representation and to be a behavioural indicator of one important aspect of the number system, i.e., the representation of the quantity that various numerical symbols represent. An understanding of number magnitude as tapped by the SDE may well form an essential basis for the development of arithmetic in typically developing children (Butterworth, 1999; Butterworth, Zorzi, Girelli, & Jonckheere, 2001).

In order to examine the role of language ability in number development and to investigate the extent to which number may be independent of other aspects of cognition, number abilities in two clinical groups with differing cognitive profiles were assessed. Individuals in the two genetic disorders, Williams syndrome (WS) and Down's syndrome (DS), have a similar degree of overall cognitive impairment but their strengths and weaknesses across domains differ. In WS, some aspects of language performance are relatively proficient, while other skills, particularly spatial cognition, are seriously impaired (Arnold, Yule, & Martin, 1985; Bellugi, Bihrle, Jernigan, Trauner, & Doherty, 1990; Donnai & Karmiloff-Smith, 2000). In DS, the opposite pattern tends to hold. Individuals with DS have particular difficulties with language, but are less impaired on tasks tapping spatial skills (Rosin, Swift, Bless, & Vetter, 1988). This different language/spatial imbalance across the two clinical groups, despite similar overall cognitive ability, should allow us to tease apart those aspects of number which are more directly constrained by language, those linked to non-verbal abilities, and those which appear to be affected by overall levels of intelligence. In addition, we can begin to examine whether these clinical populations process number via a normal developmental trajectory or whether they use different mechanisms to deal with number as a consequence of atypical development of the entire cognitive system (Karmiloff-Smith, 1998; Karmiloff-Smith et al., 2002).

To date there has been little systematic investigation of number skills in Williams syndrome, whereas several studies exist on the numerical competence of individuals with Down's syndrome. From previous research, it is clear that those with DS have difficulties with numeracy and that the degree of their problems is related to their general cognitive level. Studies of counting have suggested that developmental level, and not Down's syndrome per se, is a good indicator of success (Caycho, Gunn, & Siegal, 1991). In a more recent, detailed study, Nye, Clibbens, & Bird (1995) investigated the link between language and number in a group of 16 children with DS aged 7–12;6 years. They found that the two measures correlated highly. The tests also correlated with standardised measures of numerical competence from the British Ability Scales (Elliott, Smith, & McCulloch, 1996) and the Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983). Performance of these children with DS on number tasks was also significantly related to their grammatical comprehension, as measured by the TROG (Bishop, 1983), but did not correlate significantly with receptive vocabulary. These complex relationships are likely to be due to the type of numerical skills investigated. Different numerical tasks draw to varying degrees upon verbal skills; numerosity comparison, for instance, has no necessary verbal component, while other tasks such as arithmetic and counting involve language capacities, although to varying degrees.

In contrast to DS, there has been very little research, as mentioned above, on numerical abilities in WS. The few existing studies indicate that older children and adults with WS can count and abide by counting principles (Hughes, 1995) and that teenagers and adults with WS have difficulty with Piagetian number conservation tasks which are already mastered by 6 or 7 years in typical development. On the basis of such findings, Bellugi, Marks, Bihrle, and Sabo (1988) argued that individuals with WS have severe number impairments. However, Bellugi's early research was based on a very small number of participants.

Another WS study compared overall cognitive level with performance on numerical subtests of standardised tests. In a longitudinal study of cognitive skills in WS, Udwin, Davies, and Howlin (1996) measured numerical ability and found that at the mean chronological age of 12 years, all the participants had a test age of 8;3 years on the arithmetic subscale of the WISC R, but when tested again at a mean age of 21 years on the WISC III they had a test age of 8;1 years. These results suggest that as of adolescence performance on arithmetic scales remains at a plateau in WS, at around the level of an 8-year-old, i.e., well below chronological age. This difference is unlikely to be simply due to general cognitive impairment, because the number deficit was considerably greater than the lag between CA and language performance and between CA and overall MA. Udwin et al.'s study reveals true deficit in numerical ability in the face of better language skills, but it does not examine various numerical abilities in their own right, nor does it investigate the particular bases of the WS number impairment. Moreover, none of the aforementioned studies has made direct number comparison across different clinical groups, nor examined the relative role of language in number ability, nor assessed the infant precursors to number abilities in developmental disorders.

Given the similarity of overall level of cognitive functioning found in infants, children and adults with WS and DS of the same chronological age, we aimed to compare their performance on a variety of number-specific tasks to examine whether numerical difficulties are syndrome specific. We therefore assessed the numerical abilities of infants, older children and adults with WS and DS, keeping the methodology as close as feasible across the developmental age span. First, with both infants and adults we used tasks that tap approximate numerosity comparisons. Second, with the older children and adults, we employed tasks that tap basic numerical skills such as counting, Arabic numeral reading, matching dots to numerals, seriation of dots and numerals, and calculation. These have no obvious direct infant counterpart and so were only run with the older subjects.

If indeed language plays a role in some aspects of exact numerical cognition, we would expect individuals with WS, with their superior verbal competence, to perform better on such tasks than their counterparts with DS matched on CA and overall cognitive ability. By contrast, for tasks that do not rely on language, we would expect similar performance from both groups, or the DS group to outstrip the WS group where tasks call on visuo-spatial skills. Alternatively, the results from these clinical groups may turn out to pinpoint more than delay, indicating that they are using different mechanisms from controls to process number, as a result of atypical developmental trajectories.

Experiment 1: Numerosity comparison in infancy

Method

Participants. Fifty-nine infants took part in this experiment. Eleven infants with WS and 18 infants with DS were tested, matched on both chronological age (CA) and mental age (MA), as well as 16 MA-matched and 14 CA-matched typically developing controls. Infants were matched on overall MA using the Bayley Scales of Infant Development II (Bayley, 1993). The mean CA and MA of each of the 4 groups are presented in Table 1. Given the uneven cognitive profile in adults with WS, with language skills outstripping spatial skills, and in DS with spatial skills outstripping language skills, one hypothesis is that the DS infants will perform significantly better than the WS infants in tasks tapping magnitude rather than exact numerosity judgements.

Table 1 Chronological and mental ages for each group

	Mean CA (months)	SD	Range	Mean MA (months)	SD	Range
WS	30	5.36	24–36	16.4	2.65	12–21
DS	30	4.89	24–36	15.6	2.43	12–20
MA	15.4	2.52	12–20	15.1	2.66	11–21
CA	30.4	5.30	24–36	30.4	5.32	25–40

Procedure. Infants were tested using a replica of the basic Fagan apparatus (Fagan, 1970), in which two stimulus cards are presented simultaneously. The display was illuminated by a fluorescent light positioned out of the infant's view. In the centre of the stage was a peephole .625 cm in diameter, through which one of the experimenters, blind to the position of the stimuli, could see the direction of the visual fixations of the infant.

Each infant was tested in a special infant seat. The testing apparatus was then wheeled into position, with the display stage centred directly over the infant. At this point, the infant could no longer see the parent. The stimuli were then placed simultaneously into the two compartments by Experimenter I and, once the infant's attention was attained by talking or by shaking a rattle, the familiarisation trials began. Participants were familiarised with a sequence of pairs of stimuli depicting a variety of arrays of 2 objects, in different configurations. Each infant was shown 6 familiarisation trials. After familiarisation with sets of 2, the infant was presented simultaneously with one card displaying new objects but the old numerosity (2) and another display also showing new objects but a novel numerosity (3). The side on which the novel numerosity appeared was randomised, and Experimenter II, who measured the cumulative looking time over each trial, was blind to the position of that card. Experimenter II held a stopwatch in each hand and timed the infant's looking to the left versus the right stimulus item by observing the direction of looking from the infant's pupil. Reliability using this procedure has been shown to be high (Haaf, Brewster, de Saint Victor, & Smith, 1989; O'Neill, Jacobson, & Jacobson, 1994).

A beeper was set to a fixed length for the familiarisation and test trials, and signalled when a trial was to end (10 seconds for familiarisation, 5 seconds for test). Between each trial, Experimenter I pulled back the display stage from the infant's view, recorded the data called out by Experimenter II, changed the stimuli, obtained again the infant's attention, centred the infant's gaze, and finally closed the stage, exposing the next stimuli to the infant.

Stimuli. The stimulus set consisted of six white cards 17.7 × 17.7 cm. Differently coloured pictures of two objects were mounted on each card. These objects differed in size and position for each pair, so that density of filled space was controlled across familiarisation and test cards. The objects differed on each card and included airplanes, cats, trees, cars etc. The 2 test cards both depicted new objects, one displaying 2 items, the other 3 items.

Results

The mean looking times to the novel (3 object pictures) and familiar (2 object pictures) numerosities were calculated for each infant. The data were entered into a repeated measures Anova with group as the between-subjects factor (WS, DS, MA-matched, CA-matched) and numerosity (novel or familiar) as the within-subjects factor. An effect of numerosity $F(1,55) = 15.08, p < .001$ and of group $F(3,55) = 6.29$, $p < .001$ was found. There was also a significant interaction of group by numerosity $F(3,55) = 5.03$, $p < .001$. This suggests that looking time to each type of stimuli (novel or familiar) differed depending on the group. In order to investigate where the differences in performance resided, post hoc *t*-tests were carried out. These *t*-tests comparing the mean scores for novel and old numerosities for each group revealed a significant difference in looking time between numerosities for the WS and for the CA-matched and MA-matched groups ($t(10) = 3.26$, $p < .001$, $t(13) = 3.87$, $p < .002$ and $t(14) = 2.18$, $p < .05$, respectively). There was, however, no such difference for the DS group. The results thus show that despite some previous work suggesting a serious impairment in number in the endstate, WS participants in infancy seem to perform normally for small numerosities and look like their CA-matched counterparts. By contrast, the DS group displayed no discrimination of the difference between the novel and old numerosities. Results for the four groups of infants are presented in Figure 1.

In addition to the analyses reported above, the magnitude of difference between looking time to the novel stimuli and the familiar stimuli was examined for each group. A one-way Anova, with difference (looking time to novel stimuli – looking time to familiar stimuli) as the dependent variable and group as the independent variable, was carried out. These difference scores varied significantly between groups, $F(3,58) = 4.32$, $p < .001$. Post hoc tests were then conducted to compare the magnitude of these differences for each group. There was a significant difference in the magnitude between the WS and DS groups, Tukey HSD ($p < .05$). The WS group

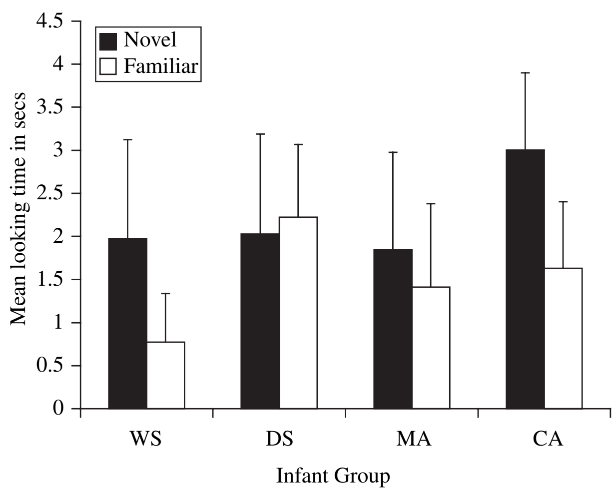

Figure 1 Mean looking time to novel and familiar numerosities: infants

difference was larger than the DS group difference. The magnitude of difference for the CA-matched group and DS group were also significantly different, Tukey HSD ($p < .05$), with the CA group exhibiting a larger difference than the DS group. There were no other differences.

Discussion

Experiment 1 investigated numerosity discrimination in infants with WS and DS and their typically developing counterparts. This enabled an initial characterisation of numerical processing early in development in these clinical populations (Paterson et al., 1999). Although the present study provides only one measure of infant number ability, it is a test which has been replicated a number of times in a variety of forms for normal development (e.g., Strauss & Curtis, 1981; Starkey, Spelke, & Gelman, 1990). Our data also reveal that despite being older than infants in previous studies, our typically developing infants – whether MA- or CA-matched – looked significantly longer at the arrays of the novel numerosity (3) after familiarisation with arrays of 2, and thus discriminated between different numerosities. The WS infants performed like their typically developing counterparts, whereas the DS infants did not. The evidence from the present study suggests that one of the fundamental underpinnings of number development is present in infants with WS, whereas it is seriously delayed in infants with DS. This is the ability to discriminate small numerosities. Interestingly, the WS group looked at the familiar stimulus for less time than the DS group. It could be that the infants with WS noticed the novel stimulus, fixated on it, and found it difficult to disengage and shift attention back to the familiar stimulus. Sticky fixation has already been reported in infants with WS compared to those with DS in tasks investigating the planning of saccadic eye movements (Brown et al., 2003). This cross-syndrome difference in performance on this task is particularly interesting, given the differing pattern of abilities that we report in the following experiment comparing older children and adults with WS and DS, to which we now turn.

Experiment 2: Symbolic Distance Effect

Like Experiment 1, Experiment 2 targets numerosity judgements, but this time in older children and adults.

Method

Participants. Eight older children and adults with WS and 9 with DS from similar socio-economic backgrounds took part in this experiment. Eight typically developing participants were matched on the British Ability Scales (BAS II), and 8 were matched on chronological age to the two clinical groups. Adults were matched using overall MA from the British Ability Scales II (Elliott, Smith, & McCulloch, 1996). This measure includes visuo-spatial, non-verbal reasoning and verbal scales. The cognitive profiles of the WS and DS groups are uneven so, when matching by overall MA, the visuo-spatial or verbal skills of the two groups may differ. A comparison of verbal skills in the two groups of adults who completed the task revealed that the WS group had significantly higher verbal scores than the DS group ($t\,(15) = 3.14$, $p < .05$) However, unusually but interestingly for our hypothesis, there was no significant difference in the visuo-spatial skills of the two groups ($t\,(15) = -1.49$, ns).

Two of the participants with DS did not complete the task; one pressed the same button continuously and would not attend to the screen, and another refused to do the task. They were removed from analyses. The mean CA for each group, including only those participants who completed the task, and the mean MA from the BAS, are reported in Table 2. Although the chronological age range of the atypically developing groups is wide, the youngest were 10, the age at which in normal development children would easily succeed on these number tasks.

Stimuli. Participants were presented simultaneously with pairs of arrays of dots on a computer screen. Their attention was captured by a cross appearing in the middle of the screen prior to the display of each stimulus. The pairs always had an inter-stimulus interval of 995 milliseconds and stimuli remained on the screen until a response was made. The dots were presented in a random configuration.

Design. Two aspects of the stimuli were varied in each experiment: (1) Participants were presented with two arrays of dots which had numerosities that were either close together (a difference of 1, 2 or 3) or

Table 2 Mean chronological and mental ages for each group

	Mean CA (years)	Range	Mean MA (years)	Range
WS	20;9	10;11–32;9	6;9	5;1–9;4
DS	24;3	11;4–35;3	5;9	5;1–6;4
MA-	6;11	5;2–8;11	–	–
CA-	21;1	9;10–29;8	–	–

Table 3 Stimulus pairs for number comparison

Split	Distance	Pairs
Close	1	2–1, 3–4, 7–6, 9–8
	2	3–1, 4–6, 9–7
	3	1–4, 9–6
Far	5	6–1, 8–3, 7–2, 9–4
	6	7–1, 8–2, 9–3
	7	8–1, 9–2

far apart (a difference of 5, 6 or 7), as in Table 3 (this was called the 'split'); and (2) each pair of arrays was presented 4 times, with the larger numerosity on the left twice and on the right twice, producing a total of 72 trials. Order of presentation varied randomly for each participant.

Procedure. Each participant was presented with a practice block before testing proper began. MA-matched controls and participants with DS and WS did a full practice block of 16 trials. Because CA controls could perform the task immediately, these participants were only given three practice trials, in order to shorten the overall test session. The test blocks consisted of 72 trials. Participants were told to respond to the stimuli by pressing one of two keys on a computer keyboard (Z and M) which were highlighted with bright velcro stickers. They were told to depress the key on the side of the array with the larger number of dots. If necessary, this was demonstrated during the practice block. For the clinical groups, these instructions were repeated where necessary. Encouragement was given throughout to keep participants on task, given the long duration of the session.

Results

Each participant's results were sorted according to the split (close or far) and then into blocks of four trials, which were identical. Incorrect responses were removed and the median reaction time for each four-trial block was calculated. Outliers for each participant were removed in order to prevent the means from being skewed. Any extreme values are unlikely to be due to inability to make numerosity comparisons but to a lapse in attention. Outliers were found using a box and whisker plot. The main body of the plot represented values between the 25th and 75th percentile. Any reaction times more than 1.5 box lengths above or below these values were excluded from the analysis, according to the instructions provided in the SPSS statistical package (Kinnear & Gray, 1999). The percentage of trials omitted as outliers was: 4.8% for the WS group, 5.6% for the DS group, 7.9% for the CA-matched group and 4.2% for the MA-matched group. The mean reaction times and standard deviations for each group are presented in Figure 2, along with the percentage of trials correct. It should be noted that because the number of participants in each study was relatively small, caution should obviously be exercised when considering the results.

Symbolic Distance Effect: Williams syndrome group and Down's syndrome group. As illustrated in Figure 2, the WS group's responses were faster overall than the DS group. *T*-tests were carried out to investigate the difference in mean reaction times for trials with a far split or a close split. For the WS

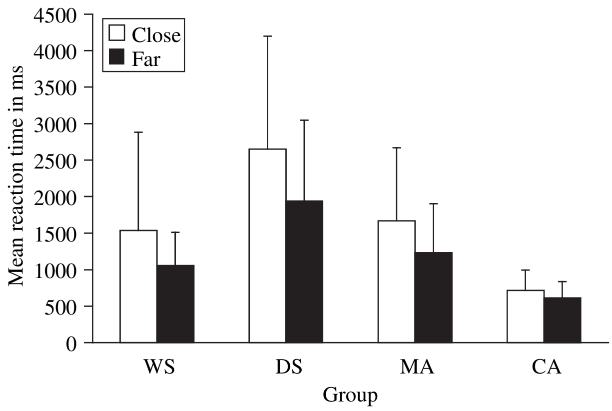

Figure 2 Distance effect for random dot comparison: older children and adults

group, there was no significant difference in reaction times for close pairs (mean 1536 ms, SD 1346 ms) and far pairs (mean 1056 ms, SD 454 ms), t (7) = 1.42, n.s., or across distance with individual distances as independent variables, $F(5,35) = 1.821$, n.s. This means that the WS group did not exhibit a robust symbolic distance effect (SDE), although the results were in the right direction. Although the DS group's responses were slower, they did by contrast display a robust symbolic distance effect, t (6) = 4.17, $p < .01$, with a mean RT of 2574 ms (SD 1570ms) for close pairs and 2096 ms (SD 1196 ms) for far pairs. They also exhibited the effect over individual distances, $F(5,30) = 6.20$, $p < .01$. An analysis of individual subjects' performance revealed that 71% DS subjects but only 50% of the WS participants exhibited the SDE. Given the much slower RTs in the DS group, data from the DS and WS group were also normalised to take response time into account. This was done for each subject by dividing the RTs of each individual trial by the mean reaction time across all trials. The SDE was then examined in each group. There was a significant difference in RT for close pairs and far pairs for both the WS, t (7) = 2.52, $p < .05$, and DS, t (7) = 5.05, $p < .01$. So in this case it appears that when RT is controlled both atypical groups show an SDE. When considering these results it should be noted that in a repeated measures Anova, with split as the within-subjects variable and group, WS or DS, as the between-subjects variable, there was no significant effect of group, $F(1,13) = 2.853$, n.s., and no interaction between group and split. However, this result may be due to the small sample sizes in this study and the large variability.

Symbolic Distance Effect: Control groups. A significant difference was found between the mean reaction time for stimuli with a close split (mean 1494 ms, SD 702.93 ms) and stimuli with a far split (mean 1203.39 ms, *SD* 626.29 ms) for the MA-matched children, t (7) = 2.18, $p < .05$. This was also true for the CA-matched group, t (7) = 3.83, $p < .01$, with a mean RT of 695 ms (*SD* 263.51) for close pairs and 613 ms (*SD* 236.7 ms) for far pairs.

Accuracy: all groups. As shown in Figure 3, all four groups made relatively few errors. Heterogeneity of variance in these data called for non-parametric statistics. The analysis indicated that for random dot arrays, the WS group made the greatest number of errors both when close pairs and the total number of errors were analysed.

A Kruskal–Wallis one-way Anova revealed that there was a significant difference between groups in the total proportion of trials correct, chi-square = 7.99; df 3, $p < .05$. When the proportion of items correct was calculated separately for close pairs and far pairs, there was no effect of group for far pairs, chi-square = 2.06; df 3, n.s. For close pairs, there was an effect of group, chi-square = 8.36; df 3, $p < .05$. Post hoc tests did not reveal where this difference lay for close pairs or for the total percentage of items correct. However, the Kruskal–Wallis ranking indicated that the WS group performed worse than other groups for close pairs and for total number of items correct.

Discussion

Let us first consider reaction time data. The participants in the WS group behaved differently from the CA, MA and DS groups on the task in that as a group they failed to exhibit a robust distance effect. In other words, they did not take significantly longer to discriminate between arrays that have close numerosities, e.g., 2 vs. 3, and those that are far apart, e.g., 2 vs. 6. All other groups, including the DS clinical group, displayed a robust distance effect. When the overall reaction time of each participant was controlled for in the analysis, to take into account the huge differences in speed between the WS and DS groups, the DS group continued to exhibit the SDE. However, the WS group exhibited the SDE only when overall reaction time was taken into account. These results suggest that the representation of the number line in WS is likely to be weaker than in the other groups.

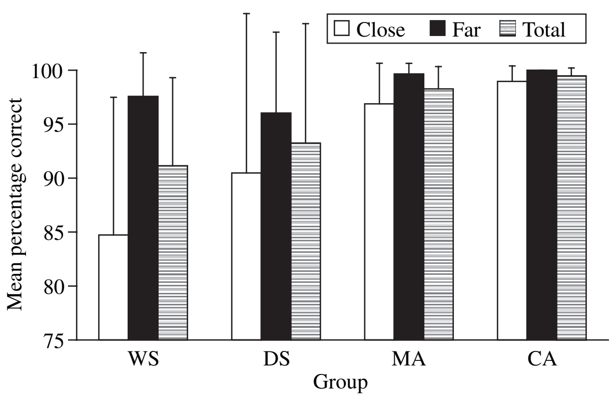

Figure 3 Accuracy for random dot comparison

Analysis of errors overall revealed that the WS group perform significantly worse than both control groups. The WS group was the least accurate of all groups. However, it is important to note that all groups performed well above chance, so participants could all do the task. No clear post hoc differences between groups were found for accuracy, although individuals with WS were the worst performers. However, again, it is important to note that the performance of the WS group was above chance level, so the results cannot be explained simply by task demands. The accuracy of both the MA-matched and CA-matched groups was high, with the expected lower accuracy for close pairs than for far pairs, because as discrimination between pairs becomes more difficult, more errors occur, together with an increase in reaction time.

While results should be interpreted with caution due to the small sample sizes and the large variability in reaction times, the results from the WS group suggest that there may be some anomalies in their basic numerical processing. In normal development, as early as 5 years old, the symbolic distance effect is a very robust indicator of the presence of a mature mental representation of numerosity. Since the WS participants had a mean chronological age of 21 and a mean mental age of nearly 7, it is clear that their numerosity representations are considerably weaker than would be predicted from their MA. By contrast, the DS group data follow a more typical trajectory. Despite their slow reaction times, they display a robust distance effect with dot displays, despite being the same chronological and mental age as the WS participants. In fact the mean MA of the WS group was somewhat higher than that of the DS group. This indicates that the weaker performance of the WS group is not simply due to general cognitive impairment. The performance patterns of the WS group were anomalous compared to the other groups. Indeed, some of the WS subjects even showed a trend towards an inverse distance effect, with close trials exhibiting faster reaction times than far trials.

The results from the current task with older childen and adults are particularly striking, given the findings from Experiment 1 testing infants. Our first experiment showed WS infants to be successful in discriminating between small numerosities. They dishabituated to a novel numerosity (3) after familiarisation with arrays of 2 objects. It is therefore possible that the individuals with WS develop their representations of *small* numerosities in a relatively normal way, since early on they are sensitive to changes in numerosity. However, it seems that they cannot clearly represent the *precise* nature of numerical changes later in development. By contrast, the group of infants with DS was unable to detect the change in small numerosities but, as our subsequent experiment showed, by later childhood and adulthood they perform better than their counterparts with WS. The pattern of results for DS point to an initial delay in the development of numerosity discrimination, but once they become able to represent small numbers, the majority of participants seem to use similar processes to normal controls in order to discriminate larger numbers. By contrast, the WS group's results suggest that many individuals follow an atypical developmental trajectory whereby the early proficiency with small numerosities is not built upon for the subsequent processing of large numbers.

If magnitude representations, or the mental number line, were normal in individuals with WS, then one would expect a typical symbolic distance effect. Its weakness in this clinical group suggests that number-related representations for these individuals are imprecise. The fact that the WS group performed significantly worse than the DS group indicates that, as predicted (Dehaene, Dupoux, & Mehler, 1990), there is no advantage of having better language abilities for numerosity judgement tasks that involve differences in magnitude.

In this sample, the WS group had higher verbal scores on the BAS II than the DS group (see Figure 4). Interestingly, when the association between three skills was assessed for each group: verbal performance, as assessed by verbal similarities and word definitions, non-verbal performance, as assessed by pattern construction and recall of designs, and quantitative reasoning on the BAS, there was no correlation between either verbal or non-verbal skills and the number scale for the DS group, but there was a significant correlation between both verbal skills, $r = .899$, $p < .01$, and non-verbal skills, $r = .73$, $p < .05$) and performance on the quantitative reasoning subscale for the WS group.

The correlations found in the data from the WS group suggest that these individuals may be relying more on non-numerical skills than the DS group when tackling quantitative reasoning tasks. A

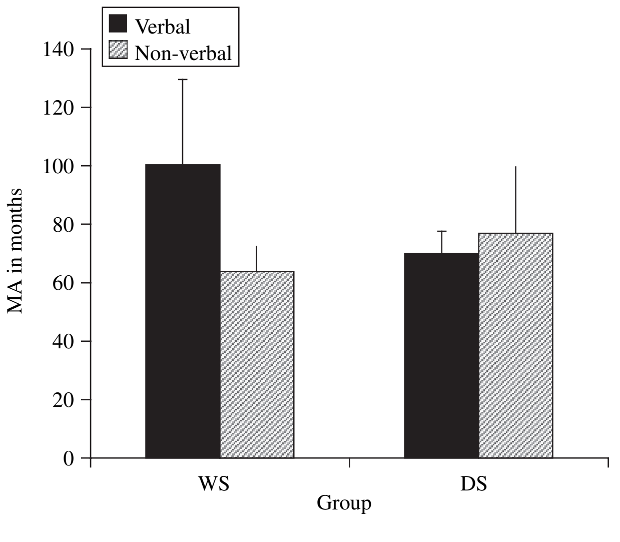

Figure 4 Verbal and non-verbal performance on the BAS II

correlation between spatial and quantitative reasoning skills is not surprising. The early tasks in the quantitative scale are based on completing a series which is laid out spatially. For example, participants with WS may be trying to use spatial skills to complete a series of dominoes by treating dots as individual features of a shape rather than as numerosities. This may be why the individuals with better visuo-spatial skills are more successful. It is also likely, that given their superior verbal abilities, these individuals would try to use language-based strategies to complete the tasks.

The lack of correlations in the data from the DS group may have several bases. In general the verbal abilities of the DS group were poor and there was little variability across the group. It is thus unlikely that this group was using language strategies to solve quantitative problems given that language is weaker than spatial ability in this group. The lack of correlation with spatial skills is interesting, therefore. It suggests that the DS group are not treating the quantitative tasks merely as pattern recognition exercises, as may be happening with the WS group. Instead, they may be relying on a true number-based strategy. Of course, caution should be exercised when interpreting these results because at the early stages, the quantitative reasoning subscale is not particularly numerically based. In addition, this subtest on a standardised scale taps only a limited aspect of number ability. It is thus clear that although our results are suggestive, further research is needed. Indeed, to begin to shed further light on the differences in numerical representations in WS and DS, we have conducted a more in-depth investigation of other number skills, to which we now turn.

Experiment 3: Number processing and calculation battery

In order to gain a wider-ranging picture of number abilities in our two clinical groups, the participants completed a battery of number processing and calculation tasks, comprising two parts: assessment of number knowledge and assessment of arithmetic knowledge.

The assessment of number knowledge consisted of the following tasks:

1. Counting task: participants were required to recite the number sequence from 1 to 20, 25 to 35 and backward from 20 to 1.
2. What comes next/before task: participants were required to say the number that followed or preceded each of 14 numbers spoken aloud by the experimenter.
3. Arabic numeral seriation: 15 series of four stimuli (1 to 3 digit numerals) were presented on cards to be ordered from the smallest to the largest.
4. Dot seriation: 6 series of patterns of dots varying in size and numerosity were presented on cards to

be ordered from the smallest to the largest amounts.

5. Matching dots to Arabic numerals: 20 arrays of dots (10 canonical and 10 in random order) of 1 to 9 numerosity were presented on cards to be matched to the corresponding Arabic numerals. Three alternatives were presented: the correct number, a numerically close number and a numerically far number.
6. Reading Arabic numerals: 22 Arabic numerals (9 single-digit, 11 2-digit and 3 3-digit numerals) were visually presented to be read aloud.

The assessment of arithmetic knowledge consisted of the following tasks:

1. Addition: 25 single-digit problems were presented on single cards to be answered verbally or in writing.
2. Subtraction: 25 single-digit problems were presented on single cards to be answered verbally or in writing.
3. Multiplication: 24 single-digit problems were presented in a multiple-choice task with three alternative answers: the correct one, a table error (e.g., $2 \times 4 = 10$) or a non-table error. Participants were required to select the correct answer.

Unlike Experiments 1 and 2 which focused on approximate magnitude representations of number, the majority of the tasks in the number battery of Experiment 3 tap precise aspects of number which are thought to be verbally mediated.

Method

Participants. The participants in the WS, DS, CA-matched and MA-matched groups were the same as those in Experiment 2, except that in the DS group, 2 individuals refused to complete the battery. The mean chronological age of the remaining participants from the DS group was 26;4 years, range 17;11–35;3. Their mean MA on the BAS II was 5;10 years, range 5;7–6;4. For all other characteristics of the WS, CA and MA controls, see Table 2 from Experiment 2.

Procedure. The battery was administered to each participant individually. A full description of the tasks is provided in Table 4. Where necessary, participants were free to use paper and pencil for answering the arithmetical tasks. They completed the items on their own, but were given encouragement or hints if they did not know what to do. Before each task, examples were given to clarify the requirements. For motivational reasons or time constraints, a shorter version of the Arabic seriation task was occasionally used.

For the CA- and MA-matched groups, if the participant could complete the three most difficult items, then easier items were not tested and their performance was deemed 100% correct. The dot seriation task was administered in the same way. If performance was not 100% correct, all items were presented and individually scored.

Control groups. The majority of the MA- and CA-controls performed at or near ceiling. In particular, the CA-control group performance was error free on almost all tasks, with very rare exceptions in the matching Arabic-to-dot task and in multiplication. In the MA-group, some of the younger participants (below 6 years) displayed more systematic difficulties in both the number processing and calculation tasks. This holds for the most difficult counting

Table 4 Tasks in the number battery

Task	Example
Arabic numeral seriation (15 items)	6, 3, 1, 8,
Put numbers in serial order in order.	20, 6, 9, 38 9, 25, 17, 112
Dot seriation (6 items)	Dots vary in size.
	An example sequence: 5,2,3,1
What comes next/before? (14 items each)	E.g. 1–11–65–17–10–
Experimenter says x, participant says what comes next/before.	
Matching (20 items – 10 random/10 canonical)	Array of 5 random dots.
Point to the Arabic numeral which matches the dot array.	Choice of: 2, 6, 5
One option is correct, one is incorrect but close,	Array of 3 dots in a canonical pattern
the other incorrect but far.	Choice of: 6, 4, 3
Reading Arabic numerals (22 items)	
9 single digit items	4 3
11 double digit items	11 80
3 triple digit items	100 250
Rote counting	forward 1–20 forward 25–35
	backwards 20–1
Addition (25 items possible)	
Presented on cards (read out if necessary) verbal response	2 + 3 6 + 1 8 + 8 (maximum sum 16)
Subtraction (as above – 25 items)	2–1 7–0 9–4
Multiplication	$2 \times 4 =$
Alternative answers provided: 1 correct,1 incorrect in the table,	8 (correct) 6 (in either 2 or 4 times table)
one incorrect not in the table	9 (not in table)

series and the seriation task, where the youngest MA-matched children only managed to put single digits into the correct order. With regard to the arithmetic tasks, the performance of MA-matched controls was rather variable. While the addition and subtraction tasks were solved by most of the oldest children, at least to their ability level, several individuals did not even attempt the multiplication tasks. This is in accordance with the levels of the UK National Curriculum, in that this younger age group would not yet be expected to master the multiplication tables.

Williams syndrome and Down's syndrome groups. As would be expected, the performance of the WS and DS groups was poorer than that of their typically developing counterparts, but interestingly there were marked differences between the two clinical groups. Overall, the WS group displayed considerably more difficulty with the number processing and calculation tasks than the DS group. Because some of the tasks were not completed by all individuals, this prevented a full statistical analysis. However, we first summarise the results in Figure 5 and then discuss them task by task, comparing the errors in the WS and DS groups qualitatively.

In the counting task, only the most automatic sequence (counting forward from 1 to 20) was completed successfully by both the WS and DS groups. By contrast, in counting forward from 25 to 35 and counting backward, the WS group showed considerable difficulties compared to the DS group. Whereas all DS participants performed flawlessly, only 4 WS individuals were able to count from 25 to 35. Similarly, whereas 66% of DS participants could count backwards, only a single participant with WS was able to do so successfully.

Similar difficulties in accessing the number sequence emerged in the 'what comes next/before' task. In the 'what comes next' task, the WS group performed worse than the DS group (t (12) $p = 2.24$, $p < .01$). In the 'what comes before' task, both groups displayed some difficulties. However, the low mean score for the DS group (67% correct) was due to two individuals refusing to complete the task. By contrast, the errors of the WS group reflected their broader problems in accessing the number sequence; the majority of the difficulties were with 2-digit numbers and some errors involved saying what came before in a 'next' task or vice versa (41.9% of errors). Other WS errors consisted in shifting from counting one by one to counting by tens (e.g., $40 \rightarrow 50$), and more frequently they resulted from skipping numbers in the sequence (e.g., from $11 \rightarrow 20$, $91 \rightarrow 89$). Individuals in the WS group failed to preserve the class of the number in 19.4% of cases, for example skipping from the teens to tens. By contrast, the DS errors always preserved the class identity of the number. This indicates a more clearly specified representation of number magnitude in the DS group. All the errors of the DS group involved making a lexical substitution, for example 34–38.

Participants with WS also made errors when reading Arabic numerals aloud, whereas the DS group performed at ceiling with both single and multi-digit numbers. The WS group was error free only with single-digit numbers, their performance dropping to 87.5% on the simple task of reading aloud the multi-digit numbers. Participants with WS made various types of error. Some again involved changing the class of a number, for example reading 17 as 27 or 82 as 28 (23%). This latter subclass of inversion errors were common for the WS group, but not for the DS. Furthermore, unlike both the control groups and the DS group, individuals with WS also made syntactic errors (38.5% of errors). For example, they read 250 as 2500. The other main errors made by the WS group were in-class errors where, al-

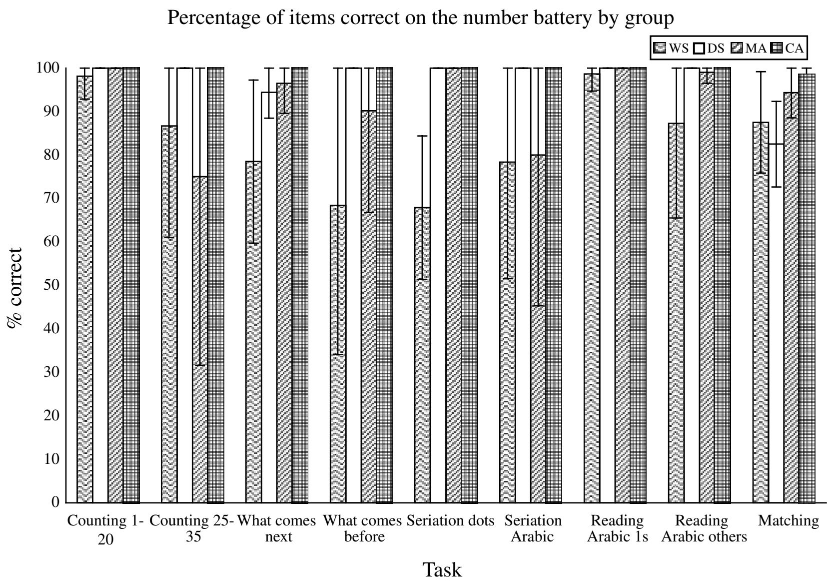

Figure 5 Percentage of items correct on the number battery by group

though the response remained within the same class, the lexical elements were incorrect, e.g., 9 read as 6, and 57 read as 27 (23% of errors).

In both seriation tasks, the WS group's performance was also significantly less accurate than the DS group's, for dots: t (12) $p = 4.71$, $p < .01$, and for Arabic numerals: t (12) $= 1.95$, $p < .01$. Again, these results point to problems for older children and adults with WS with both the written number sequence and the sequence of numerosities that this represents.

Matching numerosities to Arabic numerals was a difficult task not only for the WS group (87.5% correct) but also, for the first time, for the DS group (dropping from 100% to 82.5% correct).

There were only a limited number of results from the arithmetic tasks, due to the fact that many of the participants with WS and DS could not complete a full set of problems. Unsurprisingly, both groups seemed more knowledgeable about addition and subtraction than multiplication, but those in the WS group had a weaker knowledge of number bonds and resorted to using their fingers and making dots on the page more frequently than the DS group. Two individuals with WS did not even seem to understand the underlying concepts of addition and subtraction. They had to be told by the experimenter how these operations worked. Their attempts were not scored.

Discussion: Experiment 3

In summary, despite the fact that many of the tasks in the number battery appear to tap verbal aspects of the number system, e.g., the number sequence and arithmetic facts, the performance of the DS group again outstrips that of the WS group. Given the much poorer language abilities of the DS participants, it is clear that these number skills do not rely solely on proficient verbal skills.

The data from the number battery highlight different patterns of errors in the two atypical groups. Many of the WS group had problems with the 'What comes before?' task. This could be for two reasons: (1) they had difficulty understanding the instructions and often offered the next number in the sequence, even when reminded not to, and (2) because they could not resort for this task to their usual rote strategy for forward counting. Their problems with 2-digit numbers, skipping 40 to 50 and swapping between number classes, e.g., from the ones to the teens, 8–11, suggest that their knowledge of the rules governing the counting sequence is poorly specified and particularly vulnerable at class boundaries. Individuals with WS also made errors that changed the syntax of numbers, both in the reading task and in the 'what comes next/before' task. This was due either to an inversion problem, for example from 14 to 41, or to putting elements of a number in the wrong place, e.g., reading 250 as 2500. It is important to note that not a single error of this type emerged in any of the responses of the entire DS group. Instead, their errors involved lexical substitutions, e.g., 17 for 19, thereby preserving the approximate size of the correct number in their responses. By contrast, the responses of the WS group were less constrained by size, reflecting weaker numerosity representations and less constraint from the numerosities underlying numerical symbols during number processing.

The WS problem with reading numbers involves poor transcoding, i.e., the switch from written Arabic codes to spoken numbers. This conversion between the Arabic code and the phonological speech code may be achieved in two different ways. The first involves two steps: (1) conversion from Arabic code to numerosity and (2) conversion from the numerosity to phonological output. A second possible mechanism proposed by Dehaene (1997), as well as by Cipolotti & Butterworth (1995), involves an asemantic route between Arabic input to phonological output, without recourse to numerosity representations. This is more likely to be the way in which individuals with WS deal with the transcoding problem, thereby explaining why their responses are not constrained by the semantics of numerosity. Interestingly, despite impressive vocabulary scores, weak semantics also characterise the language and reading of people with WS (Laing, Hulme, Grant, & Karmiloff-Smith, 2001; Thomas et al., 2002). In sum, the WS problems in the transcoding process differ from those in individuals with DS, who exhibit very few syntactic or out of class errors.

In the typically developing groups, very few errors occurred. Those that did arise were as expected. It is obviously easy to make errors when judging numerosity of arrays of 9 or 10 dots randomly arranged. In addition, some of the very youngest controls had difficulty with tasks involving multi-digit numbers because these were beyond their cognitive level.

The data from the simple operations like addition and subtraction are somewhat more difficult to classify. The lack of an equal number of trials for all the groups was an unavoidable outcome of this part of the study. Because of the greater complexity of the task, it was often difficult to motivate the clinical participants to attempt the arithmetic tasks, especially at the end of a long test session. These were also the items that all participants found most challenging, and people with learning difficulties often have avoidance strategies for tasks they know that they will find difficult. Despite this being true for all groups, again the WS participants performed worse than the others and displayed little awareness of the calculation process.

Performance across tasks. In order to investigate consistency of ability in the number domain, the performance of individuals with WS and DS on the dot comparison tasks and the number battery was compared. No correlations were found between performance on the SDE tasks, as measured by the difference in RT between close and far pairs, and performance on the number battery. (See Table 5 for the individual scores of each participant.) However, the correlation between the performance of the DS group on matching dots to Arabic numerals and their SDE, approached significance, $r = .753$, $p = .08$. This hints at a link between the ability to discriminate between numerosities and the more advanced ability to make a correspondence between quantities and their Arabic numeral labels in this clinical group which is absent in the WS group.

General discussion

The present series of experiments offers a more thorough examination of numerical abilities than has hitherto been reported in the literature on Williams syndrome. In addition, our comparison between the performance of CA- and MA-matched participants with Down's syndrome pinpoints syndrome-specific differences in the number abilities of the two clinical groups, despite similar overall levels of cognitive functioning. Our study also addresses the role that language may play in numerical cognition by comparing the results of two clinical populations who differ significantly in their language abilities. Finally, we covered a wide developmental trajectory from infancy to adulthood.

The computer-based numerosity comparison task in Experiment 2 is likely to rely upon non-linguistic aspects of numerical cognition. Moreover, the existence of a robust symbolic distance effect is taken in the normal population to be a marker of good analogue magnitude representations. The typically developing participants all exhibited the symbolic distance effect. Results revealed that several individuals with DS clearly displayed the symbolic distance effect, suggesting good analogue magnitude representations in this clinical population. By contrast, a significant effect was not present for the participants with WS, pointing to poor analogue magnitude representations in this group. Our results suggest that while DS older children and adults have a representation of numerosity similar to that of typically developing groups, this does not hold for the WS group. Overall, the considerably better language abilities of the WS group provided no advantage for their performance on this number task. Instead, it is the quality of the numerosity representations of each group that turns out to be critical for performance.

The data from the number battery paint a similar picture of impairment in WS. This is despite the fact that many of the tasks in this battery would seem to call upon verbal skills, which are a relative strength in the WS profile and much better than those in the DS group. Indeed, it was predicted that the DS group

Table 5 Percentage of items correct on the number battery by individual: WS and DS

Number battery item	WS										DS							
	JN	TS	LS	AD	CC	GM	RD	NM	Mean	SD	EB	CM	AB	DT	SH	JP	Mean	SD
Counting 1–20	85	100	100	100	100	100	100	100	98.13	5.30	100	100	100	100	100	100	100	0
Counting 25–35	*	100	upto28	100	upto34	100	100	*			100	100	100	100	100	100	100	0
Backwards 20–1	21.05	100	*	*	*	*	50	*			100	100	50	*	100	100	90	
Next	50	71	64	100	86	100	64	92.86	78.48	18.80	100	100	92.86	100	100	85.71	94.43	5.98
Before	*	100	64	92.8	.00	92.86	50	79	68.38	34.23	100	100	*	*	100	100		
Seriation dots	83	80	50	80	40	60	67	83	67.88	16.52	100	100	100	100	100	100	100	0
Seriation Arabic	80	100	33	100	47	100	67	100	78.38	26.85	100	100	100	100	100	100	100	0
Reading Arabic 1s	89	100	100	100	100	100	100	100	98.63	3.89	100	100	100	100	100	100	100	0
Others	36	100	85	100	85	100	92	100	87.25	21.74	100	100	100	100	100	100	100	0
Matching	100	80	70	100	85	95	75	95	87.5	11.65	90	90	80	65	80	90	82.5	9.87
Addition	*	54.5 (6/11)	dot counting	80	72	88	36	84	69.1	20.08	84	80	*	26.7 (4/15)	90 (9/10)	100 (10/10)	84†	80†
Subtraction	*	88.9 (8/9)	60	84	56	90 (9/10)	24	32	62.1†	27†	76	92	*	37.5 (3/8)	53.8 (7/13)	53.8 (7/13)	76†	92†
Multiplication	*	28.6 (2/7)	*	40 (6/15)	*	81.3 (13/16)	46.2 (6/13)	*	49.0†	22.69†	64	*	*	*	*	*	64†	
SDE		Yes		Yes		Yes	Yes				Yes	Yes	Yes		Yes	Yes		

*Not done.
†Full set of items was not completed by each participant.

would perform worse than the WS group on these tasks than in the number comparison task because of the greater language component in the former. But this was not the case. Although individuals with WS were able to count by rote, at least to 20, problems emerged as soon as they had to manipulate numbers and not merely reproduce a verbal string by rote. This was clear even when they were asked to put numbers in the correct order. Individuals with WS also had difficulty with the syntax of number. These problems with ordering number suggest that the links between representations of magnitude and their lexical and numerical symbols are weak. Individuals with WS also have problems with transcoding.

In sum, the number skills in older children and adults with WS are weaker than those of the DS group, especially on the number battery. It is likely that the performance of individuals with DS is immature, as evidenced by their slow reaction times, but it appears to follow the typical developmental pathway, albeit with delay. By contrast, individuals with WS have much weaker number representations and seem to follow a different developmental pathway with respect to several aspects of number.

There is also a discontinuity in performance between infants and adults with Williams syndrome. Despite the success of WS infants at discriminating between arrays of two and three objects, the adults with WS do poorly on a variety of number tasks and are significantly worse in adulthood than their counterparts with DS who performed very poorly in infancy. This decalage in performance is intriguing and could have a number of causes. One plausible account arises from the fact that older children and adults with WS focus on the features of objects in visual processing tasks (Deruelle, Mancini, Livet, Casse-Perrot, & de Schonen, 1999; Karmiloff-Smith, 1997; Karmiloff-Smith et al., 2004). This featural processing style could allow the infants to succeed at the task by creating object files. By this model, when faced with small numbers up to 4, infants set up an object file for each entity (Huttenlocher, Jordan, & Levine, 1994; Carey, 1998; Simon, 1997; Uller et al., 1999). In order to succeed in the numerosity discrimination task, infants individuate objects in each array, putting each object into a new internal memory file. Then, they compare their representations of the objects in one array with the objects in the other and detect a mismatch. Object files can be used for small numerosities, after which there are too many records to maintain in memory. Although successful in infancy, this limit would preclude adults with WS from relying on this processing strategy because our tasks included numbers greater than four on many occasions. Although the object files approach is seen by some as non-numerical, the ability to parse stimuli into component parts, which can then be given numerical tags, is a useful building block for number and may bootstrap normal number development (Carey, 1998). The atypical developmental trajectory in the WS group may arise from two processing problems. The first has its roots in infancy and involves an over-reliance in adulthood on the object files strategy. The second is the poor specification of the mental number line, i.e., the magnitude representation of numerical displays and the mapping of numerical symbols onto these displays. The pattern seen in Down's syndrome suggests that number development in this group is delayed but not fundamentally different from that of typically developing groups.

The results from our experiments highlight the importance of exploring domains in depth rather than merely noting overall delay when studying atypically developing groups. Our study also shows how 'normal behaviour', such as reciting the correct counting sequence, can be underpinned by deviant cognitive processes (see discussion in Karmiloff-Smith, 1998). Despite similar overall cognitive abilities, our two clinical groups performed differently within the number domain. Previous research has suggested that number skills in Down's syndrome are impaired, but the present study highlights the fact that this impairment is not as great as in another syndrome, Williams syndrome, despite an equivalent level of general impairment. The fact that both groups have similar IQs suggests that number difficulties cannot be attributed solely to cognitive impairment in these groups. Our study also indicates that language may play a different role in numerical cognition in these clinical groups than in the normal population. Once again, the importance of a thorough investigation of atypical development is highlighted, stressing the need to differentiate between serious delay (as in the DS infant group) and deviant cognitive processing (as likely in the WS group).

A number of theorists have suggested that the cognitive profile present in older children and adults with WS can be used to support the existence of innately specified, independently functioning, cognitive modules which are either intact or impaired. (Bellugi et al., 1988; Pinker, 1999). If, in the number domain, initial states could be simply assumed from endstates, as such claims based on endstates often imply, then WS infants should perform poorly on the infant measure of number skills and the DS infants should perform considerably better. However, the data from the present studies fail to support this logic. In fact, the opposite pattern holds, with the WS group doing poorly in adulthood and better in infancy whereas the converse is true for the DS group. Our results highlight the importance of tracing the developmental trajectories of the two clinical groups back to the origins in infancy to determine their number skills. In the Down's Syndrome group, problems with number appear to reside in more low-level processes that are seriously delayed very early in development. The delay in these precursors to number is similar to the DS delay in language. Thus, timing is a critical variable that leads to their difficulties. By contrast, in Williams syndrome at least one of the foundations of number processing for small numerosities is functioning in infancy. The problems in the WS group lie further along the developmental trajectory, perhaps for judging large quantities of dots, as Xu has studied in healthy infants (Xu & Spelke, 2000). Above all, the WS differences seem to reside in building *precise* numerical representations. The cross-syndrome differences pinpointed by our study have crucial implications for syndrome-specific intervention programmes.

Acknowledgements

We would like to thank all the infants, children, adults, and their families for their help in this research, and the staff and pupils at Gateway County Primary School, Carterton. Thanks also to the Williams Syndrome Foundation, UK, and the Down's Syndrome Association for putting us in touch with the families. This research was supported by the Down's Syndrome Association, the Williams Syndrome Foundation, MRC Programme Grant No G9715462, and NIH grant R2TW06761-01 to AK-S.

Correspondence to

A. Karmiloff-Smith, Neurocognitive Development Unit, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, UK; Tel: 020 7905 2334; Fax: 020 7242 7717; Email: a.karmiloff-smith@ich.ucl.ac.uk

References

Antell, S.E., & Keating, D.P. (1983). Perception of numerical invariance in neonates. *Child Development*, *54*, 695–701.

Arnold, R., Yule, W., & Martin, N. (1985). The psychological characteristics of infantile hypercalcaemia: A preliminary investigation. *Developmental Medicine and Child Neurology*, *27*, 49–59.

Bayley, N. (1993). *Bayley Scales of Infant Development* (2nd edn). San Antonio, TX: Psychological Corporation.

Bellugi, U., Bihrle, A., Jernigan, T., Trauner, D., & Doherty, S. (1990). Neuropsychological, neurological, and neuroanatomical profile of Williams syndrome. *American Journal of Medical Genetics*, *6*, 115–125.

Bellugi, U., Marks, S., Bihrle, A.M., & Sabo, H. (1988). Dissociation between language and cognitive functions in Williams syndrome. In D. Bishop & K. Mogford (Eds.), *Language development in exceptional circumstances* (pp. 177–189). London: Churchill Livingstone.

Bishop, D.V.M. (1983). *Test for the reception of grammar*. Newcastle-upon-Tyne: University of Newcastle-upon-Tyne.

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*, 1037–1046.

Butterworth, B. (1999). *The mathematical brain*. London: Macmillan.

Butterworth, B., Zorzi, M., Girelli, L., & Jonckheere, A.R. (2001). Storage and retrieval of addition facts: The role of number comparison. *Quarterly Journal of Experimental Psychology*, *54A*, 1005–1029.

Cappelletti, M., Butterworth, B., & Kopelman, M. (2001). Spared numerical abilities in a case of semantic dementia. *Neuropsychologia*, *39*, 1224–1239.

Carey, S. (1998). Knowledge of number: Its evolution and ontogeny. *Science*, *282*, 641–642.

Caycho, L., Gunn, P., & Siegal, M. (1991). Counting by children with Down's syndrome. *American Journal on Mental Retardation*, *95*, 575–583.

Cipolotti, L., Butterworth, B., & Denes, G. (1991). A specific deficit for numbers in a case of dense acalculia. *Brain*, *114*, 2619–2637.

Cipolotti, L., & Butterworth, B. (1995). Toward multiple model of number processing: Impaired number transcoding with preserved calculation skills. *Journal of Experimental Psychology: General*, *124*, 375–390.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. Oxford: OUP.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*, 371–396.

Dehaene, S., & Cohen, L. (1994). Dissociable mechanisms of subitizing and counting: Neuropsychological evidence from simultanagnosis patients. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 958–975.

Dehaene, S., Dupoux, E., & Mehler. J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 626–641.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*, 970–974.

Deruelle, C., Mancini, J., Livet, J.O., Casse-Perrot, C., & de Schonen, S. (1999). Configural and local processing of faces in children with Williams syndrome. *Brain and Cognition*, *41*, 276–298.

Donnai, D., & Karmiloff-Smith, A. (2000). Williams syndrome: From genotype through to the cognitive phenotype. *American Journal of Medical Genetics: Seminars in Medical Genetics*, *97*, 164–171.

Duncan, E.M., & McFarland, C.E. (1980). Isolating the effects of symbolic distance and semantic congruity in comparative judgements: An additive factors analysis. *Memory and Cognition*, *8*, 612–622.

Elliott, C.D., Smith, P., & McCulloch, K. (1996). *British Ability Scales II*. Windsor, Berkshire: NFER-Nelson.

Fagan, J.F. (1970). Memory in the infant. *Journal of Experimental Child Psychology*, *9* (2), 217–226.

Haaf, R.A., Brewster, M., de Saint Victor, C.M., & Smith, P.H. (1989). Observer accuracy and observer agreement in the measurement of visual fixation with fixed-trial procedures. *Infant Behaviour and Development*, *12*, 211–220.

Hughes, S.D. (1995). *Within-domain dissociation in Williams syndrome: The case of number*. Unpublished undergraduate project, University College, London.

Huttenlocher, J., Jordan, N., & Levine, S. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology: General*, *123*, 284–296.

Karmiloff-Smith, A. (1997). Crucial differences between developmental cognitive neuroscience and adult neuropsychology. *Developmental Neuropsychology*, *13*, 513–524.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, *2*, 389–398.

Karmiloff-Smith, A., Brown, J.H., Grice, S., & Paterson, S.J. (2002). Dethroning the myth: Cognitive dissociations and innate modularity in Williams syndrome. *Developmental Neuropsychology*, *23*, 227–242.

Karmiloff-Smith, A. Scerif, G., & Ansari, D. (2003). Double dissociations in developmental disorders? Theoretically misconceived, empirically dubious. *Cortex*, *39*, 161–163.

Karmiloff-Smith, A., Thomas, M., Annaz, D., Humphreys, K., Ewing, S., Grice, S., Brace, N., van Duuren, M., Pike, G. & Campbell, R. (2004). Exploring the Williams syndrome face processing debate: The importance of building developmental trajectories. *Journal of Child Psychology and Psychiatry*, *45*, 1258–1274.

Kaufman, A.S., & Kaufman, N.L. (1983). *Kaufman Assessment Battery for Children (K-Abc): Administration and scoring manual*. Circle Pines, MN: American Guidance Services.

Kinnear, P.R., & Gray, C.D. (1999). *SPSS for Windows made simple*. Hove: Psychology Press.

Laing, E., Hulme, C., Grant, J., & Karmiloff-Smith, A. (2001). Learning to read in Williams syndrome: Looking beneath the surface of atypical reading development. *Journal of Child Psychology and Psychiatry*, *42*, 729–739.

Moyer, R.S., & Landauer, T.K. (1973). Determinants of reaction time for digit inequality judgments. *Bulletin of the Psychonomic Society*, *1*, 167–168.

Nye, J., Clibbens, J., & Bird, G. (1995). Numerical ability, general ability and language in children with Down's syndrome. *Down's Syndrome: Research and Practice*, *3*, 92–103.

O'Neill, J.M., Jacobson, S.W., & Jacobson, J.L. (1994). Evidence of observer reliability for the Fagan Test of Infant Intelligence (FTII). *Infant Behaviour and Development*, *17*, 465–469.

Paterson, S.J., Brown, J.H., Gsödl, M.K., Johnson, M.H., & Karmiloff-Smith, A. (1999). Cognitive modularity and genetic disorders. *Science*, *286*, 2355–2358.

Pinker, S. (1999). *How the mind works*. London: Weidenfeld & Nicolson.

Rosin, M.M., Swift, E., Bless, D., & Vetter, D. K. (1988). Communication profiles of adolescents with Down syndrome. *Journal of Childhood Communication Disorders*, *12*, 49–64.

Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequality. *Child Development*, *48*, 630–633.

Simon, T.J. (1997). Reconceptualizing the origins of number knowledge: A 'non-numerical' account. *Cognitive Development*, *12*, 349–372.

Strauss, M.S., & Curtis, L.E. (1981). Infant perception of numerosity. *Child Development*, *52*, 1146–1152.

Starkey, P., Spelke, E.S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition*, *36*, 97–127.

Thomas, M.S.C., Dockrell, J.E., Messer, D., Parmigiani, C., Ansari, D., & Karmiloff-Smith, A. (2002). *Naming in Williams syndrome*. Manuscript submitted for publication.

Udwin, O., Davies, M., & Howlin, P. (1996). A longitudinal study of cognitive abilities and educational attainment in Williams syndrome. *Developmental Medicine and Child Neurology*, *38*, 1020–1029.

Washburn, D.A., & Rumbaugh, D.M. (1991). Ordinal judgments of numerical symbols by macaques (Macaca mulatta). *Psychological Science*, *2*, 190–193.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*, 155–193.

Xu, F., & Spelke, E. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, 1–11.

Manuscript accepted 2 December 2004