Understanding strategic information use during emotional expression judgments in Williams syndrome

Louise Ewing[1,3,4], Emily K. Farran[2], Annette Karmiloff-Smith[1], & Marie L. Smith[1]

[1]Department of Psychological Sciences, Birkbeck, University of London
[2]UCL Institute of Education, University College London
[3] School of Psychology, University of East Anglia
[4]ARC Centre of Excellence in Cognition and its Disorders, School of Psychology, University of Western Australia

Submitted to Developmental Neuropsychology

Corresponding author:
Dr Louise Ewing
School of Psychology
University of East Anglia
Norwich Research Park
NR4 7TJ

Date: 3rd July 2017

Word count: 5397

Keywords: face perception, emotional expressions, social motivation, reverse correlation, adults, Williams syndrome

Abstract

Detailed analysis of expression judgments in Williams syndrome reveals that successful emotion categorization need not reflect ‘classic’ information processing strategies. These individuals draw upon a distinct set of featural details to identify happy and fearful faces that differ from those used by typically developing comparison groups: children and adults. The diagnostic visual information is also notably less interlinked in Williams syndrome, consistent with reports of diminished processing of configural information during face *identity* judgments. These results prompt reconsideration of typical models of face expertise by revealing that an age-appropriate profile of expression performance can be achieved via alternative routes.

Face perception is widely characterized as a particular strength in Williams syndrome (WS). Despite cognitive impairments, including atypical spatial cognition and an average IQ of approximately 55 (Donnai & Karmiloff-Smith, 2000; Pober, 2010), individuals with WS perform well on a range of face recognition and face memory tests (Bellugi, Sabo, & Vaid, 1988; Tager-Flusberg, Plesa-Skwerer, Faja, & Joseph, 2003). There are reports of scores superior to mental-age (MA) matched comparison groups, and even within the normal range on some standardised measures, such as the Benton face perception test (e.g., Karmiloff-Smith et al., 2004), and Rivermead test of face memory (e.g., Udwin & Yule, 1991). This remarkable profile has led some to ask whether face processing might constitute an intact or preserved module in WS (Bellugi et al., 1988; Wang, Doherty, Rourke, & Bellugi, 1995), although this notion has since been strongly criticized (Karmiloff-Smith et al., 2004).

It is important to note that normative performance profiles on laboratory face processing tests need not reflect normative development of underlying expertise; these outcomes could be the product of different underlying processes. Indeed, current consensus suggests that unlike the typically developing (TD) population, for whom finely tuned holistic and/or configural processing supports expertise with this complex class of objects (Wang, Li, Fang, Tian, & Liu, 2012), individuals with WS preferentially process local or featural cues (Annaz, Karmiloff-Smith, Johnson, & Thomas, 2009; Deruelle, Mancini, Livet, Casse-Perrot, & De Schonen, 1999; Dimitriou, Leonard, Karmiloff-Smith, Johnson, & Thomas, 2014; Karmiloff-Smith et al., 2004; but see D'Souza, Booth, Connolly, Happé, & Karmiloff-Smith, 2015). A similar profile of reliance upon specific details (e.g., high spatial frequency cues around the mouth, eye-brow and jaw-line) was recently shown to be associated with immature and relatively poor face processing ability in the typically developing population (Ewing, Karmiloff-Smith, Farran, & Smith, 2017b). It follows then, that the strong performance reported in WS, if relying upon this atypical (sub-optimal) perceptual input, is intriguing. Moreover, importantly it challenges traditional conceptions of face processing expertise (Maurer, Le Grand, & Mondloch, 2002).

One intuitively appealing explanation for the unexpectedly strong performance on these face tasks in WS relates to participants' social motivation. Early motivation to view and accumulate experience with faces may critically facilitate the development of experience and expertise with this perceptually homogeneous category of visual stimuli. Indeed, selectively *diminished* social interest is argued to play a key causal role in face processing difficulties in autism spectrum disorder (ASD, Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012). Individuals with WS are often characterized as extremely friendly and gregarious, with a powerful motivation towards social interaction that may be difficult for them to inhibit (see Little et al., 2013). This ‘hypersocial’ personality profile seems to be present from early in development (Järvinen, Korenberg, & Bellugi, 2013), with a particular interest in looking at faces (Laing et al., 2002; Mervis et al., 2003) and atypically strong social approach behavior observed from early infancy (see Frigerio, 2006; Jones et al., 2000). Eye-tracking studies with adults have also empirically confirmed this strong interest in faces, and the eye region in particular (e.g., Riby & Hancock, 2008, 2009), which could systematically boost face experience across the lifespan. Therefore, following the social motivation account, individuals with WS could perform better with faces broadly (e.g., with respect to processing identity and other social cues) relative to mental-age matched peers because they show an atypical interest in these stimuli.

It is perhaps surprising then, that the processing of other facial attributes such as emotional expressions are not also a relative perceptual strength in WS. Despite anecdotal reports of heightened sensitivity to the emotions of others, the message from empirical research has been clear: skills in this domain are unremarkable. Admittedly, relatively fewer studies have investigated emotion processing in WS (cf. identity) but these laboratory tests of labelling (e.g., Gagliardi et al., 2003; Plesa-Skwerer, Faja, Schofield, Verbalis, & Tager-Flusberg, 2006), identification and matching (e.g., Lacroix, Guidetti, Rogé, & Reilly, 2009) of basic emotions using static and also dynamic stimuli (Martínez-Castilla, Burt, Borgatti, & Gagliardi, 2014) tend to report at best, only mental-age appropriate skills in the disorder. A recent EEG study investigating neural responses to differently valenced images (specifically stimuli with ‘positive and negative affective value’, which included happy and angry faces) also concluded that there was no evidence of significant differences in affective processing of social information in WS (Key & Duykens, 2016). This selectivity of face processing strengths in WS constitutes a potential challenge for a social motivation-based account of these abilities, which should benefit expertise broadly. That is, of course, unless these mental-age appropriate processing skills are observed *in spite* of processing atypicalities that are acting to diminish processing efficiency.

There is some evidence to support the notion that emotional expression processing is atypical, but well-compensated for in WS. A recent study contrasted the developmental trajectories of behavioural skills in WS and typically developing participants (Martínez-Castilla et al., 2014). The selective absence of a relationship between accuracy and mental age in participants with WS (i.e., trajectory with a gradient of zero when mental age was entered as a predictor) signaled that facial emotion recognition may reach its maximal developmental level atypically early and then plateau due to cognitive constraints. Reports also exist of atypical visual scanpaths during passive viewing of emotional faces: longer looking at the eyes than do mental age-matched controls for happy, angry and fearful faces (cf. similar viewing times across groups for the nose and mouth) (Porter, Shaw, & Marsh, 2010). Interestingly, individuals with WS also spent relatively less time looking at ‘other’ things: non-salient features, such as cheeks, hair, ears, the image background (Porter et al., 2010). Imaging studies have also observed atypical neural reactivity to at least some emotional faces in WS. Unusual profiles of orbitofrontal cortical activity have been reported during WS emotion discrimination (matching tasks involving fearful and angry faces, Meyer-Lindenberg et al., 2005; and happy and angry faces, Mimura et al., 2010) and activation of the amygdala appears attenuated when viewing fearful faces and heightened when viewing happy faces (Haas et al., 2010; Haas et al., 2009; Meyer-Lindenberg, Mervis, & Berman, 2006). Such differences relative to comparison groups constitute preliminary evidence that expression perception in WS is more atypical than suggested by processing ability metrics.

A more detailed characterization of emotion processing in WS is required in order to assess how precisely emotional information is read from faces, and the extent to which this might differ from typically developing age and mental-ability matched individuals. The ‘bubbles’ reverse correlation technique (e.g., Gosselin & Schyns, 2001) provides a particularly elegant and powerful means to characterise expression processing strategies, i.e., the specific visual information driving face judgments. Moreover, recent methodological advances have broadened the application of this approach to non-traditional participant samples, e.g., children as young as 6 years (Smith, Letizia Cesana, Farran, Karmiloff-Smith, & Ewing, 2017). By examining participants’ ability to categorise emotional faces when provided with only subsampled stimuli across trials – having to rely on specific subsets of information - we can determine the diagnostic visual cues for deciding when a face looks scared, sad, happy or angry.

This approach has been used productively to investigate expression processing strategies across a range of populations. Studies with typically developing adults (Smith, Cottrell, Gosselin, & Schyns, 2005) and children (Ewing, Karmiloff-Smith, Farran, & Smith, 2017a) indicate that from the youngest ages tested, participants flexibly and strategically draw upon distinct visual information for different expressions, e.g., relying on the wide-open eyes to accurately identify fear, the mouth for happy and the furrowed brow for anger. These particular visual cues may be optimized to discriminate between emotion categories, so that 'tuning in' to them supports processing efficiency (Smith et al., 2005). The current study is the first to use the bubbles paradigm with participants with WS. We hypothesised that it would be sufficiently sensitive to reveal the presence of any atypicalities in information-use during expression judgments in individuals with WS.

INSERT TABLE ONE ABOUT HERE

Methods

Participants

19 individuals with Williams' syndrome were recruited via the Williams Syndrome Foundation UK. All had a phenotypic diagnosis of WS from a clinician as well as based on a 'fluorescent in situ hybridization' (FISH) test for the deleted Elastin gene. One participant was removed due to insufficient cognitive capacity to complete the experimental task, leaving 18 participants (6 male, aged 18 to 50 years). See Table 1 for further age and cognitive characterization based on the British Picture Vocabulary Scale III (BPVS-III, Dunn, Whetton, & Pintilie, 1997) and the Ravens Coloured Progressive Matrices (RCPM, Raven, 1993).

Eighteen typically developing adults were matched individually to the WS participants on gender and chronological age (6 male, age M=28;4, SD=7;8, $t(34) = 0.40$, p= 0.68). This comparison group should have had the opportunity to accumulate a similar amount of lifetime experience with faces as the clinical sample. These participants completed a longer version of the experimental task, with only the first 192 trials reported here (to match the clinical group's number of trials). Given the level of cognitive ability in the WS group, we also included a comparison group of typically developing children. Here a large sample was recruited and tested (N=177) aged 6 to 12 years from primary schools in London (UK) and Perth (Australia) to first establish the age-group that most closely resembled the quantitative performance profile of the WS group on the experimental task, full details below. The final group selected for use in our analysis were 11-year olds (N=31, 20 male, mean age 11;5 months, SD=0;4).

Stimuli

Two male identities from the California Facial Expressions Database each displayed fear, sadness, happiness and anger (following the Facial Action Coding System; Ekman, 1978) with standardised locations of the eyes and the mouth (Dailey, Cottrell, & Reilly, 2001). [1]

Participants were required to categorise subsampled versions of these faces (see Gosselin & Schyns, 2001 for full methodological details). Briefly, these images were decomposed into five non-overlapping spatial frequency (SF) bandwidth scales of one octave (120–60, 60–30, 30–15, 15–7.5 and 7.5–3.8 cycles/image; remaining bandwidth = constant background). Information was sampled independently on each trial from each SF scale via randomly positioned circularly symmetric Gaussian apertures (referred to as 'bubbles'). We ensured equivalent information sampling across each SF scale by adjusting the number and size of these bubbles on each trial (6 cycles/SF scale). The sampled information from each band was recombined into one image for each trial stimulus, which thus comprised high, mid and low SF information in randomly determined locations (see Supplementary Figure 1). Stimuli appeared centrally on a grey background for 1000ms at a viewing distance of 70cm (subtending a visual angle of approximately 6.94°x 4.42°).

We used a staircase algorithm to independently adjust the sampling density (number of bubbles) for each expression and each participant to maintain 75% accuracy in expression categorisations. That is, when participants performed well (more than 75%) they would have less information (fewer bubbles) presented to them on trials, whereas when they performed poorly (less than 75%) they would have more information (more bubbles) presented, to a maximum of 200. In this way we made every effort to ensure that the task was comparably

---

[1] Note that Smith et al., 2005 found no significant difference in the diagnostic information for expression judgments across male and female faces.

challenging for all participants.

Procedure

We presented the task to participants as ‘The Puzzle Bubble Game’, which was completed in 15-20 minutes by adults and children (in quiet testing rooms at university or school) and approximately 30 minutes by individuals with Williams syndrome (at home). The aim of the game was to identify the feeling shown on each face when “cheeky puzzle bubbles” concealed parts from view[2] to make this job tricky. Stimuli were presented for 1000ms and preceded by a fixed duration fixation cross of 500ms. There was no fixed ISI beyond the fixation cross duration and any time taken for response (when it exceeded the stimulus presentation duration). Responses (no time limit) were made verbally to the experimenter or via a labelled keyboard press (labels=photographs of the 4 expressions displayed by a third model as well as a question mark for an ‘I don’t know’ response). The experimenter sat with each participant at all times and maintained attention and motivation by providing enthusiastic praise and encouragement.

Participants were provided with training before the main test trials. That is, we ensured that they could identify the expressions with at least 75% accuracy when displayed intact (without bubbles) first for an unlimited time, and then for a fixed time of 1000ms to prepare them for the test trials. Auditory performance feedback was provided for correct and incorrect responses during this training phase only. They then completed 192 bubbles test trials: 8 blocks of 24 trials. Three participants with Williams syndrome omitted the last block due to time constraints. Blocks were separated by generic “keep up the good work” screens (odd numbered blocks) and a brief entertaining game called the Puzzle Bubble Challenge (even numbered blocks). Here, participants guessed the name of a film, geographical location, or television program depicted in a masked image, progressively revealed via bubbles as the experimenter provided additional clues/information.

Results

WS and Typically Developing Adults

Performance. Performance accuracy was measured via hits (the correct emotional expression being selected as the response) and incorrect responses (any incorrect emotional

---

[2] For ease of explanation, we told participants that the bubbles concealed things from view; however, in truth the bubbles more accurately *revealed* information (like windows through an opaque mask).

expression or the don't know response being selected). By design the Bubbles paradigm updates task difficulty on a trial per trial basis to target a 75% accuracy criterion for each condition across the experiment. The modest number of trials appropriate for each participant in this particular experiment (48 trials per expression category cf. typically many hundreds when using this approach) led to a spread of results around this target. To explore these variations in performance accuracy, a two-way repeated measures ANOVA investigated the effects of emotion (fear, sadness, happiness, anger) and participant group (adults, WS) on performance accuracy[3]. A main effect of emotion, $F(2.34, 79.69) = 40.78$, $p<0.001$, $\eta_p^2 = 0.54$, confirmed significant differences in performance accuracy for each emotion with happy being the easiest (*M*=83.8%, *SD*=4.1%) and anger the most difficult (*M*=57.4%, *SD*=12.1%), see Figure 1. There was no main effect of participant group, $F(1,34) = 0.59$, $p=0.44$, $\eta_p^2= 0.01$, but there was a marginally significant interaction between participant group and emotional expression, $F(2.34, 79.69) = 2.74$, $p=0.06$, $\eta_p^2 = 0.07$. Subsequent independent samples t-tests confirmed a trend for a significant difference in performance for the categorization of sad facial expressions only, $t(29.08) = 1.76$, $p=0.08$, with typically developing adults (M=73.1%, SD=12.2%) outperforming than the WS group (M=63.8%, SD=18.9%). Otherwise performance accuracy was well-matched across the groups for happy, $t(34) = 0.03$, $p=0.97$) and fear faces, $t(22.2) = 0.89$, $p=0.38$. Angry performance was also matched – but very low - across groups, $t(34) = 0.65$, $p=0.52$.

Another key performance metric is the median number of bubbles participants required to reach mean accuracy levels for each emotion. In order to achieve matched levels of performance in the categorization task reported above, the adaptive staircase algorithm modulated task difficulty by providing poor-performing participants with more information (more bubbles, making the task easier) and well-performing participants with less information (fewer bubbles, making the task more difficult). A two-way ANOVA investigating the influence of emotion and participant group on this variable confirms a main effect of participant group ($F(1,34) = 40.0$, $p<0.001$, $\eta_p^2 = 0.54$) as well as a main effect of emotion ($F(2.29, 77.90) = 61.86$, $p<.001$, $\eta_p^2 = 0.64$) that mirrors performance accuracy differences (most bubbles for angry, least bubbles for happy). There was also a significant

---

[3] Note that one participant from the WS group had a severe and selective impairment for fearful facial expressions only, registering only 12% correct across the course of the experiment in this one condition whilst performance for the remaining three expressions was well above chance (M=58% correct). Similarly a second WS individual was selectively poor with anger (11% correct). Group average data has been used in place of the individual values for these two cells of the ANOVA.

group x emotion interaction ($F(2.29, 77.90) = 4.16$, $p= 0.015$, $\eta_p^2 = 0.10$). For happy and fearful categorizations, where performance accuracy is matched, WS individuals require significantly more information than TD adults (happy, $t(34) = 5.57$, $p<0.001$, M=88.5(19.9) vs. 56.1(14.4); fear, $t(34) = 5.56$, p<0.001, M=123.1(29.1) vs. 77.9(18.4)). For sadness, where there was a trend for the WS group to perform worse than their adult counterparts, they again used significantly more information ($t(34) = 5.47$, $p<0.001$; WS 138.5(28.7) bubbles vs. TD 88.9(25.4) bubbles). Finally, for anger, where performance accuracy was matched, but very low for both groups, there was also a trend for the WS individuals to require more information ($t(34) = 1.94$, $p=0.06$; WS 146.0(27.5) bubbles vs. TD 127.4(29.7) bubbles)[4].

INSERT FIGURE ONE ABOUT HERE

Information Use. The group differences and poor performance for sadness for the WS group (one-sample t-test for sadness vs. 75% correct: $t(17) = 2.51$, $p=0.02$) and poor performance for anger for both groups (one sample t-tests anger vs. 75% correct: WS, $t(17) = 3.87$, $p=0.001$, TD adults $t(17) = 6.11$, $p<0.001$), necessarily result in a noisier information-map solution for these categories. We therefore restrict our subsequent analysis to happy and fear emotional expressions where performance was matched and both groups attained the set levels of performance.

For each participant group, and each expression, trials were divided into two categories: those in which information presented resulted in correct vs. incorrect categorizations (‘don’t know’ was considered incorrect[5]). Then we summed the bubble masks associated with correct categorizations for fear and for happiness and divided each of these

---

[4] Pilot studies which used equivalent levels of information sampling in a separate adult control group (N=14 completed an identical task to the WS group) resulted in a group level significant difference in performance accuracy, $F(1,30) = 5.8$, $p=0.023$). That is, when the number of bubbles was matched between groups, performance accuracy was not. These new typical adults performed extremely well (79% correct on average): outperforming relative to the WS group, as well as the 75% accuracy criterion target because the task was simply too easy for them.

[5] Note that very few participants made use of the ‘don’t know’ option, with no difference between the groups: median use in WS: 2.9% vs. typical adults: 3.7%, U = 148.5 (Z=.43), p=0.67.

by the sum of all bubble masks presented for each emotion to generate probability maps revealing the specific information driving correct categorizations. The pixel value at each location represents the probability that presenting visual information at that location will facilitate a correct response. To maximise data collection with this atypically brief version of the paradigm we sampled information only from the face images (as per Smith et al, 2005) and not from any surrounding completely non-informative region. We applied a $p<.05$ peak threshold and cluster size criterion (z-critical: 3.96, 3.59, 3.18, 2.76, 2.37 from high to low SF scale respectively, threshold t-value: 2.7) to establish those regions that were statistically associated with correct categorisation performance, henceforth termed diagnostic information or diagnostic features. Please see Chauvin, Worsley, Schyns, Arguin and Gosselin (2005) for full details of the specially designed statistical tests for classification images which are based on the fMRI standard Random Field Theory approach to correcting for multiple comparisons. In order to conduct these tests a baseline region must be selected where no difference in information use are expected. Here the sampled non-face regions around the face image (e.g. the neck, the hairline) formed this region. To visualize the resulting diagnostic information we selected a representative emotional face for each emotion and revealed only the information significantly associated with correct categorization performance (Figures 2 to 4).

Regions that are significantly correlated with correct facial expression categorization performance are shown in Figure 2 (in red for individuals with WS and in blue for adult controls). Information that is used significantly by both groups is purple. Diagnostic images (bounded by red or blue boxes) represent only the significant information, combined across the spatial frequency scales, for each expression categorization and participant group.

INSERT FIGURE TWO ABOUT HERE

It is immediately apparent that there is very little information use overlap between individuals with WS and typically developing adults for these two expressions. For fear categorisations the profiles of the two groups are quite distinct and for happiness there is shared use of only some fine details around the mouth and central portion of the face. Broadly speaking, the typically developing adult data is very much in line with existing findings (e.g., Smith et al., 2005; Smith & Merlusca, 2014), whereas the features used by the individuals with WS are more unusual. Typically developing adults rely upon the wide opened eyes and coarse mouth information when categorizing fear, and the broad smiling mouth outline alongside fine details around the corners of the mouth for happy. By contrast, for the same judgments, we reveal here that individuals with WS draw upon multiple, fine-grained cues like creases in the forehead and around the chin for fear and show a left-lateralised reliance upon mid-spatial frequency eye and mouth information for their happy categorisations.

WS and Typically Developing Children

Performance. We report above that the online calibration of bubble numbers (to equate categorization performance accuracy across groups) resulted in a significant difference in the amount of information sampled for typically developing adults and those with WS. To address the possibility that it was *this* difference that was driving the observed atypicalities in the WS group, we also contrasted their profile of information use with a group matched for both performance accuracy *and* number of bubbles: children.

We explored task performance across a range of ages: testing 177 children aged from 6 to 12 years to determine the most appropriate comparison group. A one-way ANOVA to investigate the influence of age (years: 6,7,8,9,10,11,12) on the median number of bubbles supported developmental differences in the amount of information required to accurately categorize the two key expressions investigated with adults: a marginally significant age effect for fear ($F(6, 176) = 2.05$ $p=0.06$) and a significant effect for happiness ($F(6, 176) = 5.92$, $p<.001$. Follow-up tests contrasting information use in each age-group relative to the WS group identified 11 year olds as the optimal age-group within the assessed range. This group distribution was well-matched to the WS participants (above the p=0.50 criterion proposed by Mervis & Klein-Tasman, 2004) with no significant group difference in the number of bubbles required for our key judgments: fear ($t(38.41) = 0.10$, $p=0.91$) or happiness ($t(24.01) = -0.61$, $p=0.54$).

Figure 3 shows contrasts the key performance metrics in these two participant groups. For completeness, a repeated measures ANOVA investigating the effects of emotion (fear, sadness, happiness, anger) and participant group (typically developing 11 year olds – hereafter referred to as typically developing children, WS) on number of bubbles confirmed no significant main effect of group ($F(1,47) = .57$, $p=0.45$, $\eta_p^2 = 0.01$) or interaction with group ($F(1,47) = 2.38$, $p=0.12$, $\eta_p^2 = 0.04$). A marginally significant effect of emotion ($F(1,47) = 3.74$, $p=.059$, $\eta_p^2 = 0.07$) reflected numerically more bubbles required for categorisations of sadness (*M*=130.41, *SD*=29.63) compared to fear (*M*=123.67, *SD*=30.6). The same pattern of effects was observed in an equivalent ANOVA investigating categorization accuracy: percent correct. That is, a marginally significant effect of emotion ($F(1, 47) = 3.16$, p=0.08, $\eta_p^2 = 0.06$) reflected better categorization performance for fear (*M*=72.3, *SD*=15.0) than sadness (*M*=69.5, *SD*=16.2) but crucially, there was no main effect of participant group (F1,47) = 1.94, p=0.16, $\eta_p^2 = 0.04$) or interaction with group ($F(1, 47)=2.59$, p=0.11, $\eta_p^2 = 0.05$). For consistency with the adult results reported above, the remaining bubbles analyses will be restricted to happy and fearful expressions.

INSERT FIGURE THREE ABOUT HERE

Information Use. Figure 4 provides a visual representation of the face regions significantly correlated with correct categorization for the WS group (again in red) and the matched typically developing children (now those in blue). Similarly to the adults, and previous published results, these children used the wide-open eyes for fear, along with greater use of coarser mouth and nose information. For happy faces, they primarily employed the broad smiling mouth in the lower spatial frequency bands. Importantly, there is no evidence that it was the greater number of bubbles (required by both children and WS individuals) that led to the abnormal piecemeal strategies employed by the individuals with WS observed in Figure 2.

INSERT FIGURE FOUR ABOUT HERE

Discussion

The current findings provide important new insights into emotion processing in WS: a relatively understudied aspect of face perception in this group. Despite previous reports of mental-age appropriate performance on expression tasks, here for the first time we reveal a distinct and atypical profile of information-use during these judgments. Like typically developing participants, they flexibly calibrated their information-use with task requirements, i.e., drawing upon different visual features when categorizing different emotional expressions (happy and fearful faces). Crucially, however, the specific cues they relied upon for the judgments differed from those traditionally associated with expertise in this domain.

We compared the information processing strategies of individuals with WS to those of two different groups: typically developing adults and children. Both 'baselines' offered important insights into the clinical profile. The adult data (chronologically age-matched to the WS group) provided a picture of the diagnostic information driving expression judgments in individuals with a similar level of lifetime experience with faces. By contrast the eleven-year-olds' data provided the profile of a group matched for expression reading ability (metrics taken directly from the experimental paradigm: bubbles task). Results revealed that the bubbles solutions in WS did not resemble that of either group, which signals that the observed atypicalities reflect more than just an immature expression recognition system or differences in experience with faces.

Typically developing adults demonstrated a 'classic' profile of strategic information use during their emotion judgments. For accurate judgments of fear, for example, results indicated that the significantly diagnostic information was the wide eyes and lower SF shape information around the mouth. For happiness we observed a focused reliance upon the smiling mouth region. This consistency with previous research using more exhaustive testing protocols (e.g., Smith et al., 2005; Smith & Merlusca, 2014) is useful both in providing a robust behavioral benchmark in the current study, and also validating our developmental adaptation of the standard bubbles paradigm (see Smith et al., 2017).

Information-use in typically developing children was qualitatively similar to that observed in the adults. The key difference was that their strategy broadly resembled a less refined version of the adult profile. Assuming strategic information use plays a functional role in typical face expertise (Ewing et al., 2017b) then such a profile aligns with the relatively less efficient categorization performance we observed in this group. That is, relative to adults, the 11 year olds required substantially more face information (bubbles) to reach the 75% performance accuracy criterion. These results are consistent with the notion that the development of strategic information use with age contributes to face processing efficiency, e.g., through increased experience with faces specifically, or perhaps more generic changes in visual perception.

Relative to both comparison groups, participants with WS drew upon a wider and less integrated range of information. There was only limited overlap in the significantly diagnostic features, e.g., shared use of fine details around the central portion of the face and mouth for happiness. Moreover, individuals with WS showed a singular reliance upon several fine-grained facial cues, such as creases around the forehead and chin for judgments of fear, and the left eye region for happiness. Though the information value of such cues is not hard to fathom, neither were used by typically developing children or adults. These results suggest that the typical profile of strategic information-use is not the only formula for relatively strong performance during expression categorization judgments.

These results are consistent with another instance of individuals with WS accomplishing somewhat normative performance outcomes via an unusual processing route (see Karmiloff-Smith, 2009). For example, despite broadly age-appropriate (i.e., strong) face identity processing abilities in WS, there have been reports of selectively reduced perceptual sensitivity to highly informative configural (cf. featural) information (Annaz et al., 2009; Karmiloff-Smith et al., 2004). These results suggest an atypical, relatively piecemeal approach to identity judgements. Such a strategy is consistent with the detail-focused visual processing style observed across other tasks in WS (Bellugi, Lichtenberger, Jones, Lai, & St George, 2000; Farran, Jarrold, & Gathercole, 2003 but see D'Souza et al., 2015) but at odds with good performance on face identity processing tasks, which are widely considered to crucially tap ‘global’ processing skills. Holistic processing may be a hallmark of *typical* face expertise (Richler, Cheung, & Gauthier, 2011; Wang et al., 2012), but there appears to be scope for individuals with WS to achieve strong performance outcomes using an alternative approach.

The atypical expression strategies revealed in the current study with the bubbles paradigm could be functioning to limit WS performance outcomes on tasks that involve emotion judgments. The typical profile of strategic information-use has been shown to be optimized for discriminating between (at least) the basic emotional expressions (Smith et al., 2005). It follows then, that deviations from this normative profile should come at a cost to efficiency and/or accuracy. If this attenuation is occurring, then the widely observed mental age-appropriate outcomes on expression tasks in WS might actually constitute an impressive processing feat, i.e., if achieved in spite of a suboptimal strategy.

It is again tempting to speculate that the strong social motivation in WS functions to boost face processing abilities in this group. The extent to which this mechanism generates only mental-age appropriate expression skills rather than particularly strong (chronologically age-appropriate) skills observed for identity could signal that the former start from a lower baseline. It is not intuitively obvious why this would be the case, given that any performance decrements associated with diminished sensitivity to global information should impact upon identity judgments as much (if not more) than expression judgments (Richler et al., 2011).

The direct investigation of identity-related patterns of information-use would be extremely interesting, and potentially revealing on this point. Unfortunately, pilot testing from our laboratory suggests that identity categorization tasks (with unfamiliar and therefore necessarily ‘learned’ face identities) can be extremely cognitively taxing and challenging for this participant group.

The current results reveal that the power of reverse correlation paradigms to provide detailed information about information processing during a given task is not limited to typically developing groups. Still, our study has its limitations. We could not clearly interpret the results for judgments of sad and angry faces because of difficulties obtaining comparison groups that matched the low performance levels in the WS group for these expression categories. Additionally, the challenges associated with running more than the current, relatively limited number of trials with these participants (most critically, cognitive fatigue) necessarily limited the power and range of possible analyses. We maintain that the extent to which our Bubbles information use results clearly align with established findings warrants confidence in the observed results (see also Smith et al., 2017). Still it was a shame, for example, there was not scope to apply signal detection analyses and meaningfully explore any possible contributions of response criterion differences as well as perceptual sensitivity - between the groups.

More broadly associated with the reverse correlation approach, it is worth mentioning that by artificially presenting only subsets of information from each expressive face to categorize, questions can be asked regarding the generalizability of the findings to full faces. It is, however, important to note that the relatively large number of bubbles required by the WS individuals, alongside the use of a 3D sampling scheme that takes information across the spatial frequency scales to build each stimulus, meant that in this particular case a fairly large proportion of face information was presented to participants on each trial. Critically, this highly controlled task has provided the first insight into expression processing strategies in WS, revealing atypicalities that were formerly masked by mental-age appropriate outcomes on standard tasks. These results also raise some questions for typical models of face expertise by indicating that mental age-appropriate expression reading outcomes can be achieved via the use of atypical processing routes.

Acknowledgements

This research was supported by a grant from the Leverhulme Trust: RPG-2013-019 awarded to MLS, EF and AKS. Additional support was generously provided by Gillian Rhodes and the Person Perception Node of the Australian Research Centre of Excellence in Cognition and its Disorders. We thank Michael Papasavva, Inês Mares, Hayley White, Cathy Wong, Charles Wigley and Marguerite Cullity for assistance with data collection as well as the participants, their families and the schools for the time and effort expended in these studies.

References

Annaz, D., Karmiloff-Smith, A., Johnson, M. H., & Thomas, M. S. (2009). A cross-syndrome study of the development of holistic face recognition in children with autism, Down syndrome, and Williams syndrome. *Journal of Experimental Child Psychology, 102*(4), 456-486.

Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., & St George, M. (2000). The neurocognitive profile of Williams syndrome: a complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience, 12,* 7 - 29.

Bellugi, U., Sabo, H., & Vaid, J. (1988). Spatial deficits in children with Williams syndrome. In J. Stiles-Davis, M. Kritchevsky, & U. Bellugi (Eds.), *Spatial Cognition: Brain Bases and Development* (pp. 273-298). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chauvin, A., Worsley, K., Schyns, P., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision, 5*, 659 - 667.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E., & Schultz, R. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences, 16*, 231 - 239.

Dailey, M., Cottrell, G. W., & Reilly, J. (2001). *California facial expressions (CAFE). Unpublished digital images*.

Deruelle, C., Mancini, J., Livet, M., Casse-Perrot, C., & De Schonen, S. (1999). Configural and local processing of faces in children with Williams syndrome. *Brain and Cognition, 41*(3), 276-298.

Dimitriou, D., Leonard, H., Karmiloff-Smith, A., Johnson, M. H., & Thomas, M. S. (2014). Atypical development of configural face recognition in children with autism, Down syndrome and Williams syndrome. *Journal of Intellectual Disability Research*, doi: 10.1111/jir.12141.

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

Dunn, L. M., Whetton, C., & Pintilie, D. (1997). *British Picture Vocabulary Scale*. Windsor: NFER-Nelson.

Ekman, P., & Friesen, W. V. (1978). . . (1978). *Manual for the facial action coding system*: Consulting Psychologists Press.

Ewing, L., Karmiloff-Smith, A., Farran, E., & Smith, M. L. (2017a). Developmental changes in the critical information used for facial expression processing. *Cognition, in press*.

Ewing, L., Karmiloff-Smith, A., Farran, E., & Smith, M. L. (2017b). A distinct profile of information-use characterizes face identity judgments in children and low-expertise adults *Journal of Experimental Psychology: Human Perception and Performance, in press*.

Farran, E., Jarrold, C., & Gathercole, S. E. (2003). Divided attention, selective attention and drawing: processing preferences in Williams syndrome are dependent on the task administered. *Neuropsychologia, 41*, 676-687.

Frigerio, E., Burt, D. M., Gagliardi, C., Cioffi, G., Martelli, S., Perrett, D. I., & Borgatti, R. (2006). Neuropsychologia, 44(2), 254-259. (2006). Is everybody always my friend? Perception of approachability in Williams syndrome. . *Neuropsychologia, 44*, 254 - 259.

Gagliardi, C., Frigerio, E., Burt, D. M., Cazzaniga, I., Perrett, D. I., & Borgatti, R. (2003). Facial expression recognition in Williams syndrome. *Neuropsychologia, 41*(6), 733-738.

Gosselin, F., & Schyns, P. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research, 41*(17), 2261-2271.

Haas, B. W., Hoeft, F., Searcy, Y. M., Mills, D., Bellugi, U., & Reiss, A. (2010). Individual differences in social behavior predict amygdala response to fearful facial expressions in Williams syndrome. *Neuropsychologia, 48*, 1283 - 1288.

Haas, B. W., Mills, D., Yam, A., Hoeft, F., Bellugi, U., & Reiss, A. (2009). Genetic influences on sociability: heightened amygdala reactivity and event-related responses to positive social stimuli in Williams syndrome. *Journal of Neuroscience, 29*, 1132 - 1139.

Järvinen, A., Korenberg, J., & Bellugi, U. (2013). The social phenotype of Williams syndrome. . *Current Opinion in Neurobiology, 23*, 414 - 422.

Jones, W., Bellugi, U., Zona, L., Chiles, M., Reilley, J., Lincoln, A., & Adolphs, R. (2000). Hypersociability in Williams syndrome. *Journal of Cognitive Neuroscience, 12*, 30 - 46.

Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: rethinking the study of developmental disorders. *Developmental Psychology, 45*, 56 - 63.

Karmiloff-Smith, A., Thomas, M., Annaz, D., Humphreys, K., Ewing, S., Brace, N., . . . Campbell, R. (2004). Exploring the Williams syndrome face-processing debate: the importance of building developmental trajectories. *Journal of Child Psychology and Psychiatry, 45*(7), 1258-1274.

Key, A., & Duykens, E. (2016). Processing of stimulus content but not of emtional valence is altered in persons with WIlliams syndrome. *Journal of Intellectual Disability Research, 60,* 993 - 1009.

Lacroix, A., Guidetti, M., Rogé, B., & Reilly, J. (2009). Recognition of emotional and nonemotional facial expressions: A comparison between Williams syndrome and autism. . *Research in Developmental Disabilities, 30,* 976-985.

Laing, E., Butterworth, G., Ansari, D., Gsödl, M., Longhi, E., Panagiotaki, G., . . . Karmiloff-Smith, A. (2002). Atypical development of language and social communication in toddlers with Williams syndrome. *Developmental Science, 5,* 233 - 247.

Martínez-Castilla, P., Burt, M., Borgatti, R., & Gagliardi, C. (2014). Facial emotion recognition in Williams syndrome and Down syndrome: A matching and developmental study. *Child Neuropsychology*, 1-25.

Maurer, D., Le Grand, R., & Mondloch, C. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6,* 255 - 260.

Mervis, C., & Klein-Tasman, B. (2004). Methodological issues in group-matching designs: alpha levels for control variable comparisons and measurement characteristics of control and target variables. . *Journal of Autism and Developmental Disorders, 34,* 7 - 17.

Mervis, C. B., Morris, C. A., Klein-Tasman, B. P., Bertrand, J., Kwitny, S., Appelbaum, L. G., & Rice, C. E. (2003). Attentional characteristics of infants and toddlers with Williams syndrome during triadic interactions. *Developmental Neuropsychology, 23*(1-2), 243-268.

Meyer-Lindenberg, A., Hariri, A., Munoz, K., Mervis, C. B., Mattay, V., Morris, C., & Berman, K. F. (2005). Neural correlates of genetically abnormal social cognition in Williams syndrome. *8*(991 - 993).

Meyer-Lindenberg, A., Mervis, C. B., & Berman, K. F. (2006). Neural mechanisms in Williams syndrome: a unique window to genetic influences on cognition and behaviour. *Nature Reviews Neuroscience, 7,* 380 - 393.

Mimura, M., Hoeft, F., Kato, M., Kobayashi, N., Sheau, K., Piggot, J., . . . Reiss, A. L. (2010). A preliminary study of orbitofrontal activation and hypersociability in Williams Syndrome. . *Journal of Neurodevelopmental Disorders, 2,* 93 - 98.

Plesa-Skwerer, D., Faja, S., Schofield, C., Verbalis, A., & Tager-Flusberg, H. (2006). Perceiving facial and vocal expressions of emotion in individuals with Williams syndrome. *American Journal on Mental Retardation, 111*(1), 15 - 26.

Pober, B. R. (2010). williams–Beuren syndrome. *New England Journal of Medicine, 362,* 239 - 252.

Porter, M., Shaw, T. A., & Marsh, P. J. (2010). An unusual attraction to the eyes in Williams-Beuren syndrome: a manipulation of facial affect while measuring face scanpaths. . *Cognitive Neuropsychiatry, 15*(6), 505 - 530.

Raven, J. (1993). *Coloured progressive matrices*. Oxford: Information Press Ltd.

Riby, D., & Hancock, P. (2008). Viewing it differently: Social scene perception in Williams syndrome and Autism. . *Neuropsychologia, 46,* 2855 - 2860.

Riby, D., & Hancock, P. (2009). Do faces capture the attention of individuals with williams syndrome or autism? Evidence from tracked eye movements. . *Journal of Autism and Developmental Disorders, 39,* 421 - 431.

Richler, J., Cheung, O., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science, 22,* 464 - 471.

Smith, M. L., Cottrell, G., Gosselin, F., & Schyns, P. (2005). Transmitting and decoding facial expressions. *Psychological Science, 16*(3), 184-189.

Smith, M. L., Letizia Cesana, M., Farran, E., Karmiloff-Smith, A., & Ewing, L. (2017). A ‘spoon full of sugar’ helps the medicine go down and the experimental outcomes up: How a participant friendly version of a psychophysics task significantly improves task engagement, performance and data quality in a typical adult sample. *Submitted*.

Smith, M. L., & Merlusca, C. (2014). How task shapes the use of information during facial expression categorisations. . *Emotion, 14,* 478 - 487.

Tager-Flusberg, H., Plesa-Skwerer, D., Faja, S., & Joseph, R. M. (2003). People with Williams syndrome process faces holistically. *Cognition, 89*(1), 11-24.

Udwin, O., & Yule, W. (1991). A cognitive and behavioural phenotype in Williams syndrome. *Journal of Clinical and Experimental Neuropsychology, 13*(2), 232-244.

Wang, P., Doherty, S., Rourke, S. B., & Bellugi, U. (1995). Unique profile of visuo-perceptual skills in a genetic syndrome. *Brain and Cognition, 29*(1), 54-65.

Wang, R., Li, J., Fang, H., Tian, M., & Liu, J. (2012). Individual differences in holistic processing predict face recognition ability. *Psychological Science*, *23*(2), 169-177.