COGNITIVE NEUROPSYCHOLOGY, 2005, 22 (0), 1–12

Psychology Press
Taylor & Francis Group

SOCIAL-PERCEPTUAL ABILITIES IN ADOLESCENTS AND ADULTS WITH WILLIAMS SYNDROME

Daniela Plesa Skwerer, Alyssa Verbalis, Casey Schofield, Susan Faja, and Helen Tager-Flusberg

Boston University School of Medicine, USA

People with Williams syndrome (WMS) have a unique social phenotype characterised by unusually strong interest in other people and an engaging and empathic personality. Two experiments were designed to test whether this phenotype is associated with relatively spared abilities to decode mental-state information from nonverbal cues. The first experiment involved a modified version of the revised Eyes Test. The second experiment probed the ability to label emotions from brief dynamic facial displays. Adolescents and adults with WMS were compared to age-, IQ-, and language-matched participants with learning/intellectual disabilities, and age-matched normal controls. In both experiments the WMS group performed at a significantly lower level than the normal controls, and no different from the well-matched comparison-group with intellectual disabilities. These findings, contradicting earlier reports in the literature, argue against the view that in WMS social-perceptual abilities are relatively spared and can explain the social profile associated with this neurodevelopmental disorder.

Research with people with Williams syndrome (WMS) has surged in the last 15 years, yet the scientific literature is marked by the persistence of apparently contradictory claims, especially with respect to the social skills and language abilities of this population. Given the rarity of this genetic syndrome, many studies have reported results based on small samples, with participants spanning a wide age range, and variability in performance has been often overlooked in interpreting empirical results. At the outset, interest in the syndrome was sparked by case studies of unusual social and linguistic behaviour (e.g., Bellugi, Bihrle, Neville, Jernigan, & Doherty, 1992; Bellugi, Marks, Bihrle, & Sabo, 1988; Udwin, Yule, & Martin, 1987), which led researchers to search for spared abilities in the context of general mental retardation and of severe specific cognitive deficits in visual-spatial construction skills. This paper focuses on the social component of the WMS profile, addressing the question of whether the unusual sociability and friendly personality characteristic of people with this neurodevelopmental disorder (Gosch & Pankau, 1994; Jones et al., 2000; Mervis & Klein-Tasman, 2000) is associated with relatively

Correspondance should be addressed to Daniela Plesa Skwerer, PhD, Department of Anatomy & Neurobiology, Boston University School of Medicine, 715 Albany Street L-814, Boston, MA 02118, USA (e-mail: dplesas@bu.edu).

This research was supported by grants from the National Institute of Child Health and Human Development (RO1 HD 33470), the National Institute on Neurological Diseases and Stroke (RO1 NS 44824), and by grant M01-RR00533 from the General Clinical Research Center program of the National Center for Research Resources, National Institutes of Health. We express our sincere thanks to the National Williams Syndrome Association and New England regional chapter for their help in recruiting participants; and to the families and individuals who participated in this study. We gratefully acknowledge the reviewers for helpful comments and suggestions for supplementary analyses.

http://www.tandf.co.uk/journals/pp/02643294.html DOI:10.1080/02643290542000076

spared abilities in decoding mental state information from nonverbal facial cues.

Initial studies on the cognitive mechanisms that might underlie the social phenotype of WMS focused on classic theory of mind skills (e.g., Karmiloff-Smith, Klima, Bellugi, Grant, & Baron-Cohen, 1995; Tager-Flusberg & Sullivan, 2000). Based on the unusual social interest and friendly behaviour of people with WMS, some researchers have hypothesised that they would show sparing in social skills, including the capacity for mentalising, or understanding other people's minds (Karmiloff-Smith et al., 1995; Tager-Flusberg, Boshart, & Baron-Cohen, 1998). Only one study exploring theory of mind abilities in people with WMS (Karmiloff-Smith et al., 1995), in which an older group of children, adolescents, and young adults with WMS was administered a battery of standard first-order and second-order false belief tasks, an interpretation of gaze direction task, and a task involving interpretation of metaphor and sarcasm, reported that the majority of the participants with WMS passed these tasks, concluding that theory of mind may be an "islet of relatively preserved ability" in WMS (p. 202). Methodological limitations of this study, however—small number of participants from a wide age-range (9- to 23-year-olds), tested on standard theory of mind tasks that are developmentally appropriate for much younger children and are primarily language-based, lack of an appropriate and well-matched, nonautistic, comparison group—call into question the (often-quoted) evidence of sparing in theory of mind skills in WMS (see Tager-Flusberg & Sullivan, 2000, for a critique).

Recently, more judiciously designed studies involving a larger number of participants and carefully matched control groups revealed impairments in domains initially declared "intact" in WMS, such as language (e.g., Bates, 2004; Karmiloff-Smith et al., 1998) and social cognition. Evidence from a series of experiments using standard measures of theory of mind, including false belief, as well as higher-order tasks, found that children and adolescents with WMS were no different from well-matched control groups on the social-cognitive components of theory of mind (Plesa Skwerer & Tager-Flusberg, in press; Sullivan & Tager-Flusberg, 1999; Sullivan, Winner, & Tager-Flusberg, 2003; Tager-Flusberg & Sullivan, 2000). Given the complex cognitive and linguistic demands, such as working memory and inferencing, associated with these tasks, it is perhaps not surprising that theory of mind skills are not spared in a population with mental retardation. At the same time, researchers have suggested that using less demanding on-line tasks taping social *perceptual* skills may reveal unique capacities in this population, known for its sensitivity to other people's emotional states in social interaction situations (Jones et al., 2000; Tager-Flusberg & Sullivan, 2000; Thomas, Becerra, & Mervis, 2002).

One study investigated social perceptual abilities in adults with WMS (Tager-Flusberg et al., 1998) using the original version of the Eyes Task, developed by Baron-Cohen and his colleagues (Baron-Cohen, Joliffe, Mortimore, & Robertson, 1997). On this task participants are asked to choose the label that best fits the mental state expression from the eye region of the face displayed in black-and-white photographs. In their study, Tager-Flusberg et al. found that adults with WMS were significantly better than a matched group of adults with Prader-Willi syndrome, another genetically based neurodevelopmental disorder, although about half the participants with WMS performed worse than the normal adult controls. This study, however, only included a small number of participants (13 adults with WMS), and some research suggests that Prader-Willi syndrome might be associated with specific *deficits* in social-affective information processing (Beardsmore, Dorman, Cooper, & Webb, 1998; Koenig, Klin, & Schultz, in press; Whitman & Accardo, 1987), raising concerns about their suitability as a comparison group in studies exploring social perceptual abilities in WMS. The Eyes Task used by Tager-Flusberg et al. (1998) raises additional concerns. For each stimulus only two labels were presented, allowing chance performance at 50%. Moreover, the choices typically involved semantic opposites (e.g., *friendly–unfriendly*), leaving open the possibility that the relatively spared performance of some adults with WMS could have been based solely on discriminating the valence of mental state (the last two problems have been recognised and addressed by the authors of the test in a revised version of the task, see Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001, for a detailed discussion). Thus, while the results of Tager-Flusberg et al. (1998) were interpreted as evidence of relatively unimpaired social-perceptual skills in WMS, these methodological problems cast doubt on true sparing of mentalising abilities in people with WMS.

Indirect evidence for impairments in processing the social and psychological significance of facial displays in people with WMS was reported by Bellugi, Adolphs, Cassady, and Chiles (1999). They found that participants with WMS rated faces as more approachable and more trustworthy than did control groups, which is consistent with anecdotal reports about the indiscriminant trust and friendliness of people with WMS toward strangers. Furthermore, in contrast to the findings reported by Tager-Flusberg et al. (1998), several studies found that children and adults with WMS were no better than matched controls in labelling or discriminating basic emotions expressed in faces (Gagliardi, Frigerio, Burt, Cazzaniga, Perrett, & Borgatti, 2003; Tager-Flusberg & Sullivan, 2000), despite their relative strength in recognising facial identity as measured on standardised tests such as the Benton Test of Facial Recognition (Bellugi et al., 1988; Bellugi, Wang, & Jernigan, 1994; Tager-Flusberg, Plesa Skwerer, Faja, & Joseph, 2003; Wang, Doherty, Rourke, & Bellugi, 1995).

The experiments reported in this paper extend this line of research using two different kinds of tasks to tap mentalising skills in WMS. In the first experiment we followed up on the original study by Tager-Flusberg et al. (1998), using a modified version of the Eyes Task as revised by Baron-Cohen and colleagues (2001). In their revision, the authors reduced chance performance to 25% by presenting each stimulus with four choices of labels, and no longer included choices differing only in valence. We modified the task presentation and the vocabulary used for the label choices to make the task suitable for people with mental retardation. In the second experiment we presented participants with dynamic facial displays portrayed in brief colour video-clips, to explore whether people with WMS would find labelling these ecologically more natural stimuli easier, and thus might reveal greater sparing in social perception than has been found using black-and-white photographs.

EXPERIMENT 1

Method

Participants. This experiment included 43 adolescents and adults with WMS, 42 participants with learning or intellectual disabilities of mixed aetiology (LID) matched on age, IQ, and language scores, and 46 normal controls in the same age range as the participants from the clinical groups. Adolescents in each group ranged from 12; 0 to 17;11 years, and adults ranged from 18; 0 to 36;9 years. The participants with WMS (24 females and 19 males) were recruited through the Williams Syndrome Association. The diagnosis of WMS was confirmed by a geneticist and the FISH test. Participants with LID (25 females and 17 males) were recruited from a special school serving this population. Three of these 42 participants were excluded from analyses because they had an autistic profile on the parent report Social Responsiveness Scale (Constantino, 2004). The normal controls (34 females and 12 males) were recruited from local schools and universities and all had IQ and language scores within 1 *SD* of the mean.

Participants were administered the Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 1990) and the Peabody Picture Vocabulary Test (PPVT-III; Dunn & Dunn, 1997) to assess IQ and verbal knowledge. They also received the short form of the Benton Test of Facial Recognition (Benton, Hamsher, Varney, & Spreen, 1983) to assess their ability to discriminate and recognise face identity. Table 1 presents the characteristics of the participant groups on standardised measures. An analysis of variance (ANOVA) confirmed that all three groups were matched on chronological age, $F(2, 125) = 1.58$, $p = .21$. As expected, the groups differed in IQ, $F(2, 125) = 172.5$, $p < .001$, verbal knowledge (i.e., PPVT-III scores), $F(2, 125) = 93.77$, $p < .001$, and face recognition, $F(2, 117) = 28.22$, $p < .001$. Post hoc Scheffé tests indicated that the two clinical groups had comparable IQ ($p = .55$) and verbal knowledge scores ($p = .47$), but both groups scored significantly lower than the normal control group ($p < .001$) on both these measures. However, the WMS participants were significantly better on face recognition than the LID participants ($p < .001$), and were almost as proficient as the normal control group on the Benton test ($p = .16$).

Procedures. The modified version of the revised Eyes Task consisted of two blocks of 32 black-and-white photographs of the eye region of faces, representing equal numbers of men and women, presented on a computer. Each photograph was accompanied by two mental state labels, positioned below the picture of the eyes. Across the 64 trials each photograph appeared twice (one per block). In each presentation the target word was paired with a different foil, and a response was scored correct only if the target term was selected on both occurrences of the same stimulus, thus reducing chance performance to 25%, resulting in a possible total score of 0–32 points.

The target and foil words were randomly positioned on the left or right side on the screen. Table 2 presents the words used for target labels and foils in the two blocks of trials. Several mental state terms from the Baron-Cohen et al. (2001) version of the test were simplified (e.g., *reflective* was replaced with *thoughtful, regretful* with *sorry*) and others were eliminated (e.g., *dispirited, indecisive, dominant, sceptical*) because, based on preliminary testing, these terms were considered too difficult for some of our participants with WMS or LID.

During each stimulus presentation the experimenter read the two words to the participant, and a glossary with definitions of every term and examples of use was available. Participants were asked to select the label that best matched the mental state expression in the photograph.

Results

Preliminary analyses of sex and age group differences (adolescents, 12;1–17;11 years versus adults, 18 years and older) yielded no significant main effects or interaction effects for these factors; therefore, these variables were not considered further in analyses. Within each group there were no significant differences in scores as a function of the gender of the person in the photograph, or of the position on the screen (left/right) of the target mental state words.

Means and standard deviations of scores for each group are presented in Table 3. A one-way ANOVA conducted to evaluate group differences in accuracy was significant, $F(1, 125) = 55.32$,

Table 1. *Participant characteristics—Experiment 1*

	Williams syndrome (n = 43)		*Learning disabled (n = 39)*		*Normal controls (n = 46)*	
	M (SD)	*Range*	*M (SD)*	*Range*	*M (SD)*	*Range*
Chronological age	20;8 (7;5)	12;1–36;1	18;11 (7;5)	13;6–36;9	18;7 (5;7)	12;1–34;5
Full Scale IQ (KBIT)	70.2 (11.8)	52–100	72.6 (11.2)	55–100	105.6 (6.6)	91–115
Vocabulary (PPVT-III)	80.3 (10.8)	61–110	83.3 (10.4)	58–102	109.5 (11.7)	81–126
Benton Face Recognition	21.8 (2.6)	16–27	18.6[a] (3.2)	11–25	22.8 (1.8)	18–27

[a] Based on $N = 32$.

Table 2. *Mental state terms used in Experiment 1*

Block 1		*Block 2*	
interested	very scared	*upset*	annoyed
playful	comforting	alarmed	*day dreaming*
embarrassed	*day dreaming*	shy	*expecting*
joking	*insisting*	*nervous*	insisting
preoccupied	annoyed	*determined*	bored
grateful	*flirting*	relaxed	*insisting*
Thoughtful	annoyed	*careful*	too proud
sorry	*uneasy*	*sad*	excited
puzzled	*nervous*	begging	*preoccupied*
upset	very scared	*thoughtful*	annoyed
ashamed	*confident*	disappointed	*blaming*
worried	annoyed	*mean*	shy
nervous	*mean*	confused	*desire*
unsure	too proud	disappointed	*flirting*
joking	*desire*	*playful*	bored
shy	*sad*	*friendly*	shocked
thoughtful	confused	joking	*confident*
shocked	*not trusting*	daydreaming	*worried*
very careful	bored	*interested*	pleased
thoughtful	impatient	flirting	*sorry*
grateful	*preoccupied*	friendly	*worried*
worried	embarrassed	*preoccupied*	mean
amused	*determined*	*thoughtful*	caring
serious	ashamed	excited	*thoughtful*
annoyed	*thoughtful*	*unsure*	grateful
interested	joking	very sad	*interested*
*t*hreatening	*expecting*	*very careful*	shocked
friendly	guilty	alarmed	*serious*
joking	*careful*	*not trusting*	very scared
blaming	depressed	amused	*thoughtful*
very scared	*sorry*	*uneasy*	friendly
daydreaming	impatient	confused	*daydreaming*

Items in *italic* face were target terms on each trial.

$p < .001$, and the diagnostic group factor accounted for 47% of the variance in scores. Follow up pairwise comparisons (Scheffé test) showed that the normal control group performed significantly better than either of the two clinical groups, $p < .001$, while the two clinical groups were not significantly different from each other, $p = .74$.

Based on the scoring system described, scores of 13 and higher out of 32 represent significantly above-chance performance using the binomial theorem. All participants in the normal control group performed above chance level, while 72.1% of WMS group and 61.5% of the LID group performed above chance level. When considering all three groups, the omnibus chi-square test indicated significant differences among the groups in the proportion scoring above chance: $\chi^2(2, N = 128) = 20.56$, $p < .001$. Follow-up pairwise comparisons showed significant differences between the normal control group and each of the clinical groups, $\chi^2(1, N = 89) = 14.83$, $p < .001$ for WMS, and $\chi^2 (1, N = 85) = 21.48$, $p < .001$ for LID. However the WMS and LID groups were not significantly different from one another: $\chi^2 (1, N = 82) = 1.03$, $p = .31$.

Table 3. *Performance on Experiment 1*

	Mean	*SD*	*Range*
Williams syndrome	16.07	4.27	9–28
Adolescents ($N = 21$)	14.81	3.92	9–23
Adults ($N = 22$)	17.27	4.32	11–28
Learning disabled	15.33	5.40	7–26
Adolescents ($N = 21$)	15.43	5.02	7–26
Adults ($N = 18$)	15.22	5.97	7–25
Normal controls	24.04	3.15	18–30
Adolescents ($N = 27$)	22.67	3.03	18–28
Adults ($N = 19$)	26.00	2.16	21–30

Notable in both clinical groups was the wider variability in performance compared to the normal controls: the range of scores for normal controls was 18–30, while in the WMS group the range was 9–28, and in the LID group it was 7–26. We further examined the proportion of participants in each clinical group who scored within 2 *SD*s of the mean of the normal controls, which included all the normal control participants. In the WMS group 13 participants (30.23%) scored within this range, while in the LID group 11 individuals (28.21%) scored within the normal control range. The proportion of participants scoring within this range was comparable in the two clinical groups, $\chi^2 (1, N = 82) = 0.41$, $p = .84$.

We also conducted supplementary analyses using the data from only the first presentation of each stimulus, to control for possible effects of previous exposure to the same stimulus from the 1st to the 2nd experimental block (in this analysis, chance performance was 50%). Results were very similar: A one-way ANOVA conducted to evaluate group differences was significant, $F(1, 125) = 37.38$, $p < .001$, and the diagnostic group factor accounted for 37% of the variance in scores. Follow-up pairwise comparisons (Scheffé test) showed that the normal control group performed significantly better than either of the two clinical groups, $p < .001$, while the two clinical groups were not significantly different from each other, $p = .284$. When comparing the proportion of participants in each clinical group who scored within 2 *SD*s of the mean of the normal controls for the first presentation of the stimuli, we again found significant differences between the three groups, χ^2 (1, $N = 128$) = 36.63, $p < .001$. However, the proportion of participants scoring within this range was comparable in the two clinical groups, χ^2 (1, $N = 82$) = 0.28, $p = .76$, including about 44% of the WMS group and 38.5% of the LID group[1].

Analyses of responses on individual items indicated that similar target mental states were relatively more difficult for all three groups (e.g., *thoughtful, blaming, nervous*). After ranking item-related performance in order of difficulty within each group, we calculated correlations of rankings for group comparisons. Spearman rank order correlations were highly significant ($r_s = .70$, $p < .001$ for normal controls and WMS; $r_s = .710$, $p < .001$ for normal controls and LID; and $r_s = .837$, $p < .001$ for WMS and LID), suggesting similar patterns of performance across the three groups.

Finally, we explored the relations between performance on the Eyes Task and the standardised measures of intellectual functioning (IQ based on KBIT), language (PPVT-III), and facial recognition (Benton test). For the WMS group the only significant correlation was with the Benton test scores, $r(43) = .421$, $p < .01$, while IQ and vocabulary knowledge correlations were not significant, $r(43) = .116$, $p = .46$, and $r(43) = .25$, $p = .11$, respectively. For the normal controls, performance was significantly correlated with all three standardised measures, $r(46) = .548$, $p < .001$ with IQ, $r(46) = .533$, $p < .001$ with PPVT-III scores, and $r(46) = .454$, $p < .002$ with Benton test scores, whereas for the LID group the only significant correlation was with IQ: $r(39) = .401$, $p < .02$, while for PPVT-III scores, $r(39) = .295$, $p = .07$, and for the Benton test scores, $r(39) = .289$, $p = .11$).

Brief discussion

In comparison with the LID group, the adolescents and adults with WMS did not show any relative sparing on the revised Eyes Test used in this experiment. Both the WMS and LID groups performed significantly worse than the normal controls, although about one third of the participants in these clinical groups were within the same range of performance as the normal controls. These findings contrast with the earlier report by Tager-Flusberg et al. (1998), which was based on the original Eyes Test developed by Baron-Cohen et al. (1997). In the current experiment a smaller percentage of the participants with WMS performed in the same range as the normal controls (one third compared to about one half, although the proportion was also higher in this study when considering only the 1st presentation of the stimuli, where chance performance was 50%). The differences between these two studies suggest that the current Eyes Task was indeed harder than the original for the WMS participants because they could not simply rely on the semantic valence of the mental state terms presented with each stimulus, and chance performance was reduced to 25%.

The task used in this experiment included a wide range of mental states; however, the cues were restricted to the eye region of the face. Moreover, the stimuli presented were black-and-white static photographs, further limiting the cues that might be used in everyday life to decode mental states from nonverbal facial expressions. It is possible that people with WMS may perform relatively better than matched

[1]We thank an anonymous reviewer for suggesting these supplementary analyses.

control groups in reading mental states when presented with more natural and ecologically valid stimuli. Therefore, in the second experiment we developed dynamic colour stimuli of people expressing basic emotions in order to test whether we would find relative sparing in social perception in participants with WMS when given stimuli that included multiple naturalistic cues to the key emotional expressions.

EXPERIMENT 2

Participants

In this experiment, 37 adolescents and adults with WMS, 32 individuals matched on IQ, language, and age from the group with LID, and 38 normal controls in the same age-range were included. The majority of participants in the WMS and LID groups had participated in Experiment 1 in a previous session at least 2 weeks earlier, but most of the normal control participants had not. All the participants with WMS had genetic confirmation of their diagnosis (FISH test). The three groups were matched on age, $F(2, 104) = 0.308$, $p = .74$, and the two clinical groups were also matched on IQ ($p = .94$) and PPVT-III scores ($p = .51$), while both clinical groups scored significantly lower than the normal controls on these standardised measures ($p < .001$, Scheffé test). On the Benton test of facial recognition the WMS participants performed at levels comparable to the normal controls ($p = .53$) and significantly better than the LID group ($p = .002$). Details of the standardised measures for all three groups are presented in Table 4.

Procedures

Task development. Dynamic facial expressions of emotion were selected from the "Mindreaders" collection developed by Baron-Cohen and Tead (2003). The "Mindreaders" software application contains over 2000 video and audio files illustrating expressions of emotion presented by actors of various ages and races, organised into 24 groups of emotion, named according to the emotion-concept that best represents it (e.g., *furious* and *grumpy* are included in the *Angry* group). This program was developed as a training tool to teach people with autism spectrum disorders to recognise expressions of emotions. For this study we selected a group of 6 emotion categories considered to be basic emotions: happy, sad, fearful, angry, disgust, surprise, and emotionally neutral expressions (Ekman & Friesen, 1976).

For each emotion and the neutral category we preselected 16 examples from the entire collection portrayed by men (8 examples for each expression) and women (8 examples for each). Then 30 normal adults rated the 112 video-clips selected, first on how *natural* or genuine the actor looked, and then on *intensity*, with each rating on a 1–5 scale, where 1 was the *least natural or intense* and 5 the *most natural or intense*. Raters were also asked to choose the three video-clips they believed best exemplified the emotional expression, for each emotion and gender group, based on a global evaluation. Four video-clips were included

Table 4. *Participant characteristics—Experiment 2*

	Williams syndrome (n = 37)		*Learning disabled (n = 32)*		*Normal controls (n = 38)*	
	M (SD)	*Range*	*M (SD)*	*Range*	*M (SD)*	*Range*
Chronological age	19;1 (7;1)	12;0–37	18;5 (2;7)	13;8–23;1	19;6 (6;6)	12;1–35;5
Full Scale IQ (KBIT)	68.6 (12.4)	45–94	69.7 (11.7)	52–93	106.1 (13.6)	86–141
Vocabulary (PPVT-III)	79.4 (9.9)	51–103	82.6 (10.1)	54–100	102.5 (15.1)	82–141
Benton Face Recognition	21.6 (2.6)	17–27	19.1[a] (3.3)	11–25	22.4[a](2.3)	18–26

[a] Based on $N = 28$.

in the experiment for each expression (7) and gender group (2), based on consistently high natural and intensity ratings across raters, for a total of 56 video-clips. In most cases the selected video-clips corresponded with the ones chosen as best examples by the raters.

Task administration and scoring. The video-clips, edited to 5 seconds in length, were presented individually in random order on a laptop computer screen. Immediately after each clip ended, the participant was asked to name the emotion portrayed in the video. A list of 21 emotion terms (e.g., *terrified, surprised, furious, happy, delighted, distressed, neutral, calm, afraid*, etc.) was read to each participant before starting the task as examples of labels. The list remained available throughout the task administration, but participants were told they were free to use whatever word they thought best described the emotional expression viewed.

Responses were scored as passing if the labels used by participants were clearly among the emotion terms corresponding to the category depicted (e.g., *joyful* for *happy*). If the label used reflected confusion with a different emotion category, the response was scored as 0 (e.g., *sad* for *disgusted*). Because participants were encouraged to consult the list of sample emotion labels available during the experiment if needed, there were very few "don't know" responses, or responses that could not be categorised, and those were excluded from the analyses.

Results

Preliminary analyses of sex and age group differences were conducted on overall accuracy scores, and no significant main effects or interaction effects for these factors were found. These factors were not considered further in analyses. Table 5 presents the mean and standard deviations on each facial expression for the three groups. An ANOVA conducted to evaluate group differences in overall labelling accuracy was significant, $F(2, 104) = 17.69$, $p < .0001$, and the diagnostic group factor accounted for 25% of the variance in overall performance.Post hoc comparisons (Scheffé test) indicated that the normal control group was significantly more accurate than the WMS group ($p < .001$), and than the LID group ($p < .001$), but the two clinical groups were not significantly different from each other ($p = .77$).

Scores for each of the seven facial expressions were entered into a MANOVA with group as between-subjects factor. Significant differences were found among the three groups on the set of dependent variables, Wilks' lambda = .68, $F(14, 196) = 3.03$, $p < .001$. Analyses of variance on each facial expression category were conducted as follow-up tests to the MANOVA. The groups were significantly different on labelling *fear*, $F(2, 104) = 4.33$, $p < .02$, *surprise*, $F(2, 104) = 6.06$, $p < .01$, *disgust*, $F(2, 104) = 3.71$, $p < .01$, and *neutral* expressions, $F(2, 104) = 10.03$, $p < .001$. Pairwise comparisons showed that the normal

Table 5. *Accuracy on labelling facial expressions in Experiment 2*

	Williams syndrome (N = 37)		*Learning disabled (N = 32)*		*Normal controls (N = 38)*	
	M	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
Happy	95.49	(11.55)	92.71	(18.42)	97.37	(9.11)
Sad	76.58	(29.25)	81.25	(26.69)	87.72	(19.64)
Angry	81.08	(20.09)	89.58	(19.74)	87.72	(19.64)
Fear	42.34	(33.93)	52.08	(34.85)	64.91	(29.96)
Disgust	67.57	(32.85)	69.79	(30.94)	86.84	(18.24)
Surprise	73.87	(27.39)	64.58	(32.73)	84.21	(16.87)
Neutral	54.05	(39.56)	55.21	(42.84)	86.84	(22.65)
Overall accuracy	70.14	(13.91)	72.17	(13.09)	85.09	(7.44)

controls were significantly more accurate than both clinical groups on *disgust* ($p < .01$ for the WMS comparison, and $p < .04$ for the LID), and on *neutral* expressions ($p < .0001$ for the WMS comparison, and $p < .001$ for the LID). In addition, the normal control group was significantly more accurate than the WMS group in labelling *fear* ($p < .01$), and than the LID group in labelling *surprise* ($p < .01$). The two clinical groups did not differ significantly from each other on any of the facial expressions.

We further examined the proportion of participants in each group who scored within 2 *SD*s of the mean of the normal controls for overall accuracy in emotion labelling. Only one participant from the normal control group scored below 2 *SD*s from the group mean. In the WMS group 21 participants (56.8%) scored within this range, while in the LID group 16 individuals (50%) scored within the normal control range. The two clinical groups were not significantly different from each other in the proportion of participants scoring within the normal control range, χ^2 $(1, N = 69) = 0.102$, $p = .75$.

To explore the patterns of performance in labelling emotions across the groups, we rank ordered the individual faces by difficulty, within each group. Spearman rank order correlations for group comparisons were highly significant ($r_s = .711$, $p < .001$ for normal controls and WMS; $r_s = .761$, $p < .001$ for normal controls and LID; and $r_s = .777$, $p < .001$ for WMS and LID), suggesting similar patterns of performance across the three groups.

We investigated the relationship between overall accuracy on the emotion labelling task and IQ, language, and the Benton test for each group. For the WMS group performance was correlated with PPVT-III, $r(37) = .337$, $p < .05$, and almost reached significance with Benton scores, $r(37) = .307$, $p = .06$, while $r(37) = .168$, $p = .32$ with IQ. For the normal controls, performance was significantly correlated only with Benton scores, $r(28) = .414$, $p < .05$, whereas for the LID group the only significant correlation was with IQ, $r(32) = .469$, $p < .01$.

Relationship between performance on Experiments 1 and 2

Thirty-two participants with WMS and 21 with LID were administered both the Eyes Task and the labelling facial expressions in dynamic displays task, in different experimental sessions. For the WMS group, performance on these tasks was significantly correlated, $r(32) = .427$, $p < .02$, whereas for the LID group this correlation did not quite reach statistical significance, $r(21) = .403$, $p = .07$. However, a comparison of the strength of the two correlation coefficients, using Fisher's z' transformation of r, revealed that the strength of the relation between scores on the two tasks was very similar in the two clinical groups ($z = .148$, $p = .88$ for a two-tailed nondirectional test).

Brief discussion

The findings from this experiment were similar to those reported for Experiment 1: The participants with WMS performed no better than the well-matched participants with LID in labelling facial expressions of emotion and both groups performed at a significantly lower level than normal controls. These results confirm earlier studies suggesting that people with WMS are not spared in the ability to match, recognise, or name basic emotional expressions in faces (Gagliardi et al., 2003; Tager-Flusberg & Sullivan, 2000), even though we made every effort in this experiment to employ the most naturalistic dynamic stimuli, which provided multiple cues to emotions. At the same time it should be noted that about half the participants in both clinical groups performed in the same range as the normal controls. This is higher than the proportion performing in this range found in Experiment 1, suggesting that for both the WMS and LID participants it was relatively easier to decode basic emotions (compared to the complex mental states included in Experiment 1) from dynamic stimuli (relative to static photographs restricted to the eye region of the face).

GENERAL DISCUSSION

The two experiments presented in this paper explored social perceptual skills in adolescents and adults with WMS, using tasks that varied in difficulty. The first experiment used a modified version of the revised Eyes Task (Baron-Cohen et al., 2001), which had been developed as a measure sensitive to subtle impairments in decoding nonverbal cues and is therefore considered quite demanding, given the restricted cues to the complex mental states depicted in the stimuli. In contrast, the task used in the second experiment provided maximal information based on the whole face presented in brief video-clips, incorporating full colour and motion cues to the people expressing basic emotions in a more natural format. The main findings were similar across these tasks: We found no differences in performance between the adolescents and adults with WMS and the well-matched comparison group of people with LID. However, both groups performed significantly worse, overall, than age-matched normal controls, suggesting that both people with WMS and people with LID show relative impairments in the ability to decode mental state information from faces. These findings support earlier research on this topic (e.g., Gagliardi et al., 2003; Tager-Flusberg & Sullivan, 2000); however, this study included larger groups of participants and a comparison group of learning and intellectually disabled adolescents and adults that was matched on age, IQ, and language ability to the group with WMS. This more rigorous methodological approach provides strong support for the conclusion that people with WMS are no different from other individuals with disabilities in social perceptual skills involved in processing facial cues to mental states and emotions.

At the same time we also found similar *patterns* of performance across the stimuli in both experiments by all three groups. We obtained high and significant correlations between the groups on item difficulty for each experiment, suggesting that there were no atypical aspects to the abilities of the clinical groups in perceiving and labelling emotional and other mental state expressions. For several expressions in Experiment 2, namely *happy*, *sad*, and *angry*, the participants with WMS and LID performed at the same level as the normal controls (and for recognition of *happy*, performance was at ceiling in all groups), demonstrating that they do not have across-the-board deficits in recognising emotions.

Performance across the two experiments was moderately correlated for participants in the two clinical groups who received both experimental tasks, although this only reached significance for the larger WMS group. This moderate correlation provides some support for the view that the tasks tapped similar abilities for decoding mental state expressions from nonverbal facial cues. Nevertheless, there were also significant differences in the difficulty of the two tasks, as evidenced by the higher proportion of participants in the two groups who performed in the normal range (defined as within 2 *SD*s of performance by the controls) in Experiment 2, which used more naturalistic dynamic stimuli limited to basic emotional expressions.

As in previous studies on cognitive performance in WMS, as well as other neurodevelopmental disorders, performance on the social perception tasks used in this research was quite heterogeneous—some participants performed within the normal range, while others performed at chance levels, at least on the more challenging Eyes Task. What factors predict performance on social perception tasks? The only notable difference between the WMS and the LID groups on both tasks was the pattern of correlations with standardised measures of IQ and face recognition. Specifically, in the WMS group there was no significant correlation between IQ scores and performance on either task, whereas in the LID group IQ was significantly correlated with performance on both tasks. For the normal controls the Benton Test of Facial Recognition was the most reliable standardised test correlating with task performance in both experiments; this correlation with Benton scores was also found for the WMS group. These different patterns of correlation suggest that people with WMS may rely more on mechanisms that are specific to attending and analysing *social* stimuli in decoding facial expressions, like normal controls, whereas people with LID depend more on general cognitive capacities on tasks capturing the ability to interpret mental state information. Further investigations, using different kinds of tasks and methodologies, including functional neuroimaging approaches, are needed to test these claims in a more rigorous way.

The findings reported here provide clear evidence that most people with WMS do not have relatively spared social perceptual skills (cf. Tager-Flusberg et al., 1998; Tager-Flusberg & Sullivan, 2000). They are no better than well-matched controls with LID in on-line tasks that tap the ability to decode social-affective information, just as they are no better in social-cognitive tasks (Tager-Flusberg & Sullivan, 2000). Nevertheless, children and adults with WMS appear to have a unique social phenotype characterised by unusual friendliness, especially toward strangers, and a warm empathic manner in engaging with others (Dykens & Rosner, 1999; Gosch & Pankau, 1994; Jones et al., 2000). It remains to be seen whether this phenotype is best interpreted in terms of personality variables (cf. Klein-Tasman & Mervis, 2003), differences in social motivation, or unique patterns of attention to social stimuli (Tager-Flusberg & Plesa Skwerer, in press).

Manuscript received 24 October 2004
Revised manuscript received 14 January 2005
Revised manuscript accepted 6 April 2005
PrEview proof published online day/month/year

REFERENCES

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high-functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry*, *38*, 813–822.

Baron-Cohen, S., & Tead, T. (2003). *Mind reading: The interactive guide to emotions* [Computer software]. Cambridge: Human Emotions.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*, 241–251.

Bates, E. A. (2004). Explaining and interpreting deficits in language development across clinical groups: Where do we go from here? *Brain and Language*, *88*, 248–253.

Beardsmore, A., Dorman, T., Cooper, S. A., & Webb, T. (1998). Affective psychosis and Prader-Willi syndrome. *Journal of Intellectual Disability Research*, *42*, 463–471.

Bellugi, U., Adolphs, R., Cassady, C., & Chiles, M. (1999). Towards the neural basis for hypersociability in a genetic syndrome. *Neuroreport*, *10*, 1653–1657.

Bellugi, U., Bihrle, A., Neville, H., Jernigan, T., & Doherty, S. (1992). Language, cognition and brain organization in a neurodevelopmental disorder. In M. Gunnar & C. Nelson (Eds.), *Developmental behavioral neuroscience: The Minnesota symposium* (pp. 201–232). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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

Bellugi, U., Wang, P., & Jernigan, T. L. (1994). Williams syndrome: An unusual neuropsychological profile. In S. H. Broman & J. Grafman (Eds.), *Atypical cognitive deficits in developmental disorders* (pp. 23–56). Hillsdale, NJ: Lawrence Erlbaum Associates.

Benton, A. L., Hamsher, K. de S., Varney, N. R., & Spreen, O. (1983). *Contributions to neuropsychological assessment.* New York: Oxford University Press.

Constantino, J. N. (2004). *Social Responsiveness Scale.* Los Angeles: Western Psychological Services.

Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test–III.* Circle Pines, MN: American Guidance Service.

Dykens, E. M., & Rosner, B. A. (1999). Refining behavioral phenotypes: Personality-motivation in

Williams and Prader-Willi syndromes. *American Journal of Mental Retardation*, *104*, 158–169.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect.* Palo Alto, CA: Consulting Psychologists Press.

Gagliardi, C., Frigerio, E., Burt, D. M., Cazzaniga, I., Perrett, D., & Borgatti, R. (2003). Facial expression recognition in Williams syndrome. *Neuropsychologia*, *41*, 733–738.

Gosch, A., & Pankau, R. (1994). Social-emotional and behavioral adjustment in children with Williams-Beuren syndrome. *American Journal of Medical Genetics*, *52*, 291–296.

Jones, W., Bellugi, U., Lai, Z., Chiles, M., Reilly, J., Lincoln, A., & Adolphs, R., (2000). Hypersociability in Williams syndrome. *Journal of Cognitive Neuroscience*, *12 (Suppl)*, 30–46.

Karmiloff-Smith, A., Klima, E., Bellugi, U., Grant, J., & Baron-Cohen, S. (1995). Is there a social module? Language, face processing and theory of mind in individuals with Williams syndrome. *Journal of Cognitive Neuroscience*, *7*, 196–208.

Karmiloff-Smith, A., Tyler, L. K., Voice, K., Sims, K., Udwin, O., Howlin, P., & Davies, M. (1998). Linguistic dissociations in Williams syndrome: Evaluating receptive syntaxin on-line and off-line tasks. *Neuropsychologia*, *36*, 343–351.

Kaufman, A., & Kaufman, N. (1990). *Manual for the Kaufman Brief Test of Intelligence.* Circle Pines, MN: American Guidance Service.

Klein-Tasman, B. P., & Mervis, C. B. (2003). Distinctive personality characteristics of 8-, 9-, and 10-year-olds with Williams syndrome. *Developmental Neuropsychology*, *23*, 269–290.

Koenig, K., Klin, A., & Schultz, R. T. (in press). Deficits in social attribution ability in Prader-Willi syndrome. *Journal of Autism and Developmental Disabilities.*

Mervis, C. B., & Klein-Tasman, B. P. (2000). Williams syndrome: Cognition, personality, and adaptive behavior. *Mental Retardation and Developmental Disabilities Research Reviews*, *6*, 148–158.

Plesa Skwerer, D., & Tager-Flusberg, H. (in press). Social cognition in Williams syndrome. In C. A. Morris, H. M. Lenhoff, & P. Wang (Eds.), *Williams-Beuren syndrome: Research and clinical perspectives.* Baltimore: Johns Hopkins University Press.

Sullivan, K., & Tager-Flusberg, H. (1999). Second-order belief attribution in Williams syndrome: Intact or impaired? *American Journal of Mental Retardation*, *104*, 523–532.

Sullivan, K., Winner, E., & Tager-Flusberg, H. (2003). Can adolescents with Williams syndrome tell the difference between lies and jokes? *Developmental Neuropsychology*, *23*, 85–103.

Tager-Flusberg, H., Boshart, J., & Baron-Cohen, S. (1998). Reading the windows to the soul: Evidence of domain-specific sparing in Williams syndrome. *Journal of Cognitive Neuroscience*, *10*, 631–639.

Tager-Flusberg, H., & Plesa Skwerer, D. (in press). Social engagement in Williams syndrome. In P. J. Marshall & N. A. Fox (Eds.), *The development of social engagement: Neurobiological perspectives.* Oxford University Press Series in Affective Science. New York: Oxford University Press.

Tager-Flusberg, H., Plesa Skwerer, D., Faja, S., & Joseph, R. M. (2003). People with Williams syndrome process faces holistically. *Cognition*, *89*, 11–24.

Tager-Flusberg, H., & Sullivan, K. (2000). A componential view of theory of mind: Evidence from Williams syndrome. *Cognition*, *76*, 59–89.

Thomas, M. L., Becerra, A. M., & Mervis, C. B. (2002). *The development of empathy in 4-year-old children with Williams syndrome*. Long Beach, CA: 9th International Professional Conference on Williams Syndrome.

Udwin, O., Yule, W., & Martin, N. (1987). Cognitive abilities and behavioral characteristics of children with idiopathic infantile hypercalcemia. *Journal of Child Psychology and Psychiatry*, *28*, 297–309.

Wang, P. P., Doherty, S., Rourke, S. B., & Bellugi, U. (1995). Unique profile of visuo-perceptual skills in a genetic syndrome. *Brain and Cognition*, *29*, 54–65.

Whitman, B. Y., & Accardo, P. (1987). Emotional symptoms in Prader-Willi syndrome adolescents. *American Journal of Medical Genetics*, *28*, 897–905.

PCGN57-04

Queries

Daniela Plesa Skwerer, Alyssa Verbalis, Casey Schofield, Susan Faja, and Helen Tager-Flusberg

No queries