SPECIAL ISSUE: ORIGINAL ARTICLE

TYPICAL AND ATYPICAL DEVELOPMENT OF VISUAL ESTIMATION ABILITIES

Daniel Ansari[1,2], Chris Donlan[3] and Annette Karmiloff-Smith[4]

([1]Neurocognitive Development Unit, Institute of Child Health, London, UK; [2]Numerical Cognition Laboratory, Department of Psychology, University of Western Ontario, London, Ontario, Canada; [3]Department of Human Communication Sciences, University College London, London, UK; [4]Developmental Neurocognition Laboratory, Birkbeck College, University of London, London, UK)

ABSTRACT

Despite the fact that developmental impairments of number skills are common, they remain sparsely investigated. We explored low-level numerical representations and their developmental trajectory in a developmental disorder, Williams syndrome (WS). Groups of WS and typically developing (TD) individuals estimated rapidly-presented arrays of 5, 7, 9, and 11 dots. In comparison to the normal developmental trajectory, the ontogenesis of estimation skills in WS is both delayed and deviant. Whereas TD children's estimations became significantly more accurate and less variable over developmental time, only marginal developmental changes in estimation ability emerged across age in the WS groups. Our data highlight the importance of considering developmental changes in low-level components of numerical cognition in atypical development while at the same time emphasizing the importance of paying closer attention to quantitative changes and their functional role in typical development.

Keywords: numerical cognition, Williams syndrome, estimation, developmental disorders

INTRODUCTION

In recent years there has been a growing interest in the origins and neural correlates of numerical cognition (Dehaene, 1997; Butterworth, 1999). Research into how animals as well as human infants, children, and adults represent number has revealed that numerical skills are not merely the product of cultural transmission but that some aspects of numerical cognition have a long evolutionary history as well as a characteristic developmental trajectory and neural instantiation (Feigenson et al., 2004; Gallistel and Gelman, 2000; Dehaene et al., 1998). Evidence suggests that there is a substantial degree of qualitative similarity in the way in which animals, infants, children, and adults represent numerical magnitudes.

Despite the fact that developmental impairments of numerical ability are very common (Badian, 1983; Gross-Tsur et al., 1996; Shalev and Gross-Tsur, 2001) and have a significant impact on children's lives, the rich body of evidence for low-level processing abilities in typically developing (TD) infants, children, and adults has not been exploited to guide the study of how children develop impairments of numerical cognition. In contrast, in the domain of reading the discovery of the importance of low-level processing abilities in the auditory domain has contributed significantly to the understanding and remediation of higher-level reading impairments or dyslexia (Goswami et al., 2002; Tallal, 1980).

We therefore deem it vital for research on atypical number development to focus on low-level processing competencies in an attempt to uncover the reasons for subsequent impairments in numerical cognition. There are several well-replicated paradigms that have been used to describe low-level, non-verbal numerical magnitude representations. These have yielded effects that reveal the characteristics of the representations underlying the processing of numerical magnitudes.

Distance and Size Effect

It has been repeatedly found that when adults judge which of two numbers is numerically larger, their responses are more accurate and faster the further the numerical distance between the numbers (Moyer and Landauer, 1967; Dehaene, et al., 1990). This is known as the distance effect. The developmental origins of the distance effect can be traced back to infancy. Recent evidence from human infants suggests that they can discriminate between 8 *versus* 16 dots but not between 8 *versus* 12 dots (Xu and Spelke, 2000), with the same being true for comparisons of identical numerosities of sounds (Lipton and Spelke, 2003). These data from human infants carry the signature of the distance effect. Animal models are also relevant to the current debate. For example, Brannon and Terrace (1998), found that when rhesus monkeys were trained to order numerical displays, the accuracy with which animals were able to make ordinality judgments was positively related to the distance between the numerosities to be ordered.

Importantly, in the human studies, it was found that the effect of distance on reaction time and accuracy decreases over developmental time indicating that noise in magnitude representations and the overlap between represented magnitudes decreases, leading to faster and more accurate performance (Sekuler and Mierkiewicz, 1977). In addition to the distance effect, numerical magnitude comparison data reveal that performance is highly sensitive to the ratio of the magnitude (Weber's Law). Thus, even if distance is held constant, reaction time increases and accuracy decreases as the relative size of the magnitudes that are to be compared is increased. This is called the size effect. Distance and size effects are both found in data involving magnitude comparisons. In fact, there is a high correlation between distance and size as predictors of dependent measures in numerical magnitude comparison tasks. Recent neuroimaging work by Piazza et al. (2004) has investigated the neural basis of low-level magnitude representation. In a functional magnetic resonance imaging (fMRI) study, participants were presented with a stream of dot arrays and were not told what the purpose of the study was. Repeated numerosities (for example 8) were interspersed with displays deviant numerosities (such as 16 or 32). The results revealed not only that the horizontal segment of the bilateral interparietal sulci (HIPS) responds to numerosity change, but that the greater the deviance, the larger the fMRI response. This neuroimaging evidence suggests that the HIPS may contain the internal representation of non-verbal numerical magnitude that gives rise to the distance and size effects.

Estimation Variability

Numerical magnitude comparison tasks are a powerful way to tap into the characteristics underlying representations of numerical magnitude. However, another rich source of information about the features of magnitude representations and their development can be obtained by the estimation method: here, participants are asked to estimate numerosities presented either in parallel or sequentially, through the visual or auditory modalities. Early work with rats, for instance, indicated that when these animals were trained to press a lever a certain number of times, the variability in the number of their lever presses turned out to be proportional to the number of required presses (Mechner, 1958; Meck and Church, 1983; Platt and Johnson, 1971). In other words, the greater the magnitude to be estimated, the greater the variability in the estimates performed by the animals.

This positive relationship between the size of the numerical magnitude and the degree of variability in the estimates also holds for humans. Whalen et al. (1999) reported that the variability in adults' estimation of number of rapidly presented flashes increases with the number of presented flashes. Similarly, when young children judge the number of rapidly presented dots on a computer screen, their responses become less accurate and more variable as the number increases, with variability and accuracy improving over developmental time (Huntley-Fenner, 2001). Together with the data from estimation studies, the distance and size effects reveal similar features of basic magnitude representations. From the distance effect it can be inferred that the features of numerical magnitude representations that are close together overlap more than those that are far apart. The size effect reveals that variability in representations increases in proportion to the size of the represented magnitudes. However, it is only through estimation data that the noise in representations can be quantified rather than inferred indirectly. By recording trial-by-trial differences in estimates, a direct insight into the response variability can be gained. Hence, the use of an estimation paradigm provides a deeper insight into the features of magnitude representations and their associated cognitive processes.

The Importance of a Developmental Perspective

Recent advances have provided substantial insights in the parameters of basic number representations, their developmental origins in infancy and their neural instantiations. This has led to an emphasis on the *qualitative* similarity in representational features that are revealed across species and levels of analysis. Notwithstanding, there has been little investigation into the full developmental trajectories of these basic mechanisms, the quantitative changes that occur over developmental time the consequence that these ontogenetic changes might have for both typical and atypical development of numerical cognition. In fact, much of the literature is primarily concerned with establishing the existence of a number sense, whose defining features exhibit phylogenetic and ontogenetic continuity. In this paper, we explore the developmental changes in a basic number processing skill, visual estimation, and ask whether differences exist in the development of this basic competency between typically and atypically developing individuals.

Developmental Impairments of Number

In stark contrast to the ever-growing literature on basic number representation in TD individuals, surprisingly little work has targeted differences in low-level features of numerical representations in children who develop impairments of numerical cognition. Rather, the emphasis has mainly been on higher-level, culturally-mediated numerical abilities

TABLE I

Participant background data means represent age in years and months

	TD Groups (N = 63)								Clinical Group (N = 31)			
Group	4-5 Year olds (N = 15)		6-7 Year olds (N = 22)		9-10 Year olds (N = 14)		Adults (N = 12)		WS Children (N = 18)		WS Adults (N = 13)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Chronological age	4.8	.4	6.9	.6	9.5	.5	30.8	9.2	9.7	1.9	28.9	10.5
Pattern construction age equivalent	5.1	1.3	7.4	1.8	9.5	.5	n/a	n/a	4.9	1.6	6.3	2.2

Note. Typically developing (TD) adults were not tested on the Pattern Construction test, as these participants were above the ceiling age of the test. Corresponding cells are marked not applicable (n/a).

such as addition and subtraction as well as on the contribution of other cognitive capacities including attention, speed of processing, and working memory (Geary, 1993; Jordan et al., 2002; Siegler and Jenkins, 1989). This absence of data raises a number of questions: Might it be the case that children who develop difficulties with high-level number concepts and skills actually present with impairments of low-level, non-verbal representations of numerical magnitude? Are developmental impairments of number characterized by difficulties in mapping between numerals and number words and the mental magnitudes that represent them? Addressing such questions is not only important for advancing our understanding of how numerical cognition goes awry, but also has crucial implications for mathematical education and remediation (Griffin and Case, 1999; Gersten and Chard, 1999).

Case Study Disorder: Williams Syndrome

We addressed the above questions concerning the atypical and typical developmental trajectories of basic representations of number by investigating the development of estimation abilities in children and adults with the genetic developmental disorder, Williams syndrome (WS), as well as comparison groups of TD children and adults. Severe deficits of numerical cognition have previously been reported in WS (Ansari et al., 2003; Udwin et al., 1996; Paterson et al., 1999, 2006), yet their developmental origins have not, to date, been thoroughly investigated. WS is a rare genetic disorder occurring in 1 in 20,000 live births (Beuren, 1972) which is caused by a hemizygous sub-microscopic deletion of some 24 genes on chromosome 7q11.23 (Donnai and Karmiloff-Smith, 2000). While individuals with WS are relatively proficient in their language skills, they are strongly impaired in non-verbal cognitive abilities, in particular visuo-spatial cognition. This cognitive profile led to our hypothesis that weaknesses in the domain of numerical cognition may be particularly pronounced for non-verbal representations of numerical magnitude.

We examined low-level numerical abilities in WS by means of a visual estimation paradigm, closely modeled on existing paradigms (Whalen et al., 1999; Huntley-Fenner, 2001). We assessed both the variability and the accuracy in participants' numerical estimations of rapidly-presented visual arrays containing different numbers of dots. Specifically, we explored the cognitive processes underlying the mapping from mental magnitude to numerals. Most studies that compare clinical and control groups focus on comparisons of absolute levels of performance. Since our theoretical focus is on the development of estimation abilities, we sought a way to operationalize and test developmental differences statistically. In other words, we adopted a developmental perspective and methodology. Instead of comparing absolute levels of performance between the experimental and matched groups, we compared and contrasted the full developmental trajectory of estimation abilities between different age groups of TD individuals and those with WS. Thus, we compared how change in task performance over time differs between WS and TD groups, focusing on the differences and similarities of developmental changes in estimation abilities. We compared both qualitative and quantitative changes.

METHODS

Participants

We examined the estimation abilities of 3 groups of TD (n = 63) individuals and 2 WS groups (n = 31). The number of participants per group, their chronological ages, and mental age equivalent scores on the Pattern Construction Subtest of the British Abilities Scales (Elliot et al., 1996), a standardized test of visuo-spatial ability, can be found in Table I.

Tasks and Procedure

In a set of pre-tasks, participants were asked to count aloud from 1 to 20. Next, they were presented with displays of numbers 1-20 in random order on a 15-inch laptop computer screen. Stimuli were presented for 250 milliseconds each and participants were asked to name the number they saw. In addition to assessing whether individuals

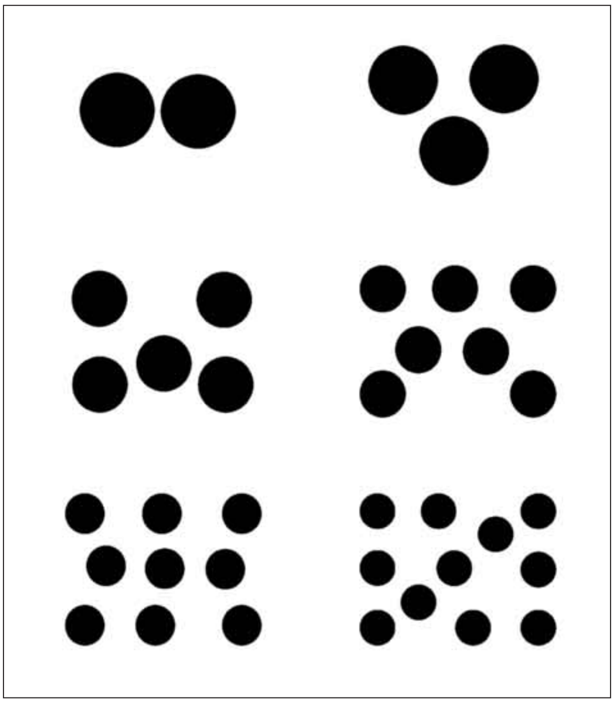

Fig. 1 – *Examples of dot stimuli.*

were able to read Arabic numerals, this pre-task also tests the extent to which all groups tested could process and respond to information that is presented for only 250 msec. This is therefore a control for our experimental task, in which the same stimulus presentation time was used.

Following the pre-tasks, participants were told that they were now going to play a different number game and that, instead of Arabic numerals, they would see different numbers of dots appearing very briefly on the screen (for 250 msec each). Each numerosity (5, 7, 9, and 11) was presented 20 times with order of presentation randomized. In addition, participants were presented with 5 displays each of 2 or 3 dots, randomly interspersed with the other displays, to get an estimate of their ability to rapidly enumerate a small number of objects (‘subitizing’) and to verify that children and adults were able to process visual dot stimuli that were presented for only 250 msec. Stimuli were designed in such a way that the total area occupied by the dots was equivalent across displays in order to avoid template matching and the use of total dot area as a cue to numerosity. Furthermore, the spatial arrangement of dots was randomly varied across all displays, again to control for pattern recognition and density cues to numerosity. Examples of the displays are illustrated in Figure 1.

The experimenter told the participants that he did not want them to count, but instead to quickly estimate or guess how many dots were presented each time. Participants were told that the game was not about getting it right or wrong, but rather, getting the best guess. The experimenter repeated questions such as “How many dots were there?” and “How many dots did it feel like?” throughout the experiment. The experimenter asked these questions immediately after dots had been displayed to ensure that participants responded quickly and did not count. Following the brief appearance of dot displays, there was a blank white display. After participants had responded and before moving onto the next presentation of a dot array, the experimenter presented the participants with a rewarding display of a cartoon. This served both as entertainment for the participants and minimized the likelihood that participants were still processing or representing an afterimage of the preceding trial carried over to the subsequent trial. Participants were given the choice between simply saying the number aloud or by pointing to the estimated number on a number line (1-20) mounted underneath the laptop screen in front of them.

RESULTS

Pre-Tasks

Participants in all groups (children and adults, typical and atypical groups) were able to count aloud from 1 to 20. Furthermore, participants in most groups were able to name the Arabic numerals presented on a computer screen for 250 msec. The exceptions were the group of typical 4-5 year olds and the group of children with WS, where 8/15 and 12/15 participants, respectively, had difficulties naming numerals 11-20, but succeeded on all numbers up to 10.

Experimental Task

Accuracy was calculated by computing the number of trials on which participants estimated correctly. In order to capture the variability in participants’ responses, a coefficient of variation (COV) was computed by dividing the standard deviation of participants’ estimate by the mean of their estimates.

Subitizing

Accuracy for estimations of arrays of 1, 2, and 3 dots was 100% in all the groups tested. Given this ceiling level of accuracy in subitizing ability across groups, no analysis of variability was undertaken. Importantly, this finding suggests that all groups were able to visually process stimuli presented for 250 msec.

Were Participants Estimating Number?

The mean estimates were found to be linearly related to the target numerosities for all groups. For all groups of TD individuals the linear fit (as indicated by the Pearson’s correlation coefficient, r)

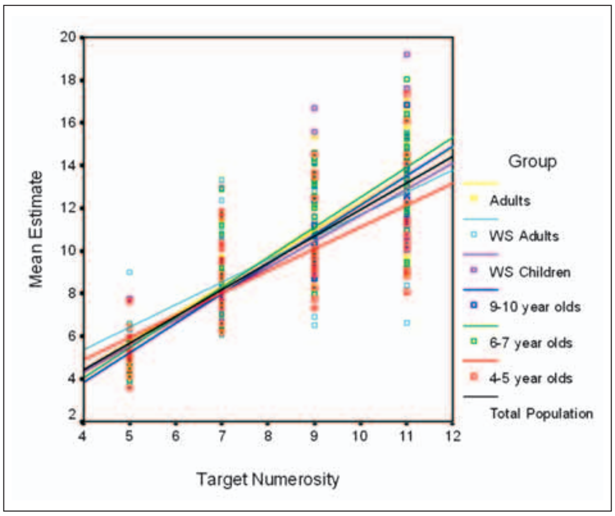

Fig. 2 – *Relationship between mean estimates and target numbers for all groups.*

was close to 1 and the intercept close to 0: 4-5 year olds, $r^2 = .76$, $\beta_0 = .68$; 6-7 year olds, $r^2 = .87$, $\beta_0 = -1.60$; 9-10 year olds: $r^2 = .94$, $\beta_0 = -1.64$; Adults: $r^2 = .96$, $\beta_0 = -.47$. The same held true for the WS groups. Children with WS: $r^2 = .70$, $\beta_0 = .68$; adults with WS: $r^2 = .80$, $\beta_0 = -.75$. All correlations were found to be significant at $p < .001$ (see Figure 2).

Developmental Changes in Estimation Accuracy

Overall accuracy changed with age in the TD groups: it was 15% in the group of 4-5 year olds, 27% for the 6-7 year olds, 43% for the 9-10 year olds, and 62% for the group of TD adults. In the group of children with WS, overall accuracy was 11% while the adults with WS gave the correct response 20% of the time. The mean proportion of accurate responses by group and numerosity is illustrated in Figure 3. Inspection of this figure suggests that estimation accuracy for arrays of 5 dots may be increasing at a faster rate than accuracy for 7, 9, and 11 dots. In view of this, the analysis presented below was computed separately with both mean total accuracy (5, 7, 9, and 11 dots) and mean accuracy without 5 dots.

To compare the developmental changes in estimation accuracy between children and adults with WS against the developmental changes observed in the groups of TD individuals, two 2-way analyses of variance (ANOVAs) were computed. The first of these analyses was run with group (WS adults and children and TD children) and development (adults *vs.* children with WS and 4-5 *vs.* 6-7 year old TD children) as the between-subjects variables. A significant main effect of group on mean total accuracy and mean accuracy without 5 was found: F (1, 60) = 5.6, $p < .021$ and F (1, 60) = 8.1, $p < .006$ respectively. The effects of development on mean total accuracy and mean accuracy without 5 were also significant: F (1, 60) = 20.4, $p < .001$ and F (1, 60) = 9.5, $p > .003$, respectively. There were no significant Group × Development interactions for either of the two dependent variables. It thus appears that the developmental changes between children and adults with WS (average chronological age difference between groups: 19 years and 5 months) do not go beyond the small developmental change seen in the two youngest group of TD participants (average chronological age difference between groups: 1 year and 3 months).

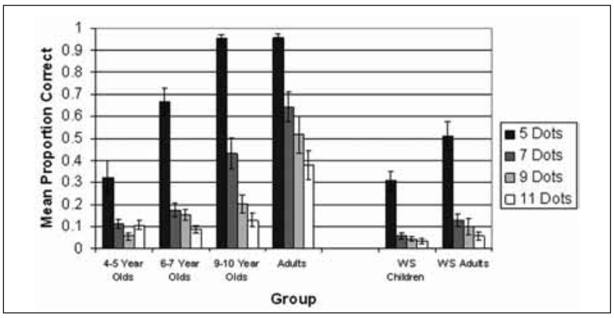

Fig. 3 – *Mean proportion of accurate estimations by group and numerosity. Error bars denote the standard error of the mean.*

In a second analysis, the developmental difference in estimation accuracy between children and adults with WS was contrasted with a wider typical developmental gap: the difference between 4-5 and 9-10 year old children. Again, significant main effects of group on mean total accuracy [F (1, 52) = 28.6, $p < .001$] and mean accuracy without 5 emerged [F (1, 52) = 17.0, $p < .001$]. The main effects of development were also significant for both mean accuracy total [F (1, 52) = 53.8, $p < .001$] and mean accuracy without 5 [F (1, 52) = 18.2, $p < .001$]. Moreover, the Group × Development interaction was also significant: F (1, 52) = 14.1, $p < .001$ for mean total accuracy and F (1, 52) = 4.9, $p < .031$. Inspection of Figures 4a and 4b suggests that the developmental difference for mean total accuracy for all arrays and mean accuracy without 5 was significantly greater between 4-5 and 9-10 year olds than between children and adults with WS.

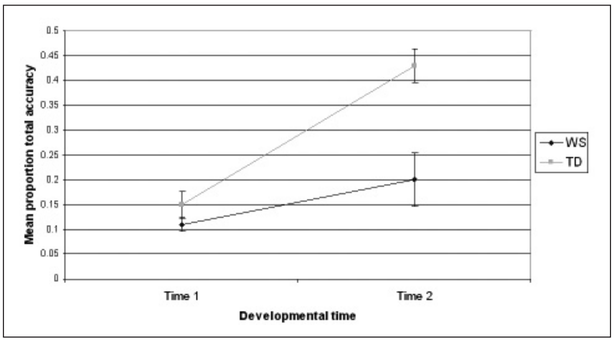

Fig. 4a – *Interaction effect on mean accuracy total of Group (WS vs. controls) and development (4-5 vs. 9-10 year old typically developing children and children vs. adults with WS). Error bars denote the standard error of the mean.*

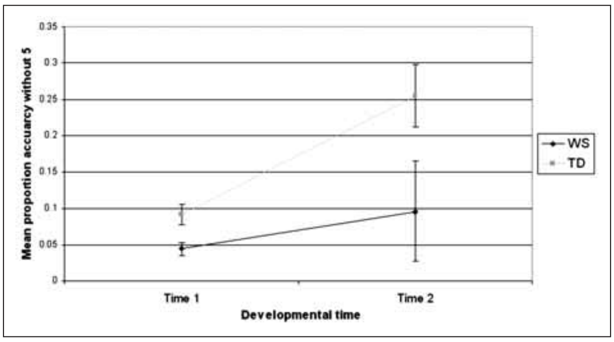

Fig. 4b – *Interaction effect on mean accuracy without 5 of Group (WS vs. controls) and development (4-5 vs. 9-10 year old typically developing children and children vs. adults with WS). Error bars denote the standard error of the mean.*

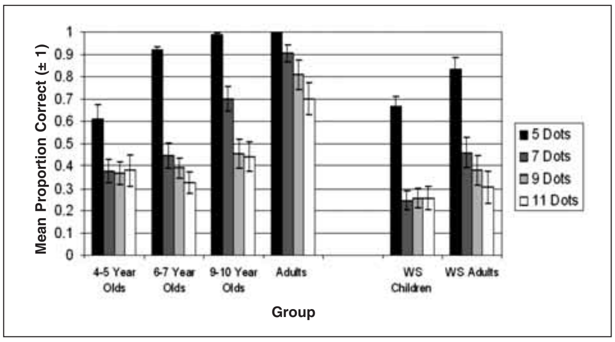

Fig. 5 – *Mean proportion of approximately accurate (± 1 of the target numerosity) estimations by group and numerosity. Error bars denote the standard error of the mean.*

Developmental Changes in Approximate Estimation Accuracy

Given the very low accuracy scores across groups, the extent to which participants' estimations were approximately distributed around the target magnitude was calculated. For this purpose, all responses that were within ± 1 of the target numerosity were scored for each target numerosity. This variable will be referred to as "*mean approximate accuracy*". Mean approximate accuracy was 43% in the group of 4-5 year olds, 52% in the group of 6-7 year olds, 64.5% in the group of 9-10 year olds, and 85% for the group of adults. In the group of children with WS, mean approximate accuracy was 35.5%, and in the group of adults with WS, 49%. The mean proportion of accurate responses (within ± 1 of the target) by group and numerosity is illustrated in Figure 5. The same statistical analyses which were undertaken for the absolute accuracy data (see above) were computed for the approximate accuracy data. In a first analysis, group (WS and TD) and development (adults *vs.* children with WS and 4-5 *vs.* 6-7 year old TD children) were entered as between-subjects variables and mean estimation accuracy approximate as the dependent variable. No significant main effect of group of mean total approximate accuracy and mean approximate accuracy without 5 was found. A significant effect of development was found for mean approximate accuracy [F (1,60) = 8.7, p < .005], but no significant main effect of development on mean approximate accuracy without 5 was revealed. Neither interactions of Group × Development of the two independent variables were revealed. In a subsequent analysis, the difference between WS and TD groups was compared by contrasting the developmental differences between children and adults with WS with cross-sectional differences between the 4-5 and 9-10 year old TD children. This analysis revealed a significant effect of group on both mean approximate accuracy [F (1, 60) = 8.2, p < .006] and mean approximate accuracy without 5 [F (1, 60) = 8.6, p < .005]. Furthermore, significant effects of development were found for both independent variables: F (1, 60) = 19.1, p < .001 and F (1, 60) = 9.4, p < .003 for mean approximate accuracy and mean approximate accuracy without 5, respectively. Both interactions between group and development were revealed to be statistically non-significant. In contrast to the effects on mean accuracy, the results from the above analysis suggest that the responses within ± 1 of the correct answer were similar between children *versus* adults with WS and the groups of 4-5 *versus* 6-7 year old TD children. This may indicate that that response variability was more closely related to the target number in the TD groups. Furthermore, the comparison of the WS groups and the 4-5 *versus* 9-10 year old TD children suggests similar developmental changes, yet overall greater approximate accuracy in the TD groups.

Estimation Variability

As mentioned in the introduction, the findings from animals, children, and adults suggest that non-verbal, low-level representations of numerical magnitude are characterized by the fact that the variability in participants' responses is directly proportional to the target magnitude. This implies that the ratio of variability to the mean estimate (COV) should be constant across numerosities. To assess whether the COV was constant across numerosities and whether there were differences in scalar variability between groups, a 4 (Numerosity) × 6 (Group), 2-way ANOVA was computed (see Figure 6). Both the main effect of number [F (3, 336) = 7.4, p < .001] and the main effect of group [F (5, 336) = 61.8, p < .001] were significant. Furthermore, the interaction effect of Group × Number was also significant: F (15, 336) = 4.8, p < .001.

To assess both main effects and interactions on COV scores, 1-way ANOVAs were computed for each group separately. A significant effect of numerosity on COV was found in the group of 6-7 year olds [F (3, 84) = 8.0, p < .001], the 9-10 year olds [F (3, 52) = 22.27, p < .001], and in the group

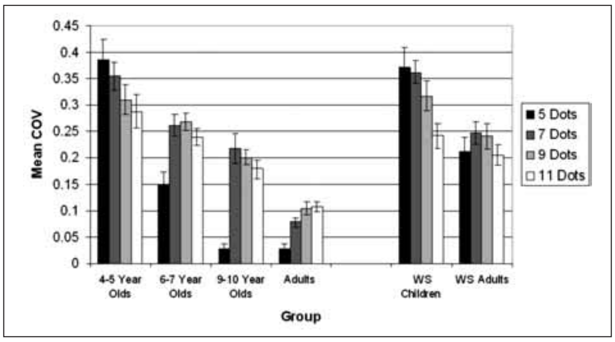

Fig. 6 – *Mean coefficient of variation (COV) by group and numerosity. Error bars denote the standard error of the mean.*

of TD adults [F (3, 44) = 13.53, p < .001]. Bonferroni *post-hoc* comparisons revealed that in all three groups the COV scores for 7, 9, and 11 dots did not differ significantly from each other, but all differed significantly from the COV scores for estimations of 5 dots (all p < .01). As can be seen from Figure 3, accuracy for estimating 5 dots increases at a faster rate than for 7, 9, and 11 dots, which explains the significant effects of numerosity on COV. In the group of 4-5 year olds, however, the effect of numerosity on COV was non-significant: F (3, 56) = 2.0, p = .125. Thus, the COV turned out to be constant for arrays of 7, 9, and 11 dots in all the groups of TD children.

A significant effect of numerosity on COV was also apparent in the group of children with WS: F (3, 56) = 4.28, p < .009. From a Bonferroni *post-hoc* comparison it emerged that the mean COV for 11 dots differed significantly from the COV for 5 and 7 but not for 9 dots. This suggests a small decrease in COV with increasing magnitude among the children with WS, which may be due to perseverations on certain numbers during the estimation of arrays of 11 dots. In other words, children with WS may have difficulties with double digits and may therefore have been responding with one or two double digits repeatedly, leading to a decrease in the observed variability of estimates. The effect of numerosity on COV in the group of adults with WS was non-significant.

Developmental Changes in Estimation Variability

Developmental changes in the COV and differences between the WS and TD developmental trajectories were analyzed in the same way as the accuracy data. The mean COV scores by group and numerosity can be found in Figure 6. Again, a 2 (Development) × 2 (Group) ANOVA was computed, with COV total and COV without 5 as the dependent variables. In a first analysis, the developmental difference between children and adults with WS was contrasted with developmental changes in COV between the groups of 4-5 and 6-7 year olds. No significant main effect of group on COV total [F (1, 60) = .2, p < .675] and COV without 5 [F (1, 60) = .8] emerged. The main effects for development on COV, however, were significant both for COV total [F (1, 60) = 28.3, p < .001] and COV without 5 [F (1, 60) = 12, p < .001]. No significant Group × Development interactions for either COV for all arrays [F (1, 60) = .04, p < .839] or COV without 5 [F (1, 60) = .1, p < .713] were found. Hence, similarly to the absolute accuracy data presented above, the developmental differences in estimation variability between children and adults with WS does not appear to exceed the very small difference between the groups of 4-5 and 6-7 year old TD children.

In a second analysis, the developmental difference between children and adults with WS was contrasted with the difference between the groups of 4-5 and 9-10 year olds. Again, no significant main effect of group on COV total [F (1, 52) = 2.2, p < .140] and COV without 5 [F (1, 52) = .2, p < .610] was found. The effects of development on COV were highly significant for both COV total [F (1, 52) = 51.0, p < .001] and COV without 5 [F (1, 52) = 22.1, p < .001]. While there was a small, but significant, interaction effect of Group × Development on COV total [F (1, 52) = 4.4, p < .039], the interaction effect on COV without 5 [F (1, 52) = 1.2, p < .307] was not significant. The absence of a significant interaction effect for COV without 5 suggests that the interaction effect for all arrays is primarily due to a qualitative difference between WS and TD individuals for the estimation of arrays of 5 dots.

DISCUSSION

The primary question addressed in this paper was whether children who develop difficulties with higher level numerical computations exhibit developmental delay and deviance in the non-verbal processing of numerical magnitude. We focused on a rare genetic developmental disorder, WS, where serious difficulties with higher-level number have been reported (Ansari et al., 2003; Paterson et al., 1999, 2006) as a case study disorder to examine these questions. We first discussed the specific results from our investigation of estimation abilities in WS and subsequently evaluated the wider limitations and implications of our work for the general study of developmental impairments of number and the consideration of ontogenetic changes in basic numerical magnitude representations.

We examined the quality and developmental trajectory of non-verbal quantity representation in children and adults with WS as well as 3 groups of TD children and one group of TD adults through an assessment of their ability to estimate number from rapidly-presented arrays of dots. We contrasted the developmental differences in absolute and approximate (± 1) estimation accuracy and variability (as measured by the COV) among individuals with WS with differences in task performance across the developmental trajectory of different age groups of TD children. Our results show that for both absolute and approximate estimation accuracy and COV, the developmental changes between children and adults with WS do not exceed those observed between groups of 4-5 and 6-7 year old TD children. The data suggest that the developmental processes that enable the rapid access of mental magnitude representation and the mapping from such representations onto external symbols (Arabic numerals and number words) are impaired in WS. The extent of this impairment we report here becomes particularly apparent when considering that the average age difference between the young and old WS groups is around 20 years while it is just 1 year between the two youngest TD groups. Furthermore, when WS developmental changes were compared to a wider developmental difference among TD individuals (4-5 *vs.* 9-10 year olds), qualitative differences between the TD and WS developmental trajectory emerged. We found that increases in estimation accuracy between the TD groups were much greater than between children and adults with WS. The same is not true, however, for estimation variability where widening the TD developmental difference to 4-5 *versus* 9-10 year olds reveals a qualitative difference between WS and TD for variability only in the estimation of 5 dots. The absolute estimation accuracy data (with the exception of 5 dots) was found to be rather low, which may indicate floor effects chance performance rather than reflecting task related cognitive processes. In view of this, we calculated another estimation accuracy measure which included both absolutely accurate measures and those responses that fell within ± 1 of the target numerosity. We called this measure "approximate estimation accuracy". We found that in all groups the number of approximately correct responses was significantly greater than the absolutely accurate responses, suggesting that participants were not simply responding at random but rather that their estimations were distributed around the target numerosity. The comparison of the developmental difference between children and adults with WS with the developmental change between 4-5 and 6-7 year old TD children revealed no significant effects of group on mean approximate accuracy without 5 and no significant effects of development of approximate estimation accuracy with or without 5. These data differ from the absolute accuracy data and reveal that approximate accuracy did undergo significant developmental changes between the two youngest groups of TD children. The comparison between the WS groups and the larger developmental difference between TD children (4-5 *vs.* 9-10 year olds), however, revealed that while the overall slope of development did not differ between WS and TD groups, TD groups exhibited greater approximate accuracy than WS groups. These findings, together with the results from the analysis of the absolute estimation accuracy highlight the developmental impairment of visual estimation abilities in WS. Interestingly, we find that while there are group differences in both absolute and approximate estimation accuracy between the WS groups and the 4-5 *versus* 9-10 year old TD groups, no such differences emerged our measure of estimation variability (COV). These findings suggest that there may be qualitative differences between our control and clinical groups in the relationship between estimation accuracy and variability. The findings suggest that while the estimation variability in the TD groups is quantitatively similar to the variability of estimates among WS children and adults, the variability is more closely related to target estimates in the TD groups. This may point to a qualitative difference in the variability of magnitude representations between our WS and TD groups.

It might be argued that the developmental differences between WS and TD children reported here are merely indicative of differences in low-level perceptual mechanisms and have little to do with non-verbal representations of numerical magnitude. In other words, the developmental impairment of visual estimation abilities we report may be a consequence of a low-level visual impairment and thus unrelated to the representation of numerical magnitude. We consider this possibility unlikely for a number of reasons. First, the results from our pre-task suggest that individuals in all groups were able to recognize stimuli (Arabic numerals) presented to them for the same duration as the dot arrays. Thus, all groups were able to visually process information presented for such a short duration. In addition, performance on the enumeration of 2 and 3 dot arrays was at ceiling for all groups tested. Thus, children and adults with WS can, like all controls, process dot stimuli presented for 250 msec. In addition, these findings suggest that any impairment of visual processing must be specific to the visual processing of arrays of objects that are greater than 3 and therefore that these processes are likely to be part of crucial input pathways to magnitude representations. In other words, if some of the difficulties reported here could be traced back to visuo-spatial processing, then these impairments are crucial components of the kind of numerical magnitude processing engaged by our visual estimation task. However, it is important to note that while both WS and TD groups were able to enumerate arrays of 2 and 3 dots with perfect accuracy, it is possible that typical and atypical groups achieved this high level of accuracy by using different strategies. Indeed, it is even possible that within the TD groups, the enumeration strategies for these small arrays changed over developmental time. In future investigations of this kind, reaction time data should be collected to allow for a greater discrimination of the enumeration performance. It is important to note that the stimuli used in the present study do not control for all the variables that are continuous with numerosity. While area was controlled by ensuring that dots in all slides occupied the same amount of space in the display, this may therefore have allowed participants to use individual dot size as a cue to numerosity. Though perhaps plausible, there need to be further empirical investigations into the cues that young children and adults use during visual enumeration. Thus far the issue of covariates of non-symbolic stimuli in numerosity paradigms has been thoroughly investigated in studies with human infants (Clearfield and Mix, 1999; Feigenson et al., 2002) and conclusions from these studies may not readily generalize to older participants. Only by pitting various cues against each other will it be become possible to establish what cues young children and adults use in visual enumeration of non-symbolic stimuli. A related issue concerns the strategies that participants may use during the visual enumeration of non-symbolic stimuli such as dots. It is possible that there are significant differences in the strategies used during enumeration between WS and TD groups as well as among TD groups of different developmental levels. It is possible that TD adults are able to gauge the likely range of targets after only a few exemplars. This in turn may enable TD adults to perform more accurately than younger children or individuals with WS who may not possess the ability to exert this kind of strategic control over their behavior during visual estimation. This makes it likely that part of the variance in participants' performance measured in our study can be explained by cognitive differences between groups that have little to do with the representations underlying the non-verbal representations of numerical magnitude. One way of controlling for the variance explained by particular strategies and deficits unrelated to numerical cognition is to choose a wider array of control tasks. For example, in the context of the present study it might have been useful to include a non-numerical visual attention tasks (such as a visual search paradigms). Covarying performance on such tasks with performance on the estimation task may have helped to narrow down the variance explained by representations specific to numerical estimation. Of course, in work with young children and individuals with developmental disorders, it is important to find a balance between control over confounding variables and the time it takes to administer experimental tasks.

Implications and Conclusions

In view of the above considerations, we contend that our findings may suggest that the representational systems underlying visual estimation abilities are impaired in WS, leading to little improvement over developmental time in representational parameters and abilities to map between magnitude representations and external, symbolic representations of numbers. It could be that this is linked to atypical neural development in WS, leading to atypical developmental changes in the firing bandwidth of the kind of neurons thought to be involved in number processing (Nieder et al., 2002; Nieder and Miller, 2004) and/or to abnormalities in the structure and function of the parietal lobes, as has recently found to be true of children with very low-birth weight who suffer from calculation difficulties (Isaacs et al., 2001) and of individuals with Turner syndrome who exhibit deficits in arithmetic (Molko et al., 2003). It remains for future studies to investigate the neural basis of numeracy deficits in WS. While previous work pinpointed the higher-level number deficits in WS, our current study sought to identify low-level impairments processing nature that may have downstream effects on the development of higher-level number concepts. These low-level deficits may well explain why children with WS have such difficulties with simple arithmetic and numbers in everyday life. In other words, we adopt a neuroconstructivist perspective, whereby it is posited that fruitful insights into the nature of developmental disorders can only be gained by understanding the developmental processes that lead to atypical outcomes (Karmiloff-Smith, 1998; Ansari and Karmiloff-Smith, 2002). While the data we present here provide an insight into atypical development of basic magnitude processing competencies, cross-syndrome comparisons and work with children who suffer from specific impairments of the mathematical skills with otherwise normal intelligence (dyscalculia) are needed to test for commonalities and similarities across different manifestations of developmental impairments of numerical cognition. In view of this, it is important to note that the specific details of the present results may not readily generalize to other developmental disorders of numerical cognition. A recent paper by Paterson et al. (2006) explores this issue by comparing the numerical abilities of individuals with Williams and Down's syndrome.

Besides revealing differences in visual estimation abilities between typically and aTD children, our results have implications for the study of the development of typical numerical magnitude representations. Our findings across the TD groups highlight the extent to which basic number abilities undergo significant quantitative developmental changes between 4 and 10. In view of the absence of such developmental changes in our atypical group, it becomes apparent how important a deeper analysis of the functional significance of such changes is. Our findings raise the question of whether such changes imply that the overlap between numerical magnitude representations decreases over developmental time in TD children, leading to more discrete numerical representations. Moreover, it may be the case that such increasing differentiation of the features of magnitude representation over developmental time allows for more automatic access to representations and more efficient mappings between representations and their external symbolizations (e.g., number words, Arabic numerals). In our view, adopting a conceptual and methodological approach that focuses on developmental change, we are now in a position to move beyond the currently dominant approach in the field of numerical cognition research, concerned with demonstrating qualitative ontogenetic continuity in basic systems of number representations, towards a serious consideration of the significance of quantitative developmental changes for the development of numerical cognition.

We conclude that our results with WS as a case study disorder highlight the importance of investigating the developmental role played by impairments to low-level magnitude representation systems in the ontogenesis of typical number abilities and the development of deficits in high-level numerical cognition.

Acknowledgments. This research was supported by a doctoral studentship from the Williams Syndrome Foundation, UK and MRC Programme Grant No. G9715642 and NIH grant R2TW06761-01 to Annette Karmiloff-Smith. We would like to thank all the families and children who volunteered for this research and the Williams Syndrome Foundation, UK for their help with recruitment. We would also like to thank Dr. Daniel Tranel and three anonymous reviewers for their insightful and helpful comments.

References

Ansari D, Donlan C, Thomas MS, Ewing SA, Peen T and Karmiloff-Smith A. What makes counting count? Verbal and visuo-spatial contributions to typical and atypical number development. *Journal of Experimental Child Psychology, 85:* 50-62, 2003.

Ansari D and Karmiloff-Smith A. Atypical trajectories of number development: A neuroconstructivist perspective. *Trends in Cognitive Sciences, 6:* 511-516, 2002.

Badian NA. Dyscalculia and nonverbal disorders of learning. In Myklebust HR (Ed), *Progress in Learning Disabilities* (vol. V). New York: Grune & Stratton, 1983.

Beuren AJ. Supravalvular aortic stenosis: A complex syndrome with and without mental retardation. *Birth Defects, 8:* 45-46, 1972.

Brannon EM and Terrace HS. Ordering of the numerosities 1 to 9 by monkeys. *Science, 282:* 746-749, 1998.

Butterworth B. *The Mathematical Brain*. London: Macmillan, 1999.

Clearfield MW and Mix KS. Number versus contour length in infants' discrimination of small visual sets. *Psychological Science, 10:* 408-411, 1999.

Dehaene S. *The Number Sense: How the Mind Creates Mathematics*. Oxford: Oxford University Press, 1997.

Dehaene S, Dehaene-Lambertz G and Cohen L. Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences, 21:* 355-361, 1998.

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

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

Elliot CD, Smith P and McCulloch K. *British Ability Scales*. Windsor: NFER-Nelson, 1996.

Feigenson L, Carey S and Spelke E. Infants' discrimination of number *versus* continuous extent. *Cognitive Psychology, 44:* 33-66, 2002.

Feigenson L, Dehaene S and Spelke E. Core systems of number. *Trends in Cognitive Sciences, 8:* 307-314, 2004.

Gallistel CR and Gelman R. Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences, 4:* 59-65, 2000.

Geary DC. Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114:* 345-362, 1993.

Gersten R and Chard D. Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *The Journal of Special Education, 33:* 18-28, 1999.

Goswami U, Thomson J, Richardson U, Stainthorp R, Hughes D, Rosen S and Scott SK. Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proceedings of the National Academy of Sciences of the USA, 99:* 10911-10916, 2002.

Griffin S and Case R. Re-thinking the primary school math curriculum: An approach based on cognitive science. *Issues in Education, 3:* 1-49, 1999.

Gross-Tsur V, Manor O and Shalev RS. Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine and Child Neurology, 38:* 25-33, 1996.

Huntley-Fenner G. Children's understanding of number is similar to adults' and rats': Numerical estimation by 5-7-year-olds. *Cognition, 78:* B27-B40, 2001.

Isaacs EB, Edmonds CJ, Lucas A and Gadian DG. Calculation difficulties in children of very low birthweight: A neural correlate. *Brain, 124:* 1701-1707, 2001.

Jordan NC, Kaplan D and Hanich LB. Achievement growth in children with learning difficulties in mathematics: Findings from a two-year longitudinal study. *Journal of Educational Psychology, 94:* 586-597, 2002.

Karmiloff-Smith A. Development itself is the key to understanding development disorders. *Trends in Cognitive Sciences, 2:* 389-398, 1998.

Lipton JS and Spelke ES. Origins of number sense. Large-number discrimination in human infants. *Psychological Science, 14:* 396-401, 2003.

Mechner F. Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behavior, 1:* 109-121, 1958.

Meck WH and Church RM. A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9:* 320-334, 1983.

Molko N, Cachia A, Riviere D, Mangin JF, Bruandet M, Le Bihan D, Cohen L and Dehaene S. Functional and structural alterations of the intraparietal sulcus in a developmental dyscalculia of genetic origin. *Neuron, 40:* 847-858, 2003.

Moyer RS and Landauer TK. Time required for judgements of numerical inequality. *Nature, 215:* 1519-1520, 1967.

Nieder A, Freedman DJ and Miller EK. Representation of the quantity of visual items in the primate prefrontal cortex. *Science, 297:* 1708-1711, 2002.

Nieder A and Miller EK. A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences of the USA, 101:* 7457-7462, 2004.

Paterson SJ, Brown JH, Gsodl MK, Johnson MH and Karmiloff-Smith A. Cognitive modularity and genetic disorders. *Science, 286:* 2355-2358, 1999.

Paterson SJ, Girelli L, Butterworth B and Karmiloff-Smith A. Are numerical impairments syndrome specific? Evidence from Williams syndrome and down's syndrome. *Journal of Child Psychology and Psychiatry, 47:* 190-204, 2006.

Piazza M, Izard V, Pinel P, Le Bihan D and Dehaene S. Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron, 44:* 547-555, 2004.

Platt JR and Johnson DM. Localization of position within a homogenous behavior chain: Effects of error contingencies. *Learning and Motivation, 2:* 386-414, 1971.

Sekuler R and Mierkiewicz D. Children's judgments of numerical inequality. *Child Development, 48:* 630-633, 1977.

Shalev RS and Gross-Tsur V. Developmental dyscalculia. *Pediatric Neurology, 24:* 337-342, 2001.

Siegler RS and Jenkins E. *How Children Discover New Strategies*. Hillsdale: Erlbaum, 1989.

Tallal P. Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language, 9:* 182-198, 1980.

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

Whalen J, Gallistel CR and Gelman R. Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 2:* 130-137, 1999.

Xu F and Spelke ES. Large number discrimination in 6-month-old infants. *Cognition, 74:* B1-B11, 2000.

Daniel Ansari, Department of Psychology, University of Western Ontario, London, Ontario, Canada. e-mail: daniel.ansari@uwo.ca

(*Accepted 4 August 2005*)