Running head: COLOUR PERCEPTION IN WILLIAMS SYNDROME

Colour Discrimination and Categorisation in Williams Syndrome

Emily K. Farran[1]

Matthew B. Cranwell[2]

James Alvarez[3]

Anna Franklin[4]

[1]Department of Psychology and Human Development, Institute of Education, University of London

[2]Institute of Neuroscience, University of Newcastle

[3]Department of Psychology, University of Surrey

[4]School of Psychology, University of Sussex

Address correspondence to:

Emily K. Farran

Department of Psychology and Human Development

Institute of Education

25 Woburn Square

London WC1H 0AA

Tel: +44(0)207 6126272

E-mail: E.Farran@ioe.ac.uk

Abstract

Individuals with Williams syndrome (WS) present with impaired functioning of the dorsal visual stream relative to the ventral visual stream. As such, little attention has been given to ventral stream functions in WS. We investigated colour processing, a predominantly ventral stream function, for the first time in nineteen individuals with Williams syndrome. Colour discrimination was assessed using the Farnsworth-Munsell 100 hue test. Colour categorisation was assessed using a match-to-sample test and a colour naming task. A visual search task was also included as a measure of sensitivity to the size of perceptual colour difference. Results showed that individuals with WS have reduced colour discrimination relative to typically developing participants matched for chronological age; performance was commensurate with a typically developing group matched for non-verbal ability. In contrast, categorisation was typical in WS, although there was some evidence that sensitivity to the size of perceptual colour differences was reduced in this group.

Key Words: Williams syndrome, colour discrimination, colour categorisation, visual perception

Colour Discrimination and Categorisation in Williams Syndrome

Introduction

Atypical colour perception is reported in a number of neurodevelopmental disorders. For example, individuals with Autism show reduced colour memory, reduced chromatic discrimination, but typical categorical perception of colour (Franklin et al., 2008; 2010), whilst reduced blue-yellow chromatic discrimination is reported in Attention Deficit Hyperactivity Disorder (Banaschewski et al., 2006).

Williams syndrome is a rare genetic disorder, in which visuo-spatial functioning is impaired relative to verbal abilities (Jarrold, Baddeley & Hewes, 1998). Colour perception has not previously been explored in WS. This is because the visuo-spatial deficit in WS has been attributed to atypical dorsal stream functioning (e.g. Atkinson et al., 2001), and colour perception has broadly been classified as a ventral stream function (e.g. Beauchamp, Haxby, Jennings & De Yoe, 1999). It has therefore been assumed that colour perception is typical in WS. However, a brief review of the WS literature, and the literature on colour perception, questions this assumption on a number of grounds, which we detail below. The current study assessed colour perception across a battery of tasks to provide a comprehensive account of colour perception in WS.

First, we explore ventral stream processing in WS. Ventral stream activation in WS is certainly less impaired than dorsal stream function in WS and a number of studies have shown typical activation in individuals with WS when carrying out tasks that are associated with the ventral stream (e.g. Mobbs et al., 2007). However, the ventral stream cannot be described as ‘intact’. For example, Grice et al. (2003) demonstrated atypical neural activation in the ventral stream in WS when participants were presented with illusory contours, which suggests that these stimuli were being

processed in an atypical manner. Relatedly, taking a brain network approach it is important to recognise that there is much structural and functional cross-talk between dorsal and ventral streams (Van Essen et al., 1992), and so it stands to reason that a deficit in the dorsal stream in WS could impact ventral stream functions. In support of this, Sarpal et al. (2008) report impaired functional connectivity between the parahippocampal gyrus (ventral stream) and parietal cortex (dorsal stream) in WS.

Second, we consider the cortical activation associated with colour perception. Colour perception involves multiple areas of visual cortex and although activation is predominantly in the ventral stream, Claeys et al. (2004) reported dorsal activation in the Intraparietal Sulcus (IPS) and dorsal premotor cortex during a colour perception task. Exploration of primary visual cortex in WS is limited to area V1 where Galaburda, Holinger, Bellugi, and Sherman (2002) report increased cell packing and neuronal size in WS autopsy specimens, compared to control brains. The IPS is an area in WS cortex which has repeatedly been shown to be atypical in both function (Meyer-Lindenberg et al., 2005) and structure (Kippenhan et al., 2005). These findings, therefore, provide further impetus for exploring colour perception in WS.

To explore basic perceptual discrimination of colour hues, we used a widely employed colour vision test, the Farnsworth-Munsell 100 Hue test (Farnsworth, 1943). This task enables one to establish overall level of chromatic discrimination as well as to determine whether chromatic discrimination shows any distinct patterns of deficit for either the red-green or blue-yellow colour axes. Colour vision on these two axes are described as the two ‘cardinal directions’ in colour space (Krauskopf, Williams & Heeley, 1982) and derive from two physiologically distinct subsystems. The ‘red–green’ subsystem reflects a comparison of signals from retinal cone photoreceptors that are sensitive to signals from long- (L-) versus signals from medium- (M-)

wavelengths. The 'blue-yellow' subsystem reflects a comparison of signals from photoreceptors that are sensitive to short-wavelengths, with the combined signals from long and medium-wavelength cones. These two subsystems extend beyond the retina to the pathways from the retina to the cortex. In the Farnsworth-Munsell 100 Hue test, the participant is asked to put 85 coloured caps, which vary incrementally in hue around the hue circle, in order. The test is commonly used to assess accuracy of chromatic discrimination, and due to simple instructions and low task demands it is also suitable for use with children (e.g., Roy, Podgor, Collier & Gunkel, 1991; Verriest, Van Laetham & Uvijls, 1982), even for children as young as 5 years old (e.g. Kinnear & Sahraie, 2002). This test has been successfully used with people with Attention Deficit Hyperactivity Disorder (Banaschewski et al., 2006) and children with Autism (Franklin et al., 2010), but had not previously been used with people with WS. The average IQ in WS is approximately 60 (Farran & Karmiloff-Smith, 2012) and so to determine whether participants understood what was required of them, we also included a control task in which participants arranged a series of achromatic, grey caps sequentially from white to black (as in Franklin et al., 2010).

A second motivation for investigating colour perception in WS was tentative evidence from previous research that colour categorisation in WS may be atypical. In an investigation of spatial categories in WS, Farran and Jarrold (2005) included a matching-to-sample colour categorisation task as a control task. The design of the colour task was confounded by differences in the perceptual distance between a blue prototype and blue exemplars, compared to the green stimuli. The blue exemplars were more similar to the blue prototype than the green exemplars were to the green prototype, i.e. 'green' trials were harder than 'blue' trials. However, this confound revealed an interesting, albeit marginal effect which merits controlled exploration.

The WS group showed impaired categorisation when matching exemplars to the green category prototype, relative to non-verbal matched typically developing control children. A tentative conclusion from this is that the WS group did not have as strong colour categorisation as the TD group, and hence were more influenced by differences in perceptual distance than the TD group.

To further investigate this, in the current study, we included three additional tasks. The first is a replication of the matching-to-sample task employed by Farran and Jarrold (2005), correcting for the confound in perceptual distance by equating the distance of exemplars to the prototypes in chromatic perceptual space. Participants viewed a series of target coloured squares flanked by a blue prototype and a green prototype square and were asked whether the target was more similar to the square to the left or to the right. To explore potential differences in perceptual versus verbal categories, in a second categorisation task, participants viewed a randomised sequence of target squares only and were asked to verbally classify each square as green or blue.

The third task was a further measure to investigate the proposal that the WS group were more influenced by the perceptual distances between colours than the TD group in Farran and Jarrold's (2005) study. Here, we used a visual search task, where participants were asked to locate target coloured circles amongst distracters, when the perceptual chromatic difference between targets and distracters was either large or small. If increased weighting of perceptual distance can explain the atypical pattern of performance on Farran and Jarrold's matching-to-sample task, then we would expect a greater effect of perceptual distance on search accuracy in the WS group compared to control participants. Comparing overall performance for WS and controls on the chromatic search task also provides a second test, in addition to the

Farnsworth-Munsell 100 hue test, of whether chromatic discrimination is generally reduced in WS.

If one follows the assumption that colour perception is typical in WS, it is possible that individuals with WS will show the same patterns of performance as chronological age matched typically developing participants. However, in order to take task demands into account we also included a group of typically developing children matched to the WS group on non-verbal level of ability. The battery of tasks will provide evidence on whether chromatic discrimination, perceptual and linguistic colour categorisation, and the influence of perceptual distance on chromatic search are atypical in WS.

Method

Participants

Three groups of participants took part in the study: a group with Williams syndrome (WS: N=19), who had all received phenotypic diagnosis from a clinician as well as a positive Fluorescence In Situ Hybridisation (FISH) test, which confirms deletion of the elastin gene on the long arm of chromosome seven; a group who were 'typically developing' and who were individually matched in chronological age and gender to the WS group (TDCA: N=16), and a group who were 'typically developing' and who were individually matched in non-verbal cognitive ability and gender (TDMA: N=19) as assessed by the Ravens Coloured Progressive Matrices (RCPM; Raven, 1993). Note that the dataset includes sixteen, rather than nineteen TDCA participants; due to constraints on the availability of equipment, the final three TDCA participants could not be assessed. A strict solution to this would be to exclude the three WS and TDMA participants who were without a TDCA match. However, given that the groups remain extremely well matched at a group level (see Table 1), and excluding participants

would reduce power, we decided against this, and included all participant data in the analyses. Participants also completed the British Picture Vocabulary Scale II (BPVS II; Dunn, Dunn, Whetton & Burley, 1997) as a measure of verbal ability. Table 1 gives the number of males and females, mean chronological age, RCPM and BPVS scores for the three groups.

Table 1 about here

Set-up and tasks

All tasks were conducted in a darkened room, and the testing table was illuminated from above with a simulated natural daylight (D65, Gretag Macbeth lamp). Participants completed four tasks: Farnsworth-Munsell 100 hue test; matching-to-sample; naming and visual search. The order of the tasks was fixed in the following order: matching to sample, first visual search block, F-M 100 hue test, second visual search block; naming tasks. The stimuli, design and procedure for each task is outlined in turn.

Farnsworth-Munsell 100 hue test (F-M 100 hue test)

The F-M 100 hue test consists of 85 coloured caps that vary only in hue, which are distributed across four trays. When ordered correctly, adjacent hues are more similar to each other than with any other hue in the series. The test was administered following standard test instructions: one of four trays of coloured caps was presented to the participant with only beginning and end caps in place. The remaining caps were distributed in a random layout on the table in front of the participant. The participant was asked to arrange the caps in the correct order. Standardized prompts of, “What one is most like this one?” were given at the beginning of each tray. Once participants had completed the hue series of one tray,

they were presented with another until all four trays were complete. The experimenter noted down the order of caps using the numbers on the back of each cap.

A control task was also included. In this task, participants were presented with achromatic stimuli that varied only in luminance. The stimuli were printed with a colour printer (Hewlett Packard Designjet 800PS) and measured with a colorimeter (Avantes SpectroCam 75RE) to verify that stimuli varied only in luminance. Each stimulus was the same size and shape as the coloured centre of an F-M 100 hue cap. There were 23 stimuli of varying luminance (all CIE, 1931, x = 0.33, y = 0.33). The anchor stimuli at each end of the tray appeared black (Y = 2.48) and white (Y = 80.52) respectively. The average incremental variation in luminance was Y = 3.55 (SD = 2.67). The procedure and instructions were identical to the F-M 100 hue test.

Matching-to-Sample Task

Stimuli and Design. Coloured stimuli were taken from the Munsell Colour System, an approximately standardised perceptual colour space (see Hunt & Pointer, 2012). Stimuli varied only in Munsell Hue, with Munsell Value (roughly equivalent to lightness) and Munsell Chroma (roughly equivalent to saturation) kept constant (7 and 8 respectively). Eight stimuli were taken from the blue-green region of Munsell space, so that four stimuli were either side of the blue-green category boundary (7.5BG, Bornstein & Monroe, 1980). The hue step difference between each stimulus was 1.25 hue, with a step of 2.5 hue between the two stimuli straddling the boundary (5B, 1.25B, 10BG, 8.75BG, 6.25BG, 5BG, 3.75BG, 10G). Chromaticity co-ordinates (CIE x,y,Y, 1931) were generated for the stimuli using look up tables which assume an illuminant C white point (x=.33, y=.33, Y=100). Following a process of printer calibration, stimuli of these chromaticity co-ordinates were printed on an HP Photosmart printer, and x,y,Y co-ordinates were verified with an Avantes Spectrocam.

The coloured stimuli were printed as circular colour patches (diameter = 5.5cm) on a grey background (10cm x 10cm).

Procedure. Participants were presented with the best examples of blue and green from the stimulus set (5B and 10G). These were placed at left and right locations on the table. The participant was then given the stack of 6 remaining stimuli and were asked to allocate each stimulus to the stimulus on the left or right on the basis of which it appeared to go best with. This was repeated 4 times, with the stimulus order randomised each time.

Naming Task

Participants were required to name all eight of the stimuli of the matching to sample task. Stimuli were presented on a grey background (x=.33, y=.33, Y= 43.06), individually, and participants were asked to name the stimulus as either blue or green. Each stimulus was presented 4 times.

Visual Search Task

Stimuli and Design. Stimuli from the matching-to-sample task were used to form target-distracter stimulus pairs that varied the size of Munsell hue step (5 hue step or 2.5 hue step). There were 3 pairs of each hue step size (5 hue step: 5B-10BG; 10BG-5BG; 5BG-10G and 2.5 hue step: 1.25B-8.75BG; 8.75BG-6.25BG; 6.25BG-3.75BG). Search arrays were printed where each stimulus in the array was a coloured circle (diameter = 1.75cm), and the spatial location of each stimulus was pseudo-randomised so that no two circles were within 1.5cm or further than 5cm of another circle. A target was present at the top of each search array. This was separated from the search array by a black line. There were arrays with 4, 12 or 20 distracters, each with 4 targets. There were 4 search arrays per stimulus pair (total = 72), and each stimulus in a pair appeared for an equal number of arrays as the target or distracter.

Procedure. Search arrays were presented in a pseudo-randomised order within each block so that the same set size was completed before moving onto the next set size. Participants were instructed to mark off all of the circles that they thought were the same colour as the target (presented at the top of the search array). Four practice arrays consisting of black and grey coloured circles were completed before the actual trials to check that the participant understood the task.

Results

F-M 100 Hue Test

Two participants with WS failed to complete this task due to time constraints, motivation or fatigue. One TDMA and a further WS participant were excluded due to high errors on the greyscale version of the task (scores beyond the 95% confidence interval of the TDMA group), suggestive of a failure to understand the task procedure. Despite this, the groups remained well matched at a group level. This analysis included the following groups: WS, N=16; TDCA, N=16; TDMA, N=18.

The Farnsworth-Munsell 100 Test hue responses were analysed to calculate partial error scores for blue-yellow and red-green sub-systems of colour vision. These partial error scores represent the amount of errors made when ordering the hue series taking into account the magnitude of each error. They provide an error score for the blue-yellow and red-green axes of the hue series separately, and total to give an overall error score in ordering the hue series (see Smith et al., 1985 for further detail on scoring the F-M 100 test). Figure 1 gives the partial error scores for the three participant groups.

Figure 1 about here

As can be seen in Figure 1, for both red-green and blue-yellow subsystems, the TDCA group were more accurate at ordering the hues than the WS or TDMA groups, with little difference between WS and TDMA performance. This was supported by a two-way ANOVA on error scores with Axis (blue-yellow / red-green) and Group (TDCA/WS/TDMA) as factors. There was a significant difference in overall error across the groups, $F(2,47)=45.94$, $p<.001$, and Tukey post-hoc tests revealed significantly less error for TDCA (mean=18.43, SD=9.63) than TDMA (mean=211.39, SD=82.18) and WS groups (mean=161.50, SD=60.24), (both $p<.001$), and marginally less error in the WS group compared to the TDMA group ($p=.051$). There were also significantly less errors for red-green (mean=154.97, SD=108.80) than blue-yellow (mean=166.04, SD=109.72) axes, $F(1,47)=6.46$, $p=.02$. There was no significant interaction of Group and Axis, $F(2,47)=1.23$, $p=.30$.

Error scores were calculated for the greyscale control as in the F-M 100 hue test. Here, the TDCA made fewer errors (mean = 17.19, SD=7.22) than WS (mean=33.13, SD=20.71) or TDMA (mean=29.11, SD=18.26). This represented a significant group difference, $F(2,47)=4.04$, $p=.024$. Tukey comparisons demonstrated that this was due to fewer errors in the TDCA group compared to the WS group only ($p=.024$, all other comparisons, $p>.05$). We compared the effect of group on the greyscale control task to the effect of group for the chromatic stimuli. Error score per cap was calculated for the greyscale control, and for the F-M 100 hue test. The data was analysed using ANOVA with Test (control / F-M 100) and Group (TDCA/ WS/ TDMA) as factors. Unsurprisingly, this showed a main effect of Group, $F(2, 47)=38.47$, $p<.001$) and of Test, $F(1, 47)=90.14$, $p<.001$. Crucially, there was also a significant interaction, $F(2, 47)=34.57$, $p<.001$ due to a larger group effect in the F-M 100 hue test than the greyscale control task, which indicated that the differences across the groups in the F-

M 100 hue test were accounted for by differences in chromatic sensitivity over and above any effects of task demands (Group effects: F-M 100: $p<.001$; control test: $p=.024$).

Colour Categorisation: Matching-to-sample and Naming

The percentage of times that a stimulus was matched with the blue prototype, or was named as blue, was calculated for each stimulus and each participant. Figure 2 gives the classification curves for the matching-to-sample (2a) and for naming (2b) for the three participant groups.

Figure 2 about here

Sigmoid curves were fitted using nonlinear least squares to the classification curves from matching-to-sample and naming tasks[1]. For the match-to-sample classification task there were significant differences in the slope of the sigmoid curve across the three groups ($F(2,52)=3.55$, $p<.05$), with Bonferroni post-hoc tests revealing significant differences between TDMA (mean = 0.44, SD=.046) and TDCA (mean = 0.08, SD=.19) slopes, but no significant differences between WS (mean slope = 0.29, SD=0.46) and the TD groups. There were no significant differences across groups in the slope for the naming classification sigmoid curves ($F<1$).

[1] The horizontal asymptotes of the sigmoids were fixed at 0% and 100% classification, leaving two parameters: the inflection point (the point at which the curve estimates 50% classification) and the slope of the curve. The slope can be thought of as a measure of the consistency to which participants name or match the colours. When a participant's data is very consistent, the slope reduces to 0, and the curve becomes a straightforward A or B classification. Due to the slope parameter estimates from this analysis being non-normal, boot strapped confidence intervals for F were computed by sampling with replacement from the slope values of each participant (N=54) and running ANOVAs. The 95% upper confidence interval for F was found to be 3.10, demonstrating that the F value in the ANOVA of 3.55 was not likely to be solely due to the distribution of the slope values. The significance of the difference between the TDMA and TDCA group ($t=3.11$) was confirmed in a similar manner with the estimate for the upper 95% confidence to be $t = 2.12$.

Visual Search

Four participants with WS failed to complete this task due to time constraints, motivation or fatigue. These participants were excluded. The groups remained well matched at a group level. This analysis included the following groups: WS, N=15; TDCA, N=16; TDMA, N=19.

The average search accuracies were calculated for the search trials with a large and small perceptual chromatic difference between the target and distracters, for the three participant groups (see Figure 3). Search accuracy was the difference between the proportion of correct and incorrectly identified targets. As can be seen in the figure, performance of the WS group appears to be closely matched to the TDMA group, with strong TDCA group performance. Comparison to ceiling performance of 1.00 showed that no groups were performing at ceiling ($p<.05$ for all). There appears to be a general effect of perceptual difference, where search is more accurate for large than small target-distracter differences. This effect of perceptual difference on search also appears to be equivalent for the WS and TDMA group, but weakened in the TDCA group.

Figure 3 about here

A 2-way ANOVA on search accuracy, with Group (WS/TDMA/TDCA) and Perceptual Difference (large/small) as factors, supported these inferences. There was a significant difference in accuracy across groups, $F(2,47)=26.00$, $p<.001$, with Bonferroni corrected post-hoc t-tests revealing significantly greater accuracy for TDCA (mean=.95, SD=.11) than TDMA (mean=.71, SD=.11) and WS (mean=.71, SD=.10) groups ($p<.001$), but no significant difference between TDMA and WS accuracy. Search was also more accurate for large (mean=.90, SD=0.10) than small (mean=.70, SD=0.14) target-distracter perceptual differences, $F(1,47)=51.84$, $p<.001$.

The effect of Perceptual Difference also interacted with Group, $F(2,47)=6.93$, $p<.005$. To explore this interaction, the difference in accuracy for large and small conditions was calculated, and these difference scores were compared across groups. The effect of Perceptual Difference was significantly weaker for TDCA than TDMA ($p<.005$) and WS ($p<.05$), but there was no difference between TDMA and WS ($p>.99$).

Discussion

We employed a battery of tasks to measure colour perception in Williams syndrome. We were interested in whether individuals with Williams syndrome discriminated and categorised colours in a typical manner. Across the battery of tasks, the *pattern* of performance of the WS group was typical, whilst the *level* of performance was commensurate with their level of non-verbal ability (i.e. they did not differ from the TDMA group).

The F-M 100 Hue test is sensitive to chromatic discrimination across development from 5 years through to at least 79 years (Kinnear & Sahraie, 2002). In the typical population, a U-shaped function of sensitivity to differences in hue is observed with increasing age. This has been shown using both the F-M 100 Hue test (Kinnear & Sahraie, 2002) and a two-alternative forced choice discrimination test (Knoblauch, Vital-Durand & Barbur, 2001). Specifically, children become progressively stronger at chromatic discrimination with increasing age, this ability plateaus from approximately fifteen years of age, and is followed by deterioration from approximately forty years of age. Because this pattern is observed across both the F-M 100 Hue test and a forced choice test with comparatively reduced task demands, Kinnear and Sahraie (2002) attribute it to changes in chromatic discrimination ability rather than developmental changes in the ability to meet the demands of the task. To verify this, in the present study, we included a greyscale

version of the F-M 100 Hue test. Comparison of performance between the greyscale and F-M 100 Hue test broadly supported the assertion that the F-M 100 Hue test provides a valid measure of chromatic discrimination ability. The performance of the two typically developing groups was in line with previous findings. That is, the TDMA group had weaker chromatic discrimination than the TDCA group, and both groups showed weaker blue-yellow discrimination than red-green discrimination. The individuals with WS also followed this pattern. Their level of performance was weaker than the TDCA group, and only marginally stronger than the TDMA group.

However, there were also group differences in the greyscale control task in this study. In contrast to Kinnear and Sahraie's (2002) assertion, this effect suggests that maturational factors, such as attention and spatial abilities, might impact performance on the F-M 100 Hue test after all. For example, arriving at the final cap order requires spatial thinking as well as an ability to maintain attention throughout arranging each tray of caps. The TDMA group were matched to the WS group on a non-verbal task, the Raven's Coloured Progressive Matrices, and so it is likely that spatial and attentional abilities were broadly similar to those of the TDMA group (Breckenridge et al., 2012; Farran & Formby, 2012). Thus, whilst we can be confident that the comparatively larger group difference in the F-M 100 Hue test itself indicates reduced chromatic discrimination in the WS than would be expected for their chronological age, we suggest that the WS and TDMA group were performing at a lower level than the TDCA group on the F-M 100 Hue test due to poorer chromatic discrimination, but also on account of maturational differences in areas such as attention and spatial abilities. Importantly, although discrimination is reduced in WS, this group show a typical pattern of colour discrimination across the two axes of colour vision.

Reduced chromatic discrimination has also been reported in Autism (Franklin et al., 2010), and there are several potential explanations for this (Franklin & Sowden, 2011). Although the ventral stream cannot be ruled out as a candidate for impaired colour discrimination in WS, other explanations are possible. For example, increased neuronal size and increased cell packing (Galaburda et al., 2002) at V1 might impact colour discrimination in WS. Or, perhaps impairments in the Intraparietal Sulcus (IPS) in WS (Kippenhan et al., 2005; Meyer-Lindenberg et al., 2005) impact the decision making processes involved in colour discrimination in this group (see Claeys et al., 2004). This latter hypothesis emphasises the importance of co-operation between the dorsal and ventral streams in the typical population.

The matching-to-sample and naming tasks measured participants' colour categorisation. Categorisation of colour has been reported from infancy (e.g. Franklin et al., 2008). Given this, and the low task demands, unlike the F-M 100 hue test, a difference between the TDMA and TDCA was not predicted. Results showed classic sigmoid curves, which reflect that hues closer to the blue-green colour boundary were harder to classify, on both tasks. In the matching to sample task, the slope of the function was steeper in the TDCA group than the TDMA group. This could be taken to suggest stronger categorisation in the TDCA group, but this effect does not appear to be particularly robust given that it was not observed in the naming task. In the naming task, there is a stronger focus on categories because each stimulus is being labelled as 'blue' or 'green', whereas in the match-to-sample task, participants categorise based on colour similarity. We tentatively suggest that the naming task encouraged the TDMA group to draw on their categorical understanding to a greater extent than the match-to-sample task. Hence, we suggest that perhaps their

categorisation is akin to the TDCA group, but it is more fragile such that when the task demands are less explicit, performance is influenced by perceptual distance.

The categorisation tasks were included to determine whether colour categorisation is atypical in individuals with WS, and to further explore the results of Farran and Jarrold (2005) in which a WS group found green stimuli harder to categorise than blue stimuli. We were interested in whether these previous findings were simply an artefact of a confound in the task design, or whether individuals with WS do have weaker categorical perception of colours than TD children. The current results do not demonstrate any atypical processing in the WS group; the slopes of the curves in both the matching-to-sample task (which is most similar to Farran & Jarrold's task) and the naming task, do not differ from either the TDMA or the TDCA group. Thus, we can conclude that individuals with WS have typical categorisation of blue and green, and that the previous findings were likely related to experimental confounds.

The visual search task provided a final measure of sensitivity to the size of perceptual colour differences. Given that visual search performance in typical development improves with development on account of attentional maturation (Trick & Enns, 1998), it is likely that the difference in level of performance between the two TD groups relates to differences in selective attention abilities. Similarly, the WS group did not perform at an age appropriate level. They performed at the level expected of their level of non-verbal ability (i.e. at the level of the TDMA group). Previous studies which have used visual search tasks with individuals with WS report an increased rate of target-distracter confusability in targets defined by size (Scerif, Cornish, Wilding, Driver & Karmiloff-Smith, 2004) and atypical search patterns in WS (Karmiloff-Smith et al., 2012). Thus, it is likely that additional attentional factors

that are not related to colour perception contributed to the level of ability of the WS group here.

The significant interaction demonstrated differences in patterns of performance on the visual search task; we can be confident that this reflects differences in sensitivity to the size of perceptual difference. Clearly, the TDCA group found this task easy. Statistically, they did not perform at ceiling and so it appears that the reduced effect of perceptual difference in this group is real and reflects a stronger sensitivity than the TDMA or WS groups. The WS group performed at the level of the TDMA group, with both WS and TDMA groups showing comparable effects. In line with the match-to-sample and naming tasks, this does not support our earlier assertion that the pattern of performance in Farran and Jarrold (2005) might be explained by greater sensitivity to perceptual distance in WS.

In summary, individuals with WS do not show any atypical patterns in colour discrimination ability, colour categorisation or sensitivity to perceptual distance. However, on tasks where differences were observed between the typically developing groups, the individuals with WS performed at level commensurate with their non-verbal ability rather than their chronological age. That is, performance on the F-M 100 hue test indicated reduced chromatic discrimination ability in WS, relative to chronological age, and the visual search task gave some indication that individuals with WS are less sensitive to the size of colour difference than predicted by chronological age. In contrast, categorisation of colour appears typical in WS. Further research is required to determine how this profile of colour processing relates to brain function in WS; potentialcandidates are area V1 of primary visual cortex, and/or cross-talk effects of the IPS on ventral stream functioning.

Acknowledgments

Thank you to Worplesdon Primary School and the Williams Syndrome Foundation UK for enabling this research to take place. We are also grateful to Nigel Woodger or the University of Surrey for assistance with stimulus preparation. Special thanks go to the participants for taking part in the study.

References

Atkinson, J., Anker, S., Braddick, O., Nokes, L., Mason, A., & Braddick, F. (2001). Visual and visuospatial development in young children with Williams syndrome. *Developmental Medicine & Child Neurology, 43*, 330–337.

Breckenridge, K., Atkinson, J. & Braddick, O. (2012). Attention. In Farran, E.K. and Karmiloff-Smith, A. (Eds). Neurodevelopmental Disorders Across the Lifespan: A Neuroconstructivist Approach. (pp.119-134). *Oxford University Press.*

Banaschewski, T., Ruppert, S., Tannock, R., Albrecht, B., Becker, A., Uebel, H., Sergeant, J.A., & Rothenberger, A. (2006). Colour perception in ADHD. *Journal of Child Psychology and Psychiatry, 47*, 568–572.

Beauchamp, M.S., Haxby, J.V., Jennings, J., & DeYoe, E.A. (1999). An fMRI version of the Farnsworth-Munsell 100-hue test reveals multiple color-selective areas in human ventral occipito-temporal cortex. *Cerebral Cortex, 9*, 257–263.

Claeys, K.G., Dupont, P., Cornette, L., Sunaert, S., Van Hecke, P., De Schutter, E., & Orban, G.H. (2004). Color discrimination involves ventral and dorsal stream visual areas. *Cerebral Cortex, 14*, 803–822.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *British picture vocabulary scale, second edition.* Windsor, UK: NFER-Nelson.

Farnsworth, D. (1943). The Farnsworth Munsell 100-hue dichotomous tests for colour vision. *Journal of the Optical Society of America, 33*, 568–576.

Farran, E.K. & Formby, S. (2012). Visual Perception and Visuospatial Cognition. In Farran, E.K. and Karmiloff-Smith, A. (Eds). Neurodevelopmental Disorders Across the Lifespan: A Neuroconstructivist Approach. (pp. 225-246). *Oxford University Press.*

Farran, E. K., & Jarrold, C. (2005). Evidence for unusual spatial location coding in Williams syndrome: an explanation for the local bias in visuo-spatial construction tasks? *Brain and Cognition, 59,* 159-172.

Farran, E.K. and Karmiloff-Smith, A. (Eds) (2012). Neurodevelopmental disorders across the lifespan: a neuroconstructivist approach. Oxford University Press.

Franklin, A., Sowden, P., Burley, R., Notman, L., & Alder, E. (2008). Color perception in children with Autism. *Journal of Autism and Developmental Disorders, 38,* 1837-1847.

Franklin, A., Sowden, P., Notman, L., Gonzalez-Dixon, M., West, D., Alexander, I., et al. (2010). Reduced chromatic discrimination in children with autism spectrum disorders. *Developmental Science, 13,* 188-200.

Franklin, A. & Sowden, P. (2011). Colour in Autism Spectrum Disorders. In Biggam, Carole P., Hough, Carole, Kay, Christian & Simmons, David R.C. (Eds.), *New Directions in Colour Studies*. John Benjamins, Amsterdam, NL.

Galaburda, A., Holinger, D. P., Bellugi, U., & Sherman, G. F. (2002). Williams syndrome: Neuronal size and neuronal-packing density in primary visual cortex. *Achives of Neurology, 59,* 1461-1467.

Grice, S. J., de Haan, M., Halit, H., Johnson, M. H., Csibra, G., Grant, J., et al. (2003). ERP Abnormalities of Illusory Contour Perception in Williams Syndrome. *Neuroreport, 14,* 1773-1777.

Karmiloff-Smith, A., D'Souza, D., Dekker, T.M., Van Herwegen, J., Xu, F., Rodic, M., Ansari, D. (2012). Genetic and environmental vulnerabilities in children with neurodevelopmental disorders. *Proceedings of the National Academy of Science, 109,* 17261-17265.

Kinnear, P., & Sahraie, A. (2002). New Farnsworth-Munsell 100 hue test norms of normal observers for each year of age 5–22 and for age decades 30–70. *British Journal of Ophthalmology, 86*, 1408–1411.

Kippenhan, J.S., Olsen, R.K., Mervis, C.B., Morris, C.A., Kohn, P., Meyer-Lindenberg, A., Berman, K.F. (2005). Genetic contributions to human gyrification: sulcal morphometry in Williams syndrome. *Journal of Neuroscience, 25*, 7840-7846.

Knoblauch, K., Vital-Durand, F., & Barbur, J.L. (2001). Variation of chromatic sensitivity across the life span. *Vision Research, 41*, 23–26.

Krauskopf, J., Williams, D.R., & Heeley, D.W. (1982). Cardinal directions of color space. *Vision Research, 22*, 1123–1131.

Meyer-Lindenberg, A., Mervis, C. B., Sarpal, D., Koch, P., Steele, S., Kohn, P., et al. (2005). Funtional, structural, and metabolic abnormalities of the hippocampal formation in Williams syndrome. *Journal of Clinical Investigation, 115*, 1888-1895.

Mobbs, D., Eckert, M. A., Menon, V., Mills, D., Korenberg, J., Galaburda, A. M., et al. (2007). Reduced parietal and visual cortical activation during global processing in Williams syndrome. *Developmental Medicine and Child Neurology, 49*, 433-438.

Jarrold, C., Baddeley, A. D., & Hewes, A. K. (1998). Verbal and nonverbal abilities in the Williams syndrome phenotype: evidence for diverging developmental trajectories. *Journal of Child Psychology and Psychiatry, 39*, 511-523.

Raven, J. C. (1993). *Colored Progressive Matrices*. Oxford, UK: Information Press Ltd.

Roy, M.S., Podgor, M.J., Collier, B., & Gunkel, R.D (1991). Color vision and age in a normal North American population. Graefe's Archive for Clinical and Experimental Ophthalmology, 229, 139–144.

Sarpal, D., Buchsbaum, B. R., Kohn, P. D., Kippenham, J. S., Mervis, C. B., Morris, C. A., et al. (2008). A genetic model for understanding higher order visual processing: functional interactions of the ventral visual stream in Williams syndrome. *Cerebral Cortex, 18*, 2402-2409.

Scerif, G., Cornish, K., Wilding, J., Driver, J., & Karmiloff-Smith, A. (2004). Visual search in typically developing toddlers and toddlers with Fragile X or Williams syndrome. *Developmental Science, 7*, 116-130.

Trick, L. M., & Enns, J. T. (1998). Lifespan changes in attention: The visual search task. *Cognitive Development, 13*, 369-386.

Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: An integrated systems perspective. *Science, 255*, 419-423.

Verriest, G., Van Laetham, J., & Uvijls, A. (1982). A new assessment of the normal ranges of the Farnsworth-Munsell 100-Hue test scores. *American Journal of Ophthalmology, 93*, 635–642.

Table 1. Demographics of the three participant groups: the number of males and females (M:F); chronological age, Ravens Progressive Coloured Matrices (RPCM) and British Picture Vocabulary Scale (BPVS) scores for the TDCA, WS and TDMA groups.

	TDCA (N=16)	WS (N=19)	TDMA (N=19)
M:F	8:8	9:10	9:10
Chronological Age (years; months): mean (s.d.)	15; 07 (3; 02)	15; 05 (3; 05)	5; 06 (0;07)
RCPM: mean (s.d.)	34.00 (1.41)	17.37 (4.97)	17.21 (4.35)
BPVS: mean (s.d.)	122.38 (17.83)	88.58 (16.92)	65.58 (10.24)

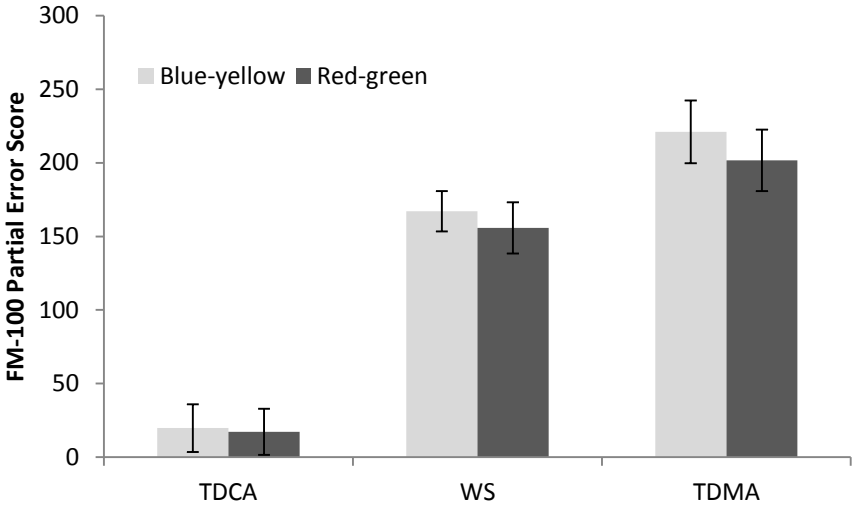

Figure 1. Partial error scores (+/-1se) for blue-yellow and red-green axes on the Farnsworth-Munsell 100 hue test, for TDCA, WS and TDMA groups.

Figure 2a: Matching-to-sample

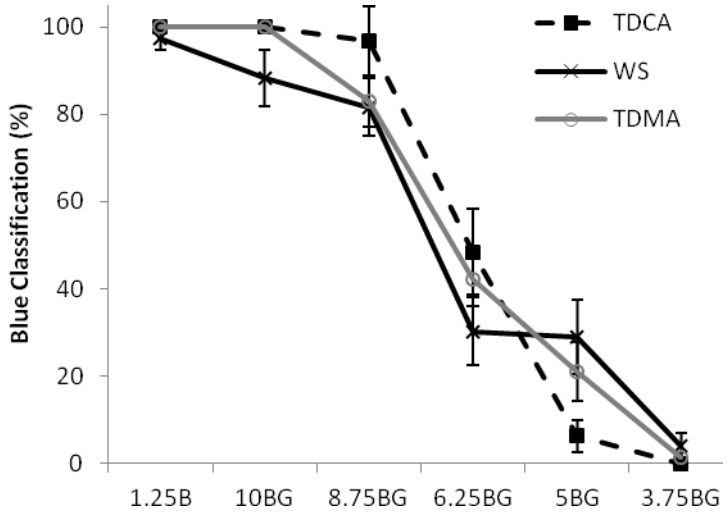

Figure 2b: Naming

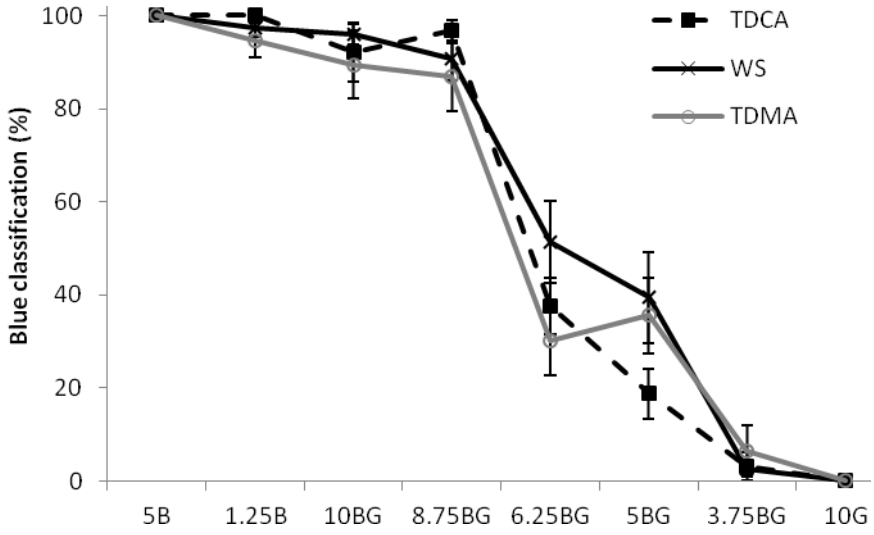

Figure 2. Classification curves: the mean percentage (+/-1se) of a times each stimulus was classified as blue, averaged for TDCA, WS and TDMA groups.

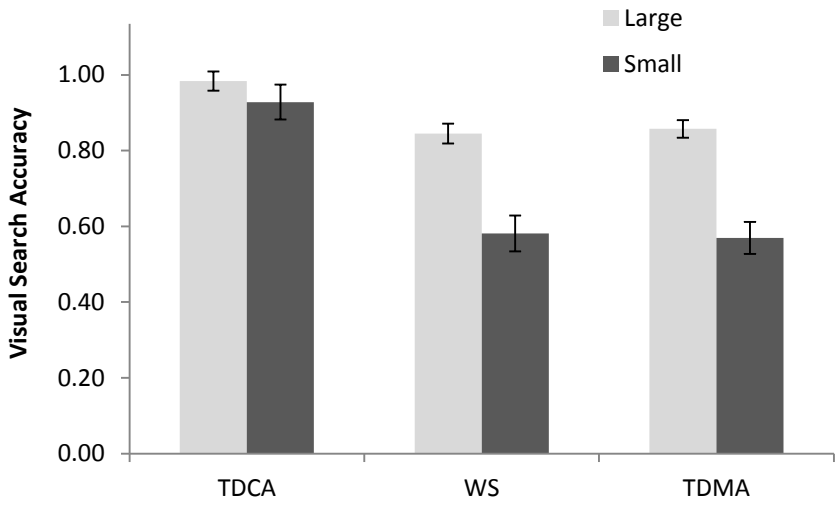

Figure 3. Search accuracy (+/-1se) when the chromatic difference between targets and distracters was large and small, for TDCA, WS and TDMA groups.