Characterizing the neural signature of face processing in Williams syndrome via multivariate pattern analysis and event related potentials

Emily K. Farran [a],[**],[1], Inês Mares [b],[*],[1], Michael Papasavva [b], Fraser W. Smith [c], Louise Ewing [c],[2], Marie L. Smith [b],[d],[2]

[a] School of Psychology, University of Surrey, UK
[b] Department of Psychological Sciences, Birkbeck College, University of London, UK
[c] School of Psychology, University of East Anglia, Norwich, UK
[d] Centre for Brain and Cognitive Development, Birkbeck College, University of London, UK

ARTICLE INFO

ABSTRACT

Face recognition ability is often reported to be a relative strength in Williams syndrome (WS). Yet methodological issues associated with the supporting research, and evidence that atypical face processing mechanisms may drive outcomes 'in the typical range', challenge these simplistic characterisations of this important social ability. Detailed investigations of face processing abilities in WS both at a behavioural and neural level provide critical insights. Here, we behaviourally characterised face recognition ability in 18 individuals with WS comparatively to typically developing children and adult control groups. A subset of 11 participants with WS as well as chronologically age matched typical adults further took part in an EEG task where they were asked to attentively view a series of upright and inverted faces and houses. State-of-the-art multivariate pattern analysis (MVPA) was used alongside standard ERP analysis to obtain a detailed characterisation of the neural profile associated with 1) viewing faces as an overall category (by examining neural activity associated with upright faces and houses), and to 2) the canonical upright configuration of a face, critically associated with expertise in typical development and often linked with holistic processing (upright and inverted faces). Our results show that while face recognition ability is not on average at a chronological age-appropriate level in individuals with WS, it nonetheless appears to be a relative strength within their cognitive profile. Furthermore, all participants with WS revealed a differential pattern of neural activity to faces compared to objects, showing a distinct response to faces as a category, as well as a differential neural pattern for upright vs. inverted faces. Nonetheless, an atypical profile of face orientation classification was found in WS, suggesting that this group differs from typical individuals in their face processing mechanisms. Through this innovative application of MVPA, alongside the high temporal resolution of EEG, we provide important new insights into the neural processing of faces in WS.

1. Introduction

The ability to process faces is critically important for successful social interaction. In the typical population, not only are we able to effortlessly recognise a friend, but we can evaluate faces for social cues such as emotion and trustworthiness. This face-selective expertise emerges early in development, with a preference for attending to the properties of a face over a non-face, observed from birth (Johnson et al., 1991). For

individuals with Williams syndrome (WS), who present with learning difficulties in the mild to moderate range (~IQ: 55; Udwin and Yule, 1991), face processing is widely reported as a relative strength (see Riby, 2012). That is, whilst poor visuo-spatial abilities are a hallmark weakness in this group (Mervis et al., 2000), they show relative proficiency in the face processing domain (e.g., Bellugi et al., 1988). Some have attributed these abilities, at least in part, to the heightened interest in faces observed in WS (see Riby, 2012, for an evaluation of this

---

* Corresponding author. Department of Psychological Sciences, Birkbeck College, University of London, London, WC1E 7HX, UK.
** Corresponding author. School of Psychology, University of Surrey, Guildford, Surrey, GU2 7XH, UK.
E-mail addresses: e.farran@surrey.ac.uk (E.K. Farran), imares01@mail.bbk.ac.uk (I. Mares).
[1] Joint first author.
[2] Joint senior author.

viewpoint) but there is increasing consensus that the mechanisms underlying their strong outcomes on lab-based face perception measures might not be typical. Research to-date in this area is quite scarce, yet behavioural findings have led to the claim that face processing in WS is accomplished via an atypical route which draws on a featural processing style (e.g., Karmiloff-Smith et al., 2004). In the current series of studies, we seek to further characterise face processing abilities behaviourally (study 1) and to probe the neural profile of activation with respect to face-selective processing in WS (study 2).

Early studies of face processing in WS employed standardised face processing tasks that lacked sufficient sensitivity to detect subtle atypicalities present in this group. For example, in the Benton Facial Perception Test (BFPT: Benton et al., 1994) the participant is asked to match faces across different viewpoints and lighting conditions. Although performance is often in the "normal" range (i.e. above the clinical cut-off score, based on a normative sample) for people with WS (e.g., Bellugi et al., 1988; Bellugi et al., 1999; Tager-Flusberg et al., 2003), it has been shown that a typical score can be achieved on the BFPT by using non-face processing strategies (see Duchaine and Weidenfeld, 2003), thus a 'typical' profile observed on this task may reflect a lack of measurement sensitivity rather than age-appropriate skills. Karmiloff-Smith et al. (2004) also identified that despite participants with WS performing on the normal range on this task, their pattern of performance was atypical, which further confirms that a typical level of performance on this task is not necessarily indicative of typical abilities. The Rivermead Face Memory Test (Wilson et al., 1989) has also been employed with individuals with WS (Udwin and Yule, 1991). This task presents faces for whom external features such as hair are not excluded. Arguably, these features can be used in place of internal facial features to recognise the previously presented faces, i.e., again, the task can be completed by using non-face processing strategies (Duchaine and Weidenfeld, 2003).

Beyond using standardised measures to estimate levels of processing ability in WS, studies have also attempted to characterise the mechanisms used to support face processing in this group. In the typical population, holistic processing is critical for face identity recognition, i.e., the automatic integration of facial information into a whole representation (Young et al., 1987). Yet as noted above, it has been proposed that individuals with WS, in contrast, rely strongly on featural cues, i.e., the individual components of a face (eyes, nose, mouth), to recognise faces (Deruelle et al., 1999). Several studies have used the face inversion effect (FIE: Yin, 1969) to draw inferences regarding face processing mechanisms. A FIE is demonstrated by a disproportionate increase (compared with other object categories) in the difficulty of identifying an inverted face compared to an upright face. This selective drop in performance may reflect the disruption of holistic processing - which is relatively more enhanced for upright than inverted faces – and subsequently increased reliance on featural information (Leder and Bruce, 2000). If true, then it would follow that the FIE should be weaker or even absent in individuals with WS on account of their purported featural processing style. Deruelle et al. (1999) supported this hypothesis, demonstrating that individuals with WS (N = 12) were less impacted by inversion than TD mental age (MA) and chronological age (CA) matched controls (see also Karmiloff-Smith et al., 1997). Similarly, Annaz et al. (2009) report no face inversion effect in WS. However, this pattern is not consistently replicated. Both Rose et al. (2006) and Riby et al. (2009) report the presence of significant face inversion effects in WS: in face recognition and in the Thatcher effect/illusion, wherein local feature changes (e.g., selectively mis-oriented eyes) are much more obvious to participants when faces are presented upright than inverted. Whilst the findings of Rose et al. (2006) are difficult to interpret because the typically developing comparison children were not matched to the WS group (N = 19), data from Riby et al. (2009) suggest that individuals with WS *are* drawing upon some holistic information when processing face identity, albeit not at the level expected from their chronological age.

Karmiloff-Smith et al. (2004) conducted a fine-grained series of studies to closely investigate how individuals with WS process configural information in faces (i.e., the spatial distances amongst internal features: second-order relational information) from a developmental perspective. In Experiment 1, the authors compared sensitivity to featural vs configural changes in faces by replacing the original features vs changing the spacing between features, respectively. In Experiment 2, they analysed upright and inverted face recognition in order to estimate configural processing. Finally, in Experiment 3, they analysed sensitivity to configural or featural changes as in Experiment 1 but this time with schematic faces and geometrical patterns. The authors determined that, whilst some configural processing is observed in individuals with WS (Experiment 1 and 3 with Ns = 12, and Experiment 2 with an N = 14), they show reduced sensitivity to configural information in upright faces (Experiment 1) and demonstrate an atypical developmental trajectory for detecting configural transformations in upright schematic faces (Experiment 3), relative to TD controls (aged 5–53 years, to encompass both mental and chronological age ranges of the participants with WS). Furthermore, they point out that across the literature, deficits in WS face processing skills are typically observed when configural processing is a crucial element of the task. They are less consistently reported in tasks for which alternative strategies (e.g., featural processing) are possible, which might mask broader behavioural deficits relative to typical comparison groups.

Whilst detailed consideration of face *expression* recognition is beyond the remit of this paper, it is also noteworthy that a similar atypical processing route has been reported with respect to face expression processing in WS. Ewing et al. (2017b) recently demonstrated that relative to TD adults and TD children (matched for overall face expression recognition ability), individuals with WS (N = 18) draw upon a wider and less integrated set of visual information when categorising faces by their emotional expression. For example, as observed using the Bubbles reverse correlation paradigm,[3] there was relatively greater use of creases in the forehead and around the chin for fear detection in WS compared to the typically developing groups, who relied more on the eye and mouth region. This ability to achieve a similar level of performance to controls, via an alternative processing route is consistent with the results mentioned above in the context of face identity processing. Taken together, this evidence strongly suggests that atypical processing mechanisms may underpin face expertise in individuals with WS.

WS provides an important demonstration that seemingly typical behavioural outcomes can be driven by different underlying mechanisms. The neuroconstructivist perspective posits that the operation of such compensatory processes - used in order to overcome limitations to typical mechanisms - reflects a developmental cascade of impairment which originates in the infant start-state. That is, over developmental time, an alternative processing route develops, which draws on available processing resources and mechanisms to circumvent the system's constraints (Farran and Karmiloff-Smith, 2012).

Not surprisingly, evidence of atypical face processing mechanisms in WS is also observed at the neural level. Using fMRI, Mobbs et al. (2004) report that right occipital regions, typically associated with

---

[3] The Bubbles paradigm seeks to establish the visual information that is used by participants to complete a particular categorization task, in this case a face recognition task. The paradigm works by presenting participants with a sub-set of the information present in a stimulus image. Information is revealed by a small number of randomly located circular Gaussian apertures (the bubbles), the rest of the stimulus is hidden from view. On each trial the information presented is different (e.g. on one trial the participant may see a part of the eyes, the mouth, and the forehead and on another a part of the mouth alongside the nose and cheek). Reverse correlation approaches can then be used to tie the behavioural response on each trial (correct/incorrect) to the information that was presented and thus make inferences about the critical information necessary to correctly perform a particular categorization task.

global/holistic processing, were activated in their TD control group, but not in their WS group (N = 11), during a face processing task. In contrast, in the WS group, face processing was associated with activation of the right prefrontal cortex, which the authors interpret as a reflection of a featural processing style. Furthermore, a key neural region critically implicated in face processing: the Fusiform Face Area (FFA) is enlarged in WS relative to the typical population (Golarai et al., 2010). Golarai et al. (2010) report that the FFA area remained face-selective in their WS group (N = 13), and among other explanations (e.g. genetic), the authors suggest that the observed enlargement might be an environmental consequence of the increased interest in faces in WS, leading to cortical expansion in the areas that represent face processing. They did not, however, observe a direct functional consequence of this enlargement on their chosen behavioural measure of face processing ability: both groups had comparable performances in the BFPT. In line with this, O'Hearn et al. (2011) report neurotypically-similar levels of ventral stream activity for faces in WS, alongside selectively reduced activity for other objects (O'Hearn et al., 2011).

A fruitful way to gain further insight into the mechanistic underpinning of face processing in WS is to investigate the electroencephalogram (EEG) signal associated with these stimuli. An extensive TD research literature has established key (face) event-related potential (ERP) components (measured via EEG) that have a distinct time course and scalp distribution. These components index specialised neural mechanisms tuned to the detection and processing of faces. Importantly, the face-selective negativity which peaks 170 ms after stimulus onset (with the whole component occurring over a 130–200 ms window), known as the N170 component, is a hallmark of typical face processing (Bentin et al., 1996). It has been shown reliably and consistently to be face selective (larger and earlier N170 response to faces than other object categories e.g. Rossion and Jacques, 2008), as well as orientation selective (larger and delayed N170 response to inverted compared to upright faces, e.g. Itier and Taylor, 2004).

There have been five studies to-date that have used EEG to investigate face identity processing in WS, two of which directly investigated the N170. Four of these studies focused on the mechanisms used to support face processing by comparing neural activity associated with viewing upright vs. inverted faces. They did not investigate the more basic question of whether there are selective responses to faces *as a category* compared with other objects. One study while focusing on face memory did include houses as a control condition but did not directly analyse the N170 component (Key and Dykens, 2016). Grice et al. (2001) presented upright and inverted faces to eight individuals with WS. They report the classic N170 profile for upright compared to inverted faces in relation to latency (but not amplitude) in WS. In contrast, Nakamura et al. (2013) reported conflicting neural profiles in two participants with WS. One participant demonstrated a typical N170 profile for inversion and the other did not – though clearly with such a small sample it is difficult to draw conclusions from this study. Mills et al. (2000) report an unusual pattern of neural activation in individuals with WS (N = 18) that showed little differentiation between upright and inverted faces (although the standard components, e.g. the N170, were not directly investigated), despite a typical behavioural FIE. Using the same task with a much larger group of individuals with WS (N = 74), Mills et al. (2013) observed similarly atypical face-related neural processing, although a direct comparison of neural activation to upright and inverted faces is not reported and again the standard components were not directly investigated.

Another early neural component that has been linked to face processing in adults is the P100 component (or visual P1). This positive component peaking approximately 100 ms after stimulus onset (running from 80 to 100 ms) appears sensitive to many low-level properties of stimuli and originates in extrastriate visual areas (Di Russo et al., 2002). In contrast to the N170, face selective effects on the P100 where a larger response is observed to faces compared to other object categories are generally less consistent and appear to be driven primarily by low level

visual features of the stimuli (Rossion and Caharel, 2011). However, given the frequency with which face selective effects are reported on the P100, we include it here for completeness and note that to-date no study has considered face selective effects on the P100 component in WS.

There is clearly scope to increase our understanding of face processing in WS, both at a behavioural and neural level. Researchers have often thus far relied on behavioural measures of face perception that can be accomplished using extraneous features to the face, wherein scores in the typical range need not reflect typical face-specific processing mechanisms. In the current study, we employed a measure that is emerging as a 'gold standard' tool in developmental face perception research: the Cambridge Face Memory Test for Children (CFMT-C; Croydon et al., 2014). The CFMT-C was adapted from its well-known adult version (CFMT, Duchaine and Nakayama, 2006), that avoids any influence of extraneous features by closely cropping the faces and excluding the hair, ears and neck, to ensure a sensitive measure of face expertise for developmental researchers. Furthermore, whilst suitable for children, the CFMT-C uses adult faces, thus avoiding any confounds relating to own-age bias (Anastasi and Rhodes, 2005), which might have disadvantaged our (adult) WS group. We have chosen to use a face processing task that is designed for children because we predicted that it would be more likely to be sensitive to the range of scores of the WS group than a task designed for adults. This decision is based on the studies reviewed above which suggest that, whilst face processing is a relative strength in WS, it is unlikely to be age appropriate. The same is true for other ostensible strengths within the WS profile, such as verbal ability, which is typically at the level of an 8- to 10-year-old (e.g. Farran et al., 2016). A measure designed for children is also appropriate with respect to the attentional and procedural task demands. We recognise that our use of this measure risks ceiling effects in some of our participants, but we considered this risk to be necessary, in comparison to the risk of floor effects with a task that is designed for adults. The aim of Experiment 1 is to provide the most accurate measure of face processing abilities in WS to-date. We will use regression functions from TD normative data (age range: 5–12 years) published in Croydon et al. (2014) large-scale validation study, coupled with TD adults data collected for the purpose of the current study, to establish the face processing mental age of the WS group. We will compare face processing mental age to verbal and non-verbal mental age in our WS group. Based on the discussion above, we do not predict face processing performance to be in the typical range for our WS group. We do, however, predict that face processing ability will represent a relative strength within the WS cognitive profile. The data from Experiment 1 also provide vital background information regarding the functional consequences of any atypicalities identified in Experiment 2. Experiment 2 provides a detailed characterisation of the neural profile associated with face processing in WS. That is, we explore the neural responses associated with passive viewing of upright and inverted faces and houses, focusing on early neural responses to faces via standard Event Related Potentials (ERPs; P1 and N170) and employing state-of-the-art multivariate pattern analysis (MVPA, Smith and Smith, 2019 & Mares et al., 2020).

2. Experiment 1

Face identity recognition ability was characterised in adults with WS using the CMFT-C (Croydon et al., 2014). By testing adults with WS we ensured that all participants were at the adult end-state of development and thus any individual differences observed should not be confounded by developmental factors.

2.1. Method

2.1.1. Participants

Nineteen adults with Williams syndrome were recruited from the records of the Williams Syndrome Foundation, UK. All had a phenotypic diagnosis of WS from a clinician as well as genetic diagnosis based on a

"fluorescent in situ hybridization" (FISH) test for the deleted Elastin gene (see Lenhoff et al., 1997). One participant did not complete the CMFT-C and so was excluded from analyses. The final sample of participants with WS (N = 18) ranged in age between 18 and 51 years (M = 29.43 ± 9.43 yrs; 13 females; see Table 1 for participant information).

WS data was compared to three groups. To determine face processing mental age, we used the linear function between CFMT-C performance and Chronological age from 282 typically developing (TD) 5-to 12-year-olds (published in Croydon et al., 2014) as well as data from 30 TD adults who completed the CFMT-C in a quiet room at the University (see Table 1 for participant information). To determine the extent to which the cognitive profile of WS performance was typical or atypical, we compared it to the cognitive profile of 28 TD 4- to 11-year-olds (Mean [SD]: 8.38 [1.65] years) who completed the CFMT-C as well as a measure of verbal and non-verbal mental age.

All participants were provided with information about the experiment and verbally assented to take part. Written informed consent was provided by the parents/guardians of individuals with WS. Participants were given a small gift voucher to thank them for their participation. Both studies reported herein were approved by the ethical committee of the Department of Psychological Sciences, Birkbeck College, University of London.

2.1.2. Procedure

The CFMT-C is a measure developed to assess face memory ability in children and closely mirrors the CFMT used in typical adult populations (see Croydon et al., 2014 for detailed information regarding the methods). This brief measure (which takes approximately 15 min) has three distinct stages in which participants are trained and tested in remembering 5 upright facial identities: Caucasian men posing with neutral facial expressions shown from three different viewpoints. In Stage 1 participants learn to identify the five faces and are then tested for their recognition of those identities in a two-alternative forced choice paradigm in which the test phase target images are always identical to those presented during the study phase (15 trials). In Stage 2, participants must recognise the learned identities when they are presented from novel viewpoints and under varied lighting conditions (25 trials). Finally, in Stage 3, participants must recognise novel images of the 5 identities with a pre-specified level of Gaussian visual noise (20 trials).

In addition to the CFMT-C, the WS participants and TD children also completed the British Picture Vocabulary Scale (BPVS-III; Dunn et al., 2009) as a measure of receptive vocabulary and the Ravens Colour Progressive Matrices (RCPM, Raven, 2008) as a measure of nonverbal reasoning. These were used to determine verbal and non-verbal Mental Age (MA) in our sample respectively. In the BPVS, the experimenter asks the participant to match a spoken word with the most representative picture (from four possible line-drawings). In the RCPM participants are presented with a visual stimulus and asked to select a picture that completes the pattern (from six possible options).

2.1.3. Results

There were two aims to this experiment. First to determine the absolute level of face processing abilities in WS, relative to their Chronological Age (CA), and second to determine where face processing abilities sit within the broader cognitive profile of this group. To meet our aims, raw scores of WS performance on each of the measures, CFMT-C, RCPM (non-verbal ability) and BPVS (verbal ability) were converted to age equivalence scores. For the RCPM and BPVS raw scores, this process was completed by using the norms data in the respective manuals. Unfortunately, the RCPM manual does not provide age equivalence data below a raw score of 14. Thus, for scores below a score of 14 (N = 4 participants), participants were credited with the lowest mental age provided by the manual, 4 years. This is a conservative estimate because it has the effect of masking a lower mental age value and reducing the range of mental age for this task. However, it was preferable to extrapolating beyond the normed age range. For the CFMT-C we used the linear function published in Croydon et al. (2014) and data from our TD adult sample to determine face processing mental age. The linear function is based on data from 282 TD children aged from 5 to 12 years. Five participants with WS scored a CFMT-C percentage accuracy of over 85.25%, which would give them a Mental Age beyond the age range of the Croydon et al. (2014) dataset. The TD adult group had a mean CFMT-C score of 93.39 (95% Confidence Interval: 90.34–96.44). Three of the participants with WS had CFMT-C scores which fell within the 95% Confidence Interval of the TD adult mean and so were credited with a Mental Age of 18 years. For the remaining two WS participants, we chose to give these participants a conservative estimate of 12 years (the maximum age of the Croydon et al., (2014) dataset), rather than extrapolating beyond the dataset (in practice, their extrapolated scores would have both been 12.51 years). We chose to provide conservative estimates for both RCPM and CFMT-C scores, as this actually risks bringing the group means closer together, an effect which is in the opposite direction to the hypothesised group difference for these two measures. Thus, we can be confident that any observed differences between these two scores in our WS sample are real.

ANOVA of the profile of Mental Age (MA) variables (face processing MA, verbal MA, non-verbal MA) risks being confounded by differences in the published normed data for each task. That is, the normative samples for each task were *different*; when the tasks are then used for the *same* individuals this does not necessarily mean that even a TD sample would accrue the same mental ages across these tasks. In order to account for this we calculated face processing MA, verbal MA, non-verbal MA for 28 typically developing 4- to 11-year-old children who had completed all of these tasks (raw scores and MA for this TD comparison group and the WS group are shown in Table 1). We then created z-scores of the WS group Mental Ages based on the mean and standard deviation of the TD Mental Ages for each task. This has the effect of standardising any differences in the normed data that was used to derive MA, such that any differences remaining constitute true differences in Mental Age. A one factor ANOVA of the profile of Mental Age variables (face processing MA, verbal MA, non-verbal MA) using z-score as a dependent variable demonstrated a main effect of task, $F(2, 34) = 27.301$, $p < 0.001$, $\eta p2 = 0.62$. Sidak post-hoc comparisons revealed that this was due to weaker MA on the non-verbal (RCPM) than the verbal (BPVS, $p < 0.001$) and face-processing (CFMT-C, $p < 0.001$) tasks, and no difference between verbal and face-processing MA ($p = 0.052$). Thus, as a group, in WS face processing MA is significantly stronger than non-verbal MA and statistically on a par with verbal MA (although note that verbal MA was

Table 1
Participant information and Mental Age profiles for WS and TD child groups.

	CA (years)	Verbal raw score[1]	Non-verbal raw score[2]	Face[3] % accurate	Verbal MA[1]	Non-verbal MA[2]	Face MA[3]
	Mean (S.D.)						
WS (N = 18)	29.43 (9.43)	126.16 (20.06)	18.47 (6.47)	79.98 (10.20)	9.83 (2.26)	5.76 (2.00)	10.36 (4.23)
TD children (N = 28)	8.33 (1.65)	116.07 (20.47)	28.82 (5.50)	81.49 (12.38)	8.52 (1.94)	8.96 (2.15)	11.42 (4.91)
TD adults (N = 30)	31.35 (10.69)	NA	NA	93.39 (8.16)	NA	NA	NA

CA – Chronological Age. 1 - British Picture Vocabulary Scale; 2 - Ravens Coloured Progressive Matrices; 3 – Cambridge Face Memory Test for Children.
Note that BPVS scores and BPVS mental ages are based on N = 27 for the TD group.

marginally stronger than face-processing MA).

Next, we determined whether face processing was a relative strength in WS at the individual level. This was carried out at the descriptive level only by categorising participants as either those for whom face processing MA was stronger than non-verbal MA, and those for whom it was equal to non-verbal MA (there were no participants who presented with weaker face processing MA than non-verbal MA). These two measures were chosen because, as discussed in the introduction, they have been proposed as a relative strength and a relative weakness in the WS cognitive profile respectively thus representing the full range of the WS cognitive profile. If the discrepancy between face processing and non-verbal abilities is borne out at the individual level it highlights the relative strength of this ability over and above individual differences in our sample. Fig. 1 shows the cognitive profile of the WS group at the individual level. In order to determine what constituted a reliable difference, we calculated difference scores by subtracting non-verbal MA from face processing MA for the 28 typically developing 4- to 11-year-old children who had completed these tasks (Table 1). The TD children had a mean difference score of 2.46 years (95% Confidence Interval: 0.65–4.26); note again that this difference is likely an artefact of their mental ages being derived from different normative samples. We categorised any difference scores of the individuals with WS which fell beyond the 95% confidence interval of the TD children as having reliably stronger face processing ability than non-verbal ability. This was the case for eight of the WS sample. (44% of the sample; i.e., their difference scores were beyond that of the estimated population mean of the TD children). For the remaining 10 participants the difference fell within (N = 8) or below (N = 2; difference scores of 0.51 and 0.22) the 95% confidence interval of the TD difference scores. Thus, whilst as a group, our data support the proposal that face processing is a relative strength in WS (when compared to non-verbal ability), at the individual level this was observed for eight participants only.

2.1.4. Discussion

Methodological issues have undermined many previous investigations of face processing ability in individuals with WS. Thus, extant claims that skills are in the typical range must be taken cautiously. Indeed, our data broadly refute this claim; using the CFMT-C, we have demonstrated that face processing expertise in adults with WS, as a group, is around the level of typically developing 10-year-olds. Importantly, however, we have identified large individual differences. For example, three of our participants with WS were performing at an adult level (17% of our sample), whilst others were performing at the

level of a TD five-year-old. This variability demonstrates that whilst it is possible for face processing performance, as measured with the CFMT-C, to be in the typical range in WS (17% of our sample), it is not universal for this group. It is worth pointing out, however, that there is also a wide range of face-processing ability observed within the typical population: the notion that all adults are similarly 'face experts' has been robustly challenged in recent years (see Young and Burton, 2018; Wilmer, 2017). The large variability in face processing ability observed in individuals with WS, however, exceeds that observed in our TD adult group, and therefore is unlikely simply a reflection of such neurotypical difference. It is possible that these individual differences also relate to the memory and attention demands of the task. For behavioural tasks, sequential presentation is vital because simultaneous presentation makes the test susceptible to the use of a feature-matching strategy (see Duchaine and Nakayama, 2004). This memory demand could be considered a confound given reports of poor visuo-spatial working memory in WS (e.g. Jarrold et al., 1999). Yet, in-depth investigation suggests that visual short term memory for objects may be independent from (impaired) spatial short term memory in WS (Vicari et al., 2006). Nevertheless, the contribution of memory demands cannot be fully dismissed. Experiment 2 addresses this issue by measuring neural activity rather than behavioural performance in response to the presentation of faces while participants engage in a simple unrelated task (spotting occasional butterflies), which minimises the impact of behavioural task demands upon outcomes observed across groups.

Our second aim was to assess the claim that face processing in WS represents a relative strength within their cognitive profile. Here, we observed that face processing was at the level of verbal ability (although note that this was a marginal p-value), but above the level of non-verbal ability. This supports the notion that face processing is a relative strength in WS when compared to other non-verbal abilities. However, again, observation of Fig. 1 demonstrates that this is not a universal finding across our participant group. It is certainly notable that all eighteen of our participants had positive difference scores (face processing mental age was always higher than non-verbal mental age), yet for 56% of our sample, the difference between face processing MA and non-verbal MA was within (or below) the range of difference scores observed in the typical population.

In conclusion, the results of Experiment 1 support the claim that recognition memory for faces is a strength relative to other non-verbal abilities in WS, which can sometimes be within the typical range. However, this result must be considered within the context of broad heterogeneity in both the level of face processing ability in WS and the
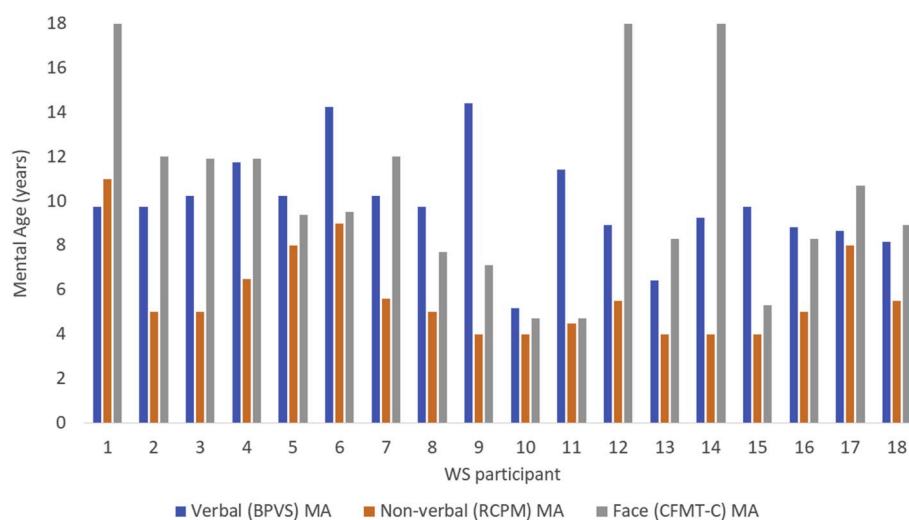

Fig. 1. Individual mental age profiles of Williams syndrome participants regarding verbal ability, non-verbal ability, and face processing ability. Note that participants 3, 4, 5, 6, 8, 9, 10, 14, 15, 16 and 18 also participated in Experiment 2.

cognitive profile of the group. This level of individual differences is not unusual in neurodevelopmental disorder groups (Farran, 2020) and can result from variation in input from multiple levels of description, be they genetic, neural, environmental or behavioural factors. These individual differences can be key to determining how the phenotypic outcome of WS in any one individual is expressed (see Karmiloff-Smith et al., 2016).

Experiment 2 extends these important behavioural findings by probing face processing at the neural level of description. We investigate the mechanisms used to support face processing and the neural (EEG) signals that are associated with face processing in WS at both the individual and group level.

3. Experiment 2

Here we characterise the neural response of individuals with WS and TD age matched controls when viewing upright and inverted faces and houses: a non-face comparison category of complex, perceptually homogeneous objects. A novel aspect of our study is the introduction of multivariate pattern analysis (MVPA). This state-of-the-art technique complements traditional univariate ERP analysis of EEG signal and is arguably more sensitive than these traditional approaches. MVPA has hitherto not been employed to investigate neural activity in WS. Experiment 2, therefore, constitutes the most comprehensive investigation of the neural signature of face processing in WS conducted to date.

An additional advantage of MVPA as a complementary analysis to the typical univariate ERP component analysis is that, in contrast to the latter, it is atheoretical with respect to time windows and the location of neural sources (see Thomas et al., 2009 for discussion of the advantages of atheoretical approaches with neurodevelopmental disorders). For groups in which the observed behaviour might be the result of an alternative developmental pathway and thus supported by neural mechanisms that are distinct from those underpinning behavioural outcomes in the typical population, an atheoretical approach can provide a richer understanding of this atypicality. Thus, whilst our analysis of hallmark ERP components is theory-driven (i.e. *Do individuals demonstrate typical N170 profiles of neural activity?*), the application of MVPA allows us also to identify distinctive patterns between categories at the individual level using a broader selection of the recorded data (e. g., *Do individuals display face-selective profiles of neural activity?*). This approach is ideal for use with individuals with WS, whose profiles of neural activation are difficult to clearly anticipate. If all analyses were determined by a theory-driven account based on typical face processing time course and location information, then we risk failing to detect abilities that are present in this atypically developing group.

Adults with WS and chronological age matched TD controls were presented with upright and inverted faces and houses. As with Experiment 1, using a sample of adults with WS limited any confounds brought about by developmental factors. Given the well-known developmental changes in the ERP waveforms and topographies (Kuefner et al., 2010; Taylor et al., 2004) that can result from multiple non-cognitive factors in children (e.g., overall head size or skull thickness), mental age matched controls could not be included. Furthermore, mental age matched controls would have demonstrated large individual differences on account of the age-related changes in the extent to which ERP components are face sensitive. Matching by chronological age avoided these difficulties, and enabled us to equate, as far as one can, for experience with faces between the groups.

In an effort to limit any confounds related to differences in cognitive ability between the groups, there was no explicit behavioural task related to the stimuli of interest. Both MVPA and standard ERP analyses were used to explore the selectivity of neural responses to 1) faces as an overall category (by examining neural activity associated with upright faces and houses), and to 2) the canonical upright configuration of a face, critically associated with expertise in typical development and often linked with holistic processing (upright and inverted faces). For

this latter question, to infer that any differences found are face-sensitive and not related to processing orientation per-se, we included a control non-face category for which no neural differences were predicted when viewed at different orientations (upright and inverted houses). Based on the studies reviewed earlier, there is a growing consensus that the relatively strong face processing in WS might be achieved through atypical mechanisms (e.g., possibly more reliance on featural processing). If the WS group show selective differences in neural activation to faces as a category (comparing faces to houses), or to face orientation (comparing upright to inverted faces) relative to a TD control group, then this would provide important evidence that face processing is indeed atypical in this group and help us to understand what is driving these differences. We predict that the WS group will show strong face-selectivity (i.e., clear differences in the neural response to faces and objects). Furthermore, we predict weaker or absent neural differentiation with reference to face orientation, relative to the TD control group, reflecting poorer differentiation of face inversion relative to upright faces in this group. Because the face inversion effect is a marker associated with face expertise in TD populations (Rezlescu et al., 2017), this group difference would strongly signal that face processing expertise in WS is supported by atypical mechanisms.

3.1. Method

3.1.1. Participants

A subset of the adults with WS reported in Experiment 1, who were willing to take part in an EEG testing session also took part in Experiment 2. The final WS group (N = 11) ranged in age between 23 and 50 years (M = 29.8 ± 9.10 yrs; 7 females; BPVS mental age: 10.45 ± 2.32 yrs; RCPM mental age: 5.74 ± 1.86 years). These individuals were individually matched for gender and chronological age (CA) with typically developing participants (7 females M = 30.27 ± 8.78 yrs; t(20) = -0.12 p = 0.91; Cohen's d = -0.05) to account for the well-known age related changes in the electrophysiological signal with healthy ageing (e.g. Taylor et al., 2004).

3.1.2. Stimuli

Greyscale photographs of six unique face identities (all male) were presented (see Schyns and Oliva, 1999). Faces were presented frontally with neutral expressions. Images were digitally manipulated to control for hairstyle and outline, using the same template for every face. Greyscale photographs of six unique houses (see Eimer, 2000) were used as a control and similarly edited to have the same outline as the face stimuli. We controlled for luminance and contrast using the Shine toolbox (Willenbockel et al., 2010). Inverted versions of the upright images were created for all stimuli. Participants were seated at a distance of approximately 70 cm from the screen in a comfortable chair in an electrically shielded and sound-proofed room such that visual stimuli subtended a visual angle of approximately 4.09° width by 6.13° height (176 × 256 pixels).

3.1.3. Procedure

Participants completed this experiment as part of a larger testing battery. Here, participants were asked to attend to a series of photographs of upright and inverted faces and houses. Stimuli were presented centrally using Eprime software, version 2.0 (Psychology Software Tools Inc.; www.pst-net.com/eprime) on a grey background (750 ms). This was followed by a black fixation cross (between 1650 and 1850 ms in steps of 25 ms). Participants completed 60 trials of each condition (for a total of 240 trials), alongside 60 catch trials in which a colourful butterfly was presented to the left or right of fixation. This (otherwise irrelevant) sub-task helped to keep participants attending to what was shown on the screen. Participants were asked to press a labelled keyboard key indicating the side of the butterflies' appearance (left or right). Specific instructions were as follows: "During this game, you will see lots of things presented on the screen, but you need to look out for

butterflies. If you see a butterfly on this side of the screen press this key, if you see one on this side of the screen press this key. Remember, you need to watch all of the pictures closely, but only press the keys when you see a butterfly". In the butterfly trials participants had an additional 1750 ms to allow for a response. The trial order was randomized per participant, and then divided into 10 blocks (30 trials each) to allow for breaks. Participants were closely monitored by the experimenter and the task was discontinued where there were concerns about lack of task engagement or fatigue in all participants irrespective of group.[4] Performance on the butterfly catch trials was close to ceiling in both participant groups. Still, the TD group (TD, M = 98.48 ± 2.17%) were numerically more accurate than those with WS (WS, M = 89.31 ± 14.79%) and this group difference was marginally significant (t(10.43) = -2.04, p = 0.07, d = −0.87). These trials were later removed from all ERP analysis.

Of note, prior to participation in this experiment, as part of the larger battery of tests, all participants also completed a short task in which they were tasked with learning three of the six identities that were used in the current experiment (see Ewing et al., 2017a for further details on this task). This small task used the Bubbles reverse correlation paradigm, as described above, and asked participants to learn by name three novel identities, and then recognise them later when obscured by variable amounts of visual noise (maximum 216 trials – approximately 10 min). In the current EEG experiment participants were only asked to pay attention to the stimuli on the screen while searching for the appearance of a butterfly. No information was given to participants regarding potential face familiarity and this point was not emphasised to participants. Although the potential effect of familiarity is certainly an interesting question in its own right we did not set out to explore it in this study and due to insufficient statistical power do not analyse this categorization via ERPs or MVPA. We note that reliable effects of face familiarity typically only begin 250 ms following stimulus onset (e.g. N250R in repetition priming of the same identity, Schweinberger et al., 2002; N250 contrasting famous and unfamiliar faces, Gosling and Eimer, 2011; Andrews et al., 2017) and are relatively small in comparison to the earlier N170 neural response to faces and face inversion (see Ramon and Gobbini, 2018 for a recent review). Further, we note that early effects of familiarity are usually tied to familiar faces whom participants know in advance of the study (not newly learned single images). Given the relatively impoverished level of true person familiarisation likely to be achieved in the current project (single images of previously unknown individuals, highly standardised exemplars), the use of an orthogonal non-face task and our primary interest in early visual components and decoding we are confident that the reported results are not driven by this factor.

3.1.4. EEG recording and pre-processing

EEG was continuously recorded using a fitted cap (EASYCAP) with 32 Ag–AgCl electrodes placed according to the international 10/10 system. Electrode impedance was lowered below 20 kΩ and an added electrode was placed below one of the eyes to monitor vertical eye movements and blinks. EEG was acquired at a sampling rate of 500 Hz and referenced online to FCz (AFz acted as ground). Data was analysed using the Matlab toolbox EEGLAB (Version 14.1.1, Delorme and Makeig, 2004). Continuous data was band pass filtered between 0.1 and 40 Hz, and epoched around stimulus onset from −200 ms to 500 ms. An average 2.27 ± 1.10 channels were interpolated per participant (maximum 4) after EEG-channel rejection using EEGLAB automated criteria and visual inspection. Epochs were baseline corrected using the

200 ms preceding stimulus onset and then visually inspected to remove artefacts such as eye blinks (large deflections observed across all electrodes), large eye movements (as observed in the two horizontal EOG electrodes), muscle/movement artefacts (observed as high-frequency activity) and electrode noise. After initial artefact rejection (M = 17 ± 2.77%), the final number of trials was equalized between groups. Given that overall the TD participants completed more trials than those with WS (M = 197 ± 15.09), we adjusted the number of trials included from the former group to equate the means of the two groups (TD M = 197 ± 10.19).[5] Catch trials were removed from all further analyses.

3.1.5. ERP analysis

Channels for ERP analysis (O1/2 and P7/8) were selected based on the maximum peak difference between the two key early visual components: the P100 and the N170 from the grand average waveform across both groups over parieto-occipital channels. The mean amplitude was calculated for the P100 in a 20 ms window centred around the peak latency (102 ms). A similar approach was conducted for the N170 component using a larger window of 40 ms, given this component's wider morphology (peak latency: 160 ms). P100 and N170 individual latencies were automatically extracted as the maximum peaks (positive and negative respectively) in a window between 70 ms and 150 ms after stimuli onset for P100 and between 150 ms and 240 ms for the N170 component.

A four-way mixed ANOVA with hemisphere (left, right), orientation (upright, inverted), category (face, house) and participant group (WS and TD) was performed to analyse the parameters of the ERP components. Effect sizes for paired t-tests were calculated using formula 3 by Dunlap et al. (1996). Means and standard errors of means are presented.

3.1.6. MVPA analysis

A multivariate pattern analysis was conducted to study possible differences in the neural activity associated with face processing between TD participants and participants with WS regarding a) face category sensitivity as analysed by the face vs house comparison, and b) expert face processing as indexed by the upright vs inverted faces comparison. To analyse the selectivity of possible effects of orientation to faces, a third (non-face) control comparison was also run, contrasting upright and inverted houses.

Linear SVM classifiers were trained for each of these planned comparisons per timesample (downsampled to 250 Hz) using single trials across posterior electrodes (O1, O2, P7, P8, P3, P4, Pz, TP9, TP10). This set of occipito-temporal electrodes should encompass the key visual neural activity relevant for processing of our face stimuli (Smith and Smith, 2019). Independent splits of the data, using 70% of the trials for training and 30% for testing (see Smith and Smith, 2019) which was repeated 20 times over independent splits of the data creating a twenty-fold cross-validation. Number of trials included per condition were equated to the category with less trials per comparison in each participant as not to bias the classifier. This was further repeated 100 times for robustness (Cauchoix et al., 2014). Accuracy was calculated by testing the trained classifier against the averaged EEG pattern of all trials from the test set of each condition, in order to increase signal to noise (Gallivan et al., 2013; Smith and Muckli, 2010; Smith and Smith, 2019).

The same procedure was repeated using permuted labels to produce an empirical measure of chance level (100 iterations). A classifier created with the actual labels was included in the chance level distribution as it is one of the possible outcomes (Pereira et al., 2009). Decoding accuracy was investigated at the group level via a paired, one-tailed t-test calculated at each timepoint (False detection rate, FDR, corrected) comparing the average decoding with the average chance

---

[4] Seven WS participants did not complete the whole 240 trials. These participants completed an average of 214.42 trials (std = 43.79; min = 150), which was sufficient for significant classification of our key comparisons (face categorization and orientation) for each single individual in the MVPA analyses. All TD participants completed the whole task.

[5] Trials were deleted from the TD participants in reverse chronological order to match the numbers completed by the WS participants. See Mares et al. (2020) for details.

level decoding. To limit the number of multiple comparisons, this analysis was only conducted for the time samples between 60 and 500 ms (111 comparisons).

We further sought to analyse the three key comparisons at the individual participant level, in order to allow for a direct statistical comparison of the profiles observed between groups (Mares et al., 2020). Individual-level decoding was similarly investigated by generating individual null distributions for each participant, using classifiers trained on permuted labels (total of 1000 iterations). Significant classification was confirmed for each sample, in each individual, when the accuracy obtained with true labels was greater than or equal to 95% of the null distribution (FDR corrected, Pereira et al., 2009; Smith and Muckli, 2010, Mares et al., 2020).

This procedure enabled us to create measures at the individual level, namely 1) *decoding onset* - first time-point where significant decoding was observed and exceeded baseline levels, 2) *sustainability of decoding* - percentage of significant decoding in the considered time-window (60–500 ms), 3) *peak decoding* - maximal peak classification accuracy in a given time window (100–300 ms) and 4) *peak decoding latency* – latency where the maximal peak classification accuracy was observed.

3.2. Results

3.2.1. ERP results

For the standard ERP analysis, we considered the amplitude and latency of two early visual components that are associated with processing faces: the P100 and the N170. ANOVAs were carried out on the amplitude and latency of these components, with within participant factors of hemisphere (left and right), stimulus category (faces vs. houses) and orientation (upright vs. inverted) and a between participant factor of participant group (WS, TD). Here, we focus solely on the contrasts of direct relevance i.e., those predicted a-priori from extant literature (a full description of the ERP results can be found in the Supplementary Materials). We therefore only report main effects of stimulus category (faces vs. houses), interactions of category with orientation (upright vs. inverted), and any significant interaction of these factors with participant group (TD vs. WS). See Fig. 2, top-panel, for the grand-average ERP plots per participant group, split by experimental stimulus category and cortical hemisphere. Fig. 2, lower-panels, depict violin plots illustrating the individual participant statistics for the critical components and experimental conditions (upright and inverted faces, faces and houses) with lines connecting participant's individual data points across conditions to visualise the consistency of any pattern of differences at the individual level.

We observed no relevant effects in either amplitude or latency of the P100 component (p > 0.05 for all). There was a main effect of stimulus category (faces vs. houses, F(1,20) = 20.40, p < 0.001, ηp2 = 0.51) observed in N170 amplitude, which was driven by a larger N170 component for faces (M = −5.01 ± 1.58 μV) compared to houses (M = −1.78 ± 1.68 μV). Stimulus category did not interact with group, F(1, 20) = 1.551, p = 0.227, ηp2 = 0.072, or with stimulus orientation, F < 1. There was, however, marginal evidence that the effect of stimulus category was in fact mediated by an interaction with stimulus orientation and participant group (F(1,20) = 3.08, p = 0.09, ηp2 = 0.13). Given the strong a-priori interest in the response profile of the N170 component to face-inversion we ran a follow up analysis in each participant group separately. As one would expect, there was an interaction between stimulus category and stimulus orientation for the TD group (TD: F(1, 10) = 6.066, p = 0.034, ηp2 = 0.378) which was not seen in the WS group (F < 1). This was driven by the typical face inversion effect in the TD group (F(1, 10) = 14.924, p = 0.003, ηp2 = 0.599) with larger amplitudes for inverted faces (M = −5.3 ± 2.38 μV) compared to upright faces (M = −7.871 ± 2.84 μV), but no house inversion effect, F < 1. In contrast, neural responses in the WS group were uniformly not impacted by inversion (F < 1). These results suggest that while the ability to categorize faces is similar between groups, sensitivity to face inversion

seems atypical in WS.

Regarding the N170 latency[6] we observed an interaction between stimulus category, stimulus orientation and group (F(1,19) = 4.47, p = 0.048, ηp2 = 0.19). This result was driven by an interaction between category and orientation in the WS group (F(1,10) = 12.02, p = 0.006, ηp2 = 0.546). A faster N170 component was observed in WS for upright faces (156.09 ± 1.66 ms) compared with inverted faces (168.18 ± 3.68 ms; F(1, 10) = 16.03, p = 0.003, ηp2 = 0.616), while no differences were observed between upright houses (163.27 ± 3.10 ms) compared to inverted houses (162.55 ± 2.86 ms; F < 1). This is evidence of a face selective inversion effect in this group, albeit atypical as revealed by the lack of an amplitude effect.

There was a faster N170 for upright faces (156.09 ± 1.66 ms) compared to upright houses (163.27 ± 3.10 ms; F(1, 10) = 7.819, p = 0.019, ηp2 = 0.439). Analysis of latency in the TD group showed no main effects or interactions (stimulus category: F < 1; stimulus category x orientation: F < 1; orientation: F < 1).

3.2.2. MVPA - stimulus category (upright faces vs upright houses)

Sustained significant face category decoding was observed earlier in TD adults than individuals with WS (from 136 ms for WS adults and from 104 ms for TD adults [after an earlier un-sustained decoding bump at 72 ms]). Both groups reached peak accuracy 156 ms after stimulus onset with the typical group reaching a higher peak decoding accuracy level of 83.88% compared to 77.61% in WS (see Fig. 3, top row for the time course of significant decoding, with significant time points highlighted). Furthermore, across the time-course, we note a larger percentage of significant category decoding for the TD group (88.29% compared to 73.87% for the WS group).

To formalise these differences, we proceeded to an individual participant level analysis. Taking such a fine-grained, more robust, approach to data analysis is unfortunately not yet standard in applications of MVPA or standard ERP analysis due to the technical challenges and intrinsic noise in the EEG data which requires sufficient signal collated over a number of trials. Indeed, when individual analysis is conducted, well-known phenomena observed at the group level are sometimes not reliably found at the individual level (Oruç et al., 2011). We observed significant decoding (p < 0.05, FDR corrected) of category (faces vs. houses) in all participants in both groups (see Supplementary Fig. 1 for all individual participant decoding time courses). Comparing the two groups, there were no significant differences in the onset of decoding (t(19) = -0.67, p = 0.51, Cohen's d = −0.29, TD: 143.27 ± 30.38 ms vs WS: 171.60 ± 28.97 ms)[7] nor in the sustainability of decoding (t(20) = 1.27, p = 0.22, Cohen's d = 0.54, TD: 61.75 ± 6.19% vs WS:49.71 ± 7.19%). We further investigated the peak of decoding and found no strong evidence of differences between TD and WS participants regarding either the latency of peak decoding (t(20) = 1.29, p = 0.21, Cohen's d = 0.55, TD: 174.55 ± 13.98 ms vs WS: 198.55 ± 12.33 ms) or the peak accuracy level (t(20) = 1.20, p = 0.25, d = 0.51, TD:94.13 ± 2.49 vs WS: 90.07 ± 2.31%). Fig. 3, lower panel, Category, presents these metrics as violin plots with each individual participant data point plotted, alongside the median and distribution. In summary, despite apparent group level differences in decoding profile, when investigated in more depth at the individual level there were no significant differences in face category decoding between the groups.

3.2.3. MVPA - face orientation

Sustained significant face orientation decoding was again observed earlier in TD adults than WS (from 132 ms for WS adults and from 116

---

[6] One participant from the typically developing adult group was excluded from the N170 latency analysis as peaks could not be established in all four experimental conditions.

[7] Note that one participant with WS did not meet the criterion for inclusion in the onset analysis and was removed from this analysis.
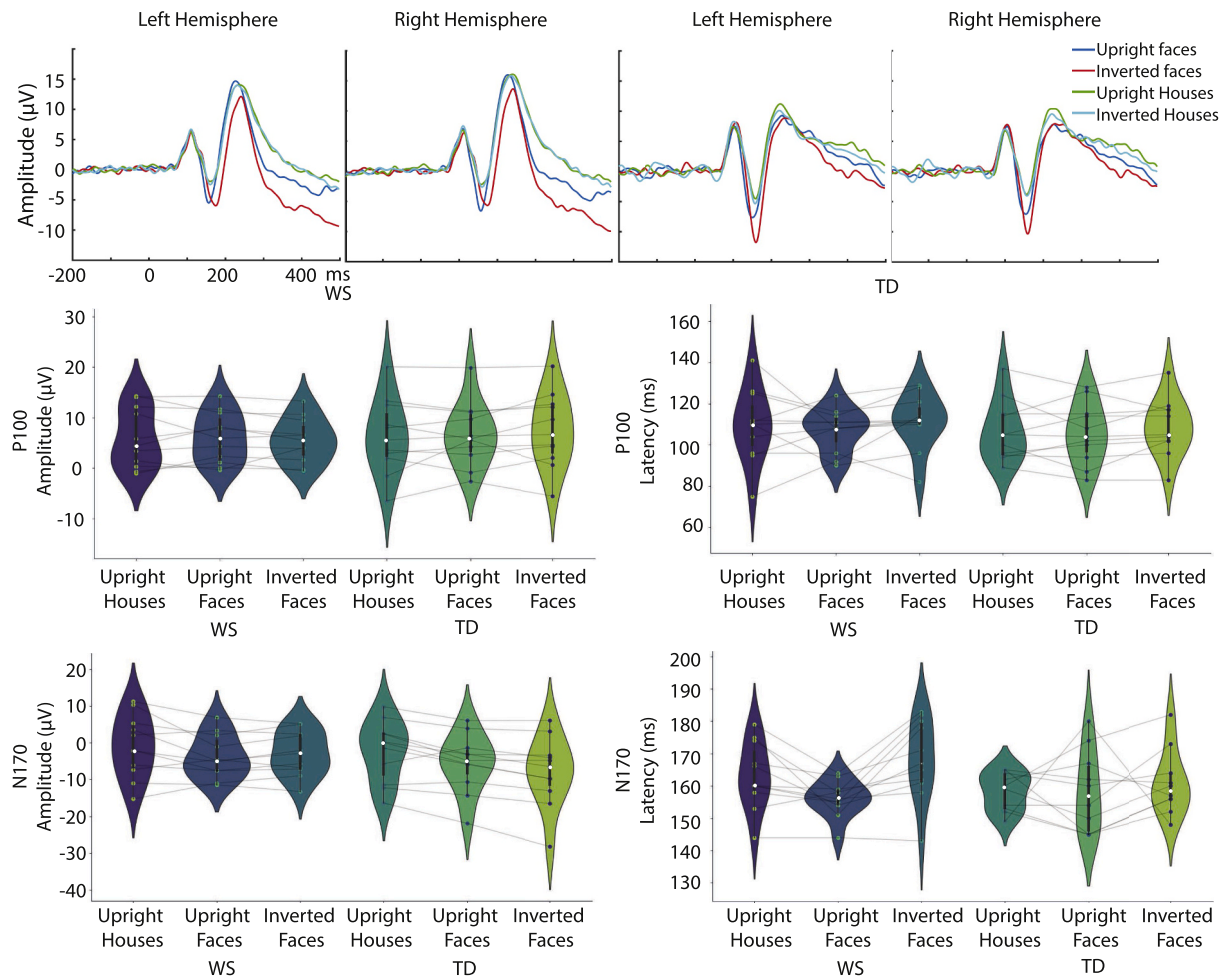
Fig. 2. ERP time course for selected electrodes per participant group (left and right hemisphere; top row). Violin plots are presented to depict the distribution underlying comparisons of the P100 (second row) and N170 (third row) amplitude and latency in the three critical categories (upright houses and upright and inverted faces. These plots highlight the spread of the kernel density estimation of the underlying data distribution (via their envelope), the median (white dots) and the individual data points. For completeness, lines link performance for the same individual in each category.

ms for TD adults [initial un-sustained decoding also noted at 80 ms]). Interestingly WS individuals displayed larger but later peak decoding accuracy compared to TD (88.9% at 204 ms vs 83.9% at 176 ms). Similarly, a larger percentage of significant orientation decoding across the time-course was observed in the WS group, 83.78%, compared with the TD group, 40.54%, see Fig. 3, second row for the time-course of decoding in each group. Note that this sensitivity for stimulus orientation appears to be face selective, with no significant decoding observed at the group level between upright and inverted house stimuli (p > 0.05, FDR corrected), see Fig. 3 third row.

As before, we further analysed sensitivity to face orientation at the individual level (given the lack of significant orientation classification for houses at the group level, a parallel analysis of the house data at the individual level was not warranted). Significant decoding of face orientation was again observed for all participants (see Supplementary Fig. 2 for all individual participant decoding time courses). Here, in support of the group level differences, a trend for a delayed onset of decoding in participants with WS was observed (WS: M = 157.45 ± 12.49 ms, TD: M = 129.20 ± 9.94 ms, t(19) = -1.77, p = 0.093, Cohen's d = 0.77),[8] alongside a significant delay in the latency of peak level decoding (WS: M = 214.18 ± 8.7 ms, TD: M = 179.63 ± 10.95 ms, t(20)

= -2.47, p = 0.023, d = −1.05). Furthermore, we observed more sustained decoding in participants with WS (M = 62.90 ± 4.93% comparatively to TD individuals (M = 39.39 ± 5.99%, t(20) = -3.031, p = 0.007, Cohen's d = −1.29). There was no difference however in peak decoding accuracy level (t(20) = -1.29, p = 0.21, d = −0.55). Fig. 3, lower panel, Orientation, presents these metrics as violin plots with each individual participant data point plotted, alongside the median, and distribution. This pattern of results seems to suggest that while participants with WS show sensitivity to face orientation, it nonetheless seems to occur in an atypical manner.

3.2.4. Discussion

Several researchers have suggested that ostensible strengths in face processing behaviour noted in individuals with WS may be underpinned by atypical mechanisms. Experiment 2 constitutes the most comprehensive investigation to date of the neural underpinnings of face-selectivity and face processing strategies (indexed indirectly by looking at EEG activity associated with face inversion) in this group. Crucially, using state-of-the-art MVPA analyses, we demonstrated that all of our participants with WS showed a distinct pattern of neural activity to faces compared to objects, reflecting a strong and clear profile of differentiated responses to faces *as a category*. Furthermore, all individuals with WS showed differential neural processing of upright vs. inverted faces. Nonetheless, the precise pattern of responses with respect to face orientation differed between the groups. Overall, this result is

---

[8] Note that one TD participant did not meet the criterion for inclusion in the onset analysis and was removed from this test.
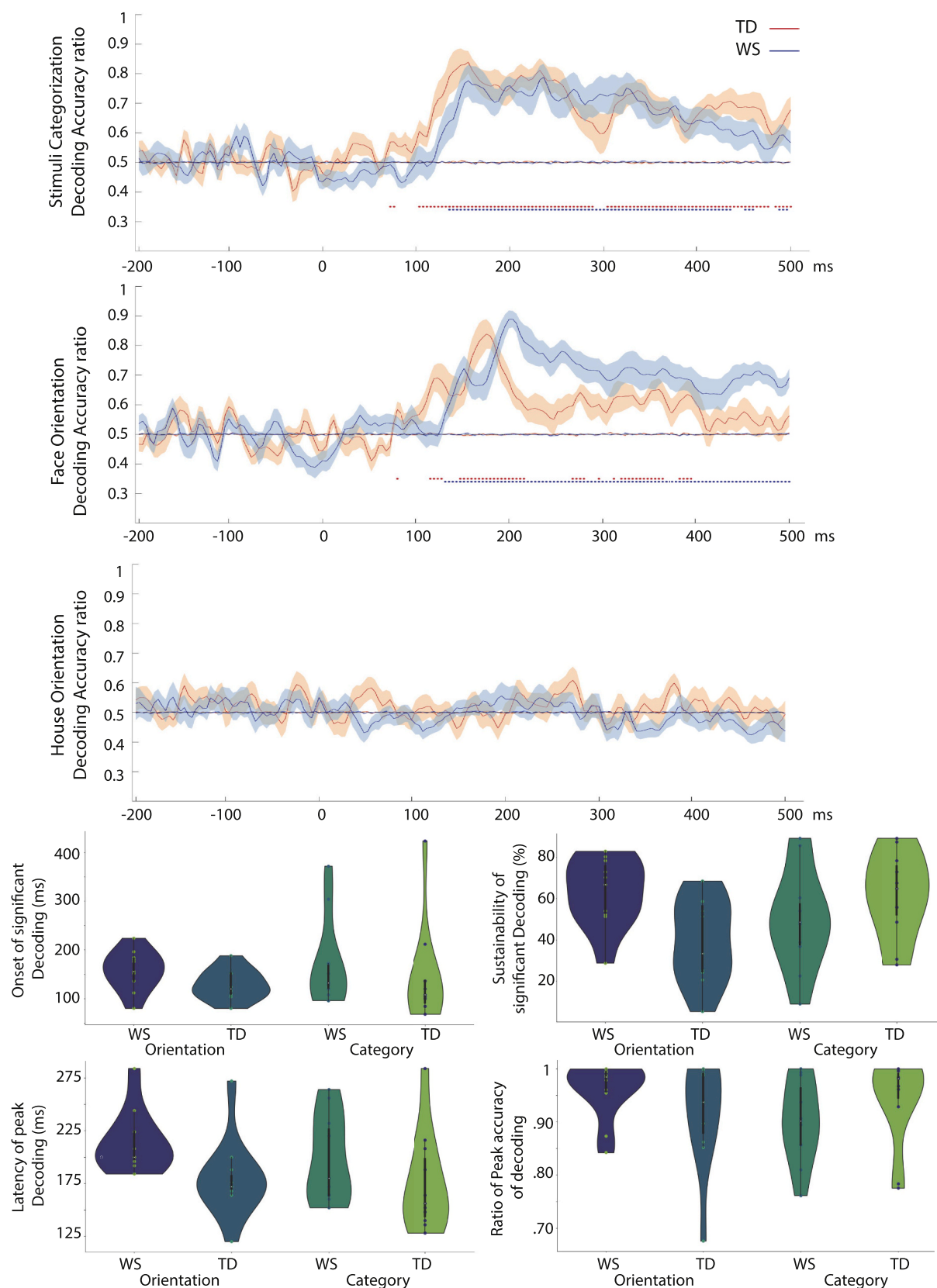
Fig. 3. Classification accuracy time-course for the three binary comparisons; i.e. stimuli categorization (upright faces vs houses; top row), face orientation (upright to inverted faces; middle row) and house orientation (upright to inverted houses; bottom row). Above chance classification are indicated by colour coded dots at the base of the curves (p < 0.05, one-tailed, Group-level, FDR corrected). Violin plots are presented in the 4th and 5th row to covering the four metrics used, onset of decoding, sustainability of decoding, latency of peak decoding and peak decoding accuracy. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

consistent with the use of holistic processing in WS – which is relatively enhanced for upright compared to inverted faces - but suggests that this group differs from typical individuals in the extent to which they rely on this face processing mechanism.

We conducted two analyses of the recorded EEG data. The first was traditional ERP analysis of the classically reported components analysed in TD literature using a constrained set of channels and time windows. The second involved an innovative, more atheoretical Multi Variate Pattern Analysis (MVPA) technique, which uses a broader selection of the recorded data across a wide post-stimulus time window to identify the presence of distinctive neural patterns of activation at the level of groups and individual participants. We investigated three comparisons. The first related to the ability to decode faces as a distinct category (compared to houses). This comparison determined whether individuals with WS have face-selective neural coding, as in the typical population. The second related to decoding of face orientation (upright vs. inverted faces). The FIE is considered an important marker of the specialised processing of faces compared to other object categories, and is widely accepted to be driven by a holistic processing of faces when in their canonical upright orientation that is not available to the same extent for faces when presented inverted or for other object categories (though see Rossion and Curran, 2010 for a suggestion that this privileged processing extends to objects of particular expertise in addition to faces). Thus, between-group differences in the neural response to face inversion could highlight a difference in the mechanisms used to support face processing in WS compared to the TD group. Finally, we also compared the ability to classify upright and inverted house stimuli. This comparison was designed to determine the selectivity of our face orientation results.

We found that it was possible to decode differential neural activity in response to faces and houses in all participants in both the WS and the TD groups. When we investigated these stimulus category effects at the group level it appeared that the TD group displayed earlier face decoding than the WS group, and a higher peak decoding accuracy, yet, these differences were not statistically significant. Comparisons of individual level data did not support group differences across any of the four measures considered: onset and sustainability of decoding; peak decoding latency and peak decoding accuracy. Thus, the ability to broadly classify a face (cf. a house) seems to follow a typical neural pattern in WS. It is interesting to note that developmental research with neurotypical individuals suggests that this face selective expertise matures in early childhood: from at least 6 years of age there are no significant differences relative to adults on these measures (Mares et al., 2020).

Although speculative, the early maturation of face-selective processing in the typical population could partially account for the relative strength in face processing in WS. That is, early maturing processes are thought to be less vulnerable to atypical development (see Farran and Formby, 2012). This argument has been made for the development of the dorsal visual stream vs. the ventral visual stream (Braddick et al., 2003). In WS (and other disorder groups), there is a relative impairment in the functions of the dorsal stream when compared to the ventral stream. This has been attributed to the relative vulnerability of the dorsal visual stream to atypical development because it has a more protracted development, in the typical population, than the ventral visual stream (Atkinson et al., 1997; Braddick et al., 2003). This argument also holds some utility for understanding face-expertise, although it must be considered within the context that competency in a skill is the product of complex coordination of multiple brain areas and multiple cognitive processes (see Farran and Formby, 2012), and thus that this relative strength is not observed across all atypical groups.

The relative strength in face processing can also be considered within the context of the theoretical arguments for the relative impairment in other areas of non-verbal ability. As discussed by Landau and Hoffman (2012)), the critical deficit in visuo-spatial processing in WS reflects difficulties in visuo-spatial construction, a function of the dorsal visual stream. The authors explain that the visuo-spatial deficit observed in WS

does not reflect broader atypicalities in ventral stream functioning, areas important for face processing. As discussed earlier, this is reflected by O'Hearn et al. (2011) who demonstrated that ventral stream activity for faces in WS was the same as chronological aged matched TD controls.

The MVPA results for face-selective processing are largely echoed in the ERP analyses. Both the TD and WS groups showed a larger N170 amplitude for faces than for houses, and the WS group also showed sensitivity to faces with respect to N170 latency (faster N170 to faces relative to houses). The presence of the hallmark face selective N170 signature in the WS group demonstrates that not only do individuals with WS differentiate faces and houses categorically at a neural level, but they process faces as a category in a typical manner with respect to neural activation. Interestingly an earlier N170 peak for faces compared to houses was observed in WS but not in TD. An earlier N170 component for faces compared to other objects has previously been reported in adults (Itier and Taylor, 2004), albeit not consistently (Sadeh et al., 2008), suggesting that this effect is not robust. Few studies have previously investigated face-selective processing using EEG in WS, and not specifically the N170 component (Key and Dykens, 2016). Key and Dykens (2016) also found more negative amplitudes in the left hemisphere for faces compared to houses in participants with WS, albeit in an earlier time window (108–156 ms, this effect was not present on the right hemisphere). The current results are further consistent with fMRI findings reporting that although the Fusiform Face Area is enlarged in WS, it is associated with typical face-selective processing (Golarai et al., 2010).

The MVPA stimulus orientation analysis showed that the neural response in all participants, in both groups, was sensitive to face orientation, which supports holistic face processing in WS, to some extent. Crucially, this effect was face sensitive with no differences in neural activation observed for upright compared with inverted houses. This finding is important because it suggests, using a very sensitive analysis, at an individual level, that an account of face processing in which individuals with WS exclusively rely on featural processing is not supported.

There were also group differences in both MVPA and ERP variables in response to face-orientation. These subtle differences provide a valuable insight into the atypicalities in face processing in WS. Individual level analyses demonstrated statistically that the TD group decoded orientation marginally earlier than the WS group, and also that there was significantly more sustained decoding in the WS group relative to the TD group. Similarly, the time point of peak face orientation decoding was significantly earlier in the TD group than the WS group. This set of results demonstrates subtle atypicalities in the neural response to faces in WS, relative to the typical population, however, given the novelty of the analysis approach interpreting these findings is challenging until further detailed work characterising both typical and atypical profiles can be undertaken. Emerging evidence suggests that mature face orientation processing (in adults compared to children) is associated with both an increased peak decoding coupled with more sustained decoding, while face expertise in typical adults is associated with both these factors as well as an earlier decoding onset (Mares et al., 2020). Here, while there was a trend for earlier onset decoding of face orientation in TD adults, this delay in WS emerged significantly on the latency of peak decoding. This is in line with an atypical, and perhaps less expert face processing in the WS group. Yet the WS group showed considerably more sustained decoding than the TD adults. It is difficult to interpret the latter result as an index of more advanced face processing ability in WS than TD adults, given the superior face recognition ability observed in TD adults. It is possible that this pattern reflects the increased social interest reported in WS and reflects attention to faces rather than strong face processing abilities per se. Golarai et al. (2010) made a similar suggestion of a brain-behaviour association with respect to attention to faces and the enlarged Fusiform Face Area in WS. We note, however, that while both groups showed sustained decoding at the group level for face category decoding, with no significant differences

between groups, the TD adult group did not show this sustained pattern for face orientation decoding. This stands in contrast to recent evidence in a larger sample of young TD adults (N = 38, 18–35yrs, Mares et al., 2020), where significant face-orientation sensitivity decoding at the group level extended across the entire time course. It may then be that differences in the composition of the current control group (ranging in age between 20 and 49 yrs) i.e. including a number of older aged participants are a contributing factor to this difference. Indeed, some studies have revealed a decreased neural sensitivity to face inversion in older participants, concomitant with a preserved sensitivity to faces compared to other object categories (e.g. Daniel and Bentin, 2012; Gao et al., 2009). Of note, these studies had older samples than the current one (>61 yrs), and although they support a decreased sensitivity to face inversion in aging, more studies are needed to understand how this effect changes through the lifespan.

Further group differences were observed for face orientation processing from the classic ERP analyses. Both the WS and TD adults demonstrated neural sensitivity to face inversion in terms of their N170 response profiles. However, for the WS group, this was reflected by differences in N170 latency (faster response to upright compared to inverted faces), whilst for the TD adults, this was reflected more classically in large N170 amplitude differences (larger response to inverted than upright faces with no similar difference for houses). The lack of sensitivity with respect to N170 latency in the TD adult group is consistent with the face-selectivity pattern of results for this group. As suggested above differences in N170 latency may not be as robustly face-selective effects as differences in N170 amplitude. The absence of significant inversion effects in the WS group for either category at the level of the amplitude of the N170 component is consistent with the results of Grice et al. (2001) who also demonstrated the classic N170 pattern for face inversion for latency, but not amplitude in WS.

This consistency in findings, coupled with evidence from the behavioural data (e.g., Deruelle et al., 1999; Karmiloff-Smith et al., 2004; Annaz et al., 2009; Riby et al., 2009) again suggests that individuals with WS demonstrate some holistic processing of faces, but may not be as reliant on this style of processing as TD individuals. Inverted faces are thought to be more heavily processed via featural information (Tanaka and Farah, 1993). Whilst the MVPA analysis clearly demonstrated differential processing of upright and inverted faces in WS, the lack of the typical increased amplitude N170 for inverted faces compared with upright in the WS group suggests an alteration of their holistic face processing, even when a face is presented upright.

We note that there were no effects of stimulus category for either group for the P100 component. This is unsurprising given mixed support for the P100 as reflecting face-selective neural activity (see Rossion and Jacques, 2008 for a discussion).

4. General discussion

Experiments 1 and 2 provide behavioural and neural data respectively. In Experiment 1, we used a highly sensitive, widely-used and well-normed test of face recognition memory to robustly establish that face identity processing abilities are not at a chronological age-appropriate level in adults with WS. This finding supports the notion that reports of 'typical range' performance on the BFPT (e.g., Bellugi et al., 1988) might reflect a lack of task sensitivity rather than age-appropriate levels of face processing ability (see Riby, 2012). Here, we demonstrated that face processing performance in adults with WS, as a group, is similar to that of a 10-year-old typically developing child, but that this ranges from the 5-year-old level to adult levels of performance. Thus, whilst it is possible for an individual with WS to have face processing performance in the typical range, it is not usual. That said, our results indicate that face processing broadly represents a relative strength when compared to non-verbal ability in this group. This relative strength is consistent with the relative strength in verbal ability in this

group, which should also not be described as "spared" (see Karmiloff-Smith, 2009). However, again, this pattern of strengths and weaknesses was not consistently shown for all individuals in our sample. While we have currently focused on face memory ability due to its behavioural relevance, the addition of purely perceptual tasks in future studies would allow for a more direct comparison with Experiment 2, where the neural underpinnings of face perception were analysed in the absence of a corresponding behavioural measure. The use of MVPA in this second experiment provided a more fine-grained understanding of face processing in WS than would have been possible from ERP or behavioural data alone. The inclusive nature of the MVPA method was beneficial, allowing for a broader and more atheoretical analysis of the neural signal, including at the level of individuals. We have presented robust neural evidence that individuals with WS demonstrate face-selective perceptual processing (i.e., their neural patterns allow for the distinction between whether an individual is viewing face and house categories), even during a largely passive task. A neural profile of significant category selectivity was observed in all of our participants with WS, which did not differ from the TD adult controls. Thus, with respect to neural classification of faces *as a category*, the underlying mechanisms employed in the WS group appear to be largely typical.

Our neural investigation of the *mechanisms* used to process faces suggested that there were some typical and some atypical features to the strategies employed by individuals with WS. MVPA clearly demonstrated that all of our WS group are sensitive to face inversion, suggesting reliable holistic coding. This is an important finding within the WS literature, in which there have been claims of a particular reliance on featural processing (Karmiloff-Smith et al., 2004). Importantly, our EEG analyses also demonstrated that the neural signal for coding face orientation includes atypical features. The latency of peak decoding was significantly later in participants with WS than in age-matched TD adults, while a more sustained decoding was observed in participants with WS. Similarly, the hallmark N170 signature for face inversion was only partially present in the WS group. We suggest that individuals with WS are also sensitive to the featural properties of a face, and that face-selective expertise in WS is supported by the use of holistic as well as featural processing. This contrasts to TD adults, where expertise may be more critically linked with holistic processing (e.g. DeGutis et al., 2013; Wang et al., 2012). This result also suggests that the behavioural data presented in Experiment 1 might be supported by atypical mechanisms in WS.

The results of Experiments 1 and 2 powerfully extend our current understanding of both the behavioural and neural signatures of face processing in WS. This range of methods has enabled us to better understand face processing in the neurodevelopmental disorder at a more fine-grained level than has hitherto been considered. Furthermore, by using analysis of both group level and individual level data we were also able demonstrate the robust finding that all individuals with WS in our sample were sensitive both to face category and to face inversion (a hallmark of specialised face processing), but that this was coupled with broad heterogeneity in the extent to which face processing can be characterised as a relative strength within their broader cognitive profile.

CRediT authorship contribution statement

Emily K. Farran: Writing - original draft, Writing - review & editing, Formal analysis, Funding acquisition, Methodology, Conceptualization. **Inês Mares:** Writing - original draft, Writing - review & editing, Formal analysis, Data curation, Methodology. **Michael Papasavva:** Investigation, Writing - original draft, Writing - review & editing. **Fraser W. Smith:** Writing - review & editing, Formal analysis, Methodology. **Louise Ewing:** Writing - review & editing, Investigation, Conceptualization, Methodology, Funding acquisition, Project administration. **Marie L. Smith:** Writing - review & editing, Formal analysis, Conceptualization, Methodology, Funding acquisition, Project administration.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2020.107440.

References

Anastasi, J.S., Rhodes, M.G., 2005. An own-age bias in face recognition for children and older adults. Psychon. Bull. Rev. 12 (6), 1043–1047. https://doi.org/10.3758/BF03206441.

Andrews, S., Burton, A.M., Schweinberger, S.R., Wiese, H., 2017. Event-related potentials reveal the development of stable face representations from natural variability. Q. J. Exp. Psychol. 70 (8), 1620–1632. https://doi.org/10.1080/17470218.2016.1195851.

Annaz, D., Karmiloff-Smith, A., Johnson, M.H., Thomas, M.S.C., 2009. A cross-syndrome study of the development of holistic face recognition in children with autism, Down syndrome, and Williams syndrome. J. Exp. Child Psychol. 102, 456–486.

Atkinson, J., King, J., Bradick, O., Nokes, L., Anker, S., Braddick, F., 1997. A specific deficit of dorsal stream function in Williams syndrome. Neuroreport: Cognitive Neurosci. Neuropsychol. 8, 1919–1922.

Bellugi, U., Lichtenberger, E., Mills, D., Galaburda, A., Korenberg, J.R., 1999. Bridging cognition, brain, and molecular genetics: evidence from Williams syndrome. Trends Neurosci. 5, 197–208.

Bellugi, U., Sabo, H., Vaid, J., 1988. Spatial deficits in children with Williams syndrome. In: Stiles-Davis, J., Kritchevshy, U., Bellugi, U. (Eds.), Spatial Cognition: Brain Bases and Development. Lawrence Erlbaum, Hillsdale, New Jersey, pp. 273–297.

Bentin, S, Allison, T, Puce, A, Perez, E, McCarthy, G, 1996. Electrophysiological studies of face perception in humans. J. Cognit. Neurosci. 8 (6), 551–565.

Benton, A., Sivan, A., Hamsher, K., Varney, N., Spreen, O., 1994. Contributions to Neuropsychological Assessment. Oxford University Press, New York.

Braddick, O., Atkinson, J., Wattam-Bell, J., 2003. Normal and anomalous development of visual motion processing: motion coherence and 'dorsal stream vulnerability'. Neuropsychologia 41, 1769–1783.

Cauchoix, M., Barragan-Jason, G., Serre, T., Barbeau, E.J., 2014. The neural dynamics of face detection in the wild revealed by MVPA. Journal of Neuroscience 34 (3), 846–854.

Croydon, A., Pimperton, H., Ewing, L., Duchaine, B.C., Pellicano, E., 2014. The cambridge face memory test for children (CFMT-C): a new tool for measuring face recognition skills in childhood. Neuropsychologia 62, 60–67. https://doi.org/10.1016/j.neuropsychologia.2014.07.008.

Daniel, S., Bentin, S., 2012. Age-related changes in processing faces from detection to identification: ERP evidence. Neurobiol. Aging 33 (1). https://doi.org/10.1016/j.neurobiolaging.2010.09.001, 206.e1-206.e28.

DeGutis, J., Wilmer, J., Mercado, R.J., Cohan, S., 2013. Using regression to measure holistic face processing reveals a strong link with face recognition ability. Cognition 126 (1), 87–100.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

Deruelle, C., Mancini, J., Livet, M.O., Casse-Perrot, C., deSchonen, S., 1999. Configural and local processing in Williams syndrome. Brain Cognit. 41, 276–298.

Di Russo, F., Martínez, A., Sereno, M.I., Pitzalis, S., Hillyard, S.A., 2002. Cortical sources of the early components of the visual evoked potential. Hum. Brain Mapp. 15 (2), 95–111. https://doi.org/10.1002/hbm.10010.

Duchaine, B., Nakayama, K., 2006. The Cambridge Face Memory Test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. Neuropsychologia 44 (4), 576–585. https://doi.org/10.1016/j.neuropsychologia.2005.07.001.

Duchaine, B.C., Nakayama, K., 2004. Developmental prosopagnosia and the Benton facial recognition test. Neurology 62, 1219–1220.

Duchaine, B.C., Weidenfeld, A., 2003. An evaluation of two commonly used tests of unfamiliar face recognition. Neuropsychologia 41, 713–720.

Dunlap, W.P., Cortina, J.M., Vaslow, J.B., Burke, M.J., 1996. Meta-analysis of experiments with matched groups or repeated measures designs. Psychol. Methods 1 (2), 170–177. https://doi.org/10.1037/1082-989X.1.2.170.

Dunn, D.M., Dunn, L.M., Styles, B., Sewell, J., 2009. The British Picture Vocabulary Scale, third ed. GL Assessment, London.

Eimer, M., 2000. Effects of face inversion on the structural encoding and recognition of faces - evidence from event-related brain potentials. Cognit. Brain Res. 10 (1–2), 145–158. https://doi.org/10.1016/S0926-6410(00)00038-0.

Ewing, L., Karmiloff-Smith, A., Farran, E.K., Smith, M.L., 2017a. Distinct profiles of information-use characterize identity judgments in children and low-expertise adults. J. Exp. Psychol. Hum. Percept. Perform. 43, 1937–1943. https://doi.org/10.1037/xhp0000455.

Ewing, L., Farran, E.K., Karmiloff-Smith, A., Smith, M.L., 2017b. Understanding strategic information use during emotional expression judgments in Williams syndrome. Dev. Neuropsychol. 42, 323–335. https://doi.org/10.1080/87565641.2017.1353995.

Farran, E.K., 2020. An analytical approach to visuospatial cognition: What can neurodevelopmental disorders tell us about developmental pathways? Psyarxiv. https://doi.org/10.31234/osf.io/yb76e.

Farran, E.K., Formby, S., 2012. Visual perception and visuospatial cognition. In: Farran, E.K., Karmiloff-Smith, A. (Eds.), Neurodevelopmental Disorders across the Lifespan: A Neuroconstructivist Approach. Oxford University Press, pp. 225–246.

Farran, E.K., Karmiloff-Smith, A. (Eds.), 2012. Neurodevelopmental Disorders across the Lifespan: A Neuroconstructivist Approach. Oxford University Press.

Farran, E.K., Formby, S., Daniyal, F., Holmes, T., Van Herwegen, J., 2016. Route-learning strategies in typical and atypical development; eye tracking reveals atypical landmark selection in Williams syndrome. J. Intellect. Disabil. 60, 933–944. https://doi.org/10.1111/jir.12331.

Gallivan, J.P., McLean, D.A., Valyear, K.F., Culham, J.C., 2013. Decoding the neural mechanisms of human tool use. ELife 2 (2), 1–29. https://doi.org/10.7554/eLife.00425.

Gao, L., Xu, J., Zhang, B., Zhao, L., Harel, A., Bentin, S., 2009. Aging effects on early-stage face perception: an ERP study. Psychophysiology 46 (5), 970–983. https://doi.org/10.1111/j.1469-8986.2009.00853.x.

Golarai, G., Hong, S., Haas, B.W., Galaburda, A.M., Mills, D.L., Bellugi, U., Reiss, A.L., 2010. The fusiform face area is enlarged in Williams syndrome. Journal of Neuroscience 30 (19), 6700–6712.

Gosling, A., Eimer, M., 2011. An event-related brain potential study of explicit face recognition. Neuropsychologia 49 (9), 2736–2745. https://doi.org/10.1016/j.neuropsychologia.2011.05.025.

Grice, S.J., Spratling, M.W., Karmiloff-Smith, A., Halit, H., Csibra, G., de Haan, M., Johnson, M.H., 2001. Disordered visual processing and oscillatory brain activity in autism and Williams syndrome. Neuroreport 12, 2697–2700.

Itier, R.J., Taylor, M.J., 2004. N170 or N1? Spatiotemporal differences between object and face processing using ERPs. Cerebr. Cortex 14 (2), 132–142. https://doi.org/10.1093/cercor/bhg111.

Jarrold, C., Baddeley, A.D., Hewes, A.K., 1999. Genetically dissociated components of working memory: evidence from Down's and Williams syndrome. Neuropsychologia 37, 637–651.

Johnson, M.H., Dziurawiec, S., Ellis, H., Morton, J., 1991. Newborns' preferential tracking of face-like stimuli and its subsequent decline. Cognition 40, 1–19.

Karmiloff-Smith, A., 1997. Crucial differences between developmental cognitive neuroscience and adult neuropsychology. Dev. Neuropsychol. 13, 513–524.

Karmiloff-Smith, A., 2009. Nativism versus neuroconstructivism: rethinking the study of developmental disorders. Dev. Psychol. 45, 56–63.

Karmiloff-Smith, A., Al-Janabi, T., D'Souza, H., Groet, J., Massand, E., Mok, K., Strydom, A., 2016. The importance of understanding individual differences in Down syndrome. F1000Research 5. https://doi.org/10.12688/f1000research.7506.1.

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.P., Dykens, E.M., 2016. Face repetition detection and social interest: an ERP study in adults with and without Williams syndrome. Soc. Neurosci. 11 (6), 652–664. https://doi.org/10.1080/17470919.2015.1130743.

Kuefner, D., de Heering, A., Jacques, C., Palmero-Soler, E., Rossion, B., 2010. Early Visually Evoked Electrophysiological Responses Over the Human Brain (P1, N170) Show Stable Patterns of Face-Sensitivity from 4 years to Adulthood. Frontiers in Human Neuroscience 3 (January), 67. https://doi.org/10.3389/neuro.09.067.2009.

Landau, B, Hoffman, E, 2012. *Spatial representation: From gene to mind.*. Oxford University Press.

Leder, H., Bruce, V., 2000. Inverting line drawings of faces. Swiss J. Psychol. Schweiz. Z. Psychol. Rev. Suisse Psychol. 59, 159–169.

Lenhoff, H.M., Wang, P.P., Greenberg, F., Bellugi, U., 1997. Williams syndrome and the brain. Sci. Am. 277 (6), 42–47.

Mares, I., Ewing, L., Farran, E.K., Smith, F.W., Smith, M., 2020. Developmental changes in the processing of faces as revealed by EEG decoding. Neuroimage, 116660.

Mervis, C.B., Robinson, B.F., Bertrand, J., Morris, C.A., Klein-Tasman, B.P., Armstrong, S.C., 2000. The Williams syndrome cognitive profile. Brain Cognit. 44, 604–628.

Mills, D.L., Dai, L., Fishman, I., Yam, A., Appelbaum, L.G., St George, M., et al., 2013. Genetic mapping of brain plasticity across development in Williams syndrome: ERP markers of face and language processing. Dev. Neuropsychol. 38 (8), 613–642 pmid: 24219698.

Mills, D.L., Alvarez, T.D., St George, M., Appelbaum, L.G., Bellugi, U., Neville, H., 2000. Electrophysiological studies of face processing in Williams syndrome. J. Cognit. Neurosci. 12, 47–64.

Mobbs, D., Garrett, A.S., Menon, V., Rose, F.E., Bellugi, U., Reiss, A.L., 2004. Anomalous brain activation during face and gaze processing in Williams syndrome. Neurology 62, 2070–2076.

Nakamura, M., Watanabe, S., Inagaki, M., Hirai, M., Miki, K., Honda, Y., et al., 2013. Electrophysiological study of face inversion effects in Williams syndrome. Brain Dev. 35, 323–330. https://doi.org/10.1016/j.braindev.2012.05.010.

O'Hearn, K., Roth, J.K., Courtney, S.M., et al., 2011. Object recognition in Williams syndrome: uneven ventral stream activation. Dev. Sci. 14 (3), 549–565. https://doi.org/10.1111/j.1467-7687.2010.01002.x.

Oruç, I., Krigolson, O., Dalrymple, K., Nagamatsu, L.S., Handy, T.C., Barton, J.J.S., 2011. Bootstrap analysis of the single subject with event related potentials. Cogn. Neuropsychol. 28 (5), 322–337. https://doi.org/10.1080/02643294.2011.648176.

Pereira, F., Mitchell, T., Botvinick, M., 2009. Machine learning classifiers and fMRI: a tutorial overview. Neuroimage 45, S199–S209. https://doi.org/10.1016/j.neuroimage.2008.11.007.

Ramon, M., Gobbini, M.I., 2018. Familiarity matters: a review on prioritized processing of personally familiar faces. Vis. Cognit. 26 (3), 179–195.

Raven, J., 2008. Coloured Progressive Matrices and Crichton Vocabulary ScaleManual. Pearson, London.

Rezlescu, C., Susilo, T., Wilmer, J.B., Caramazza, A., 2017. The inversion, part-whole, and composite effects reflect distinct perceptual mechanisms with varied relationships to face recognition. J. Exp. Psychol. Hum. Percept. Perform. 43 (12), 1961–1973. https://doi.org/10.1037/xhp0000400.

Riby, 2012. Face processing and social interaction. In: Farran, E.K., Karmiloff-Smith, A. (Eds.), Neurodevelopmental Disorders across the Lifespan: A Neuroconstructivist Approach. Oxford University Press, pp. 225–246.

Riby, D.M., Doherty-Sneddon, G., Bruce, V., 2009. The eyes or the mouth? Feature salience and unfamiliar face processing in Williams syndrome and autism. Q. J. Exp. Psychol. 62 (1), 189–203.

Rose, F.E., Lincoln, A.J., Lai, Z., Ene, M., Searcy, Y.M., Bellugi, U., 2006. Orientation and affective expression effects on face recognition in Williams syndrome and autism. J. Autism Dev. Disord. 37, 513–522.

Rossion, B., Caharel, S., 2011. ERP evidence for the speed of face categorization in the human brain: disentangling the contribution of low-level visual cues from face perception. Vis. Res. 51 (12), 1297–1311. https://doi.org/10.1016/j.visres.2011.04.003.

Rossion, B., Curran, T., 2010. Visual expertise with pictures of cars correlates with rt magnitude of the car inversion effect. Perception 39 (2), 173–183. https://doi.org/10.1068/p6270.

Rossion, B., Jacques, C., 2008. Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. Neuroimage 39, 1959–1979.

Sadeh, B., Zhdanov, A., Podlipsky, I., Hendler, T., Yovel, G., 2008. The validity of the face-selective ERP N170 component during simultaneous recording with functional MRI. Neuroimage 42 (2), 778–786. https://doi.org/10.1016/j.neuroimage.2008.04.168.

Schweinberger, R, Pickering, C, Jentzsch, I, Burton, M, Kaufmann, M, 2002. Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. Cognit. Brain Res. 14 (3), 398–409.

Schyns, P.G., Oliva, A., 1999. Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. Cognition 69 (3), 243–265. https://doi.org/10.1016/S0010-0277(98)00069-9.

Smith, F.W., Muckli, L., 2010. Nonstimulated early visual areas carry information about surrounding context. Proc. Natl. Acad. Sci. Unit. States Am. 107 (46), 20099–20103. https://doi.org/10.1073/pnas.1000233107.

Smith, F.W., Smith, M.L., 2019. NeuroImage Decoding the dynamic representation of facial expressions of emotion in explicit and incidental tasks. Neuroimage 195 (September 2018), 261–271. https://doi.org/10.1016/j.neuroimage.2019.03.065.

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

Tanaka, J.W., Farah, M.J., 1993. Parts and wholes in face recognition. Q. J. Exp. Psychol. 46, 225–245.

Taylor, M.J., Batty, M., Itier, R.J., 2004. The faces of development: a review of early face processing over childhood. J. Cognit. Neurosci. 16 (8), 1426–1442. https://doi.org/10.1162/0898929042304732.

Thomas, M.S.C., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., Karmiloff-Smith, A., 2009. Using developmental trajectories to understand genetic disorders. J. Speech Lang. Hear. Res. 52, 336–358.

Udwin, O., Yule, W., 1991. A cognitive and behavioural phenotype in Williams syndrome. J. Clin. Exp. Neuropsychol. 13, 232–244.

Vicari, S., Bellucci, S., Carlesimo, G.A., 2006. Evidence from two genetic syndromes for the independence of spatial and visual working memory *Developmental Medicine and child*. Neurology 48, 126–131.

Wang, R., Li, J., Fang, H., Tian, M., Liu, J., 2012. Individual differences in holistic processing predict face recognition ability. Psychol. Sci. 23 (2), 169–177.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G.O., Gosselin, F., Tanaka, J.W., 2010. Controlling low-level image properties: the SHINE toolbox. Behav. Res. Methods 42 (3), 671–684. https://doi.org/10.3758/BRM.42.3.671.

Wilmer, J.B., 2017. Individual differences in face recognition: a decade of discovery. Curr. Dir. Psychol. Sci. 26 (3), 225–230. https://doi.org/10.1177/0963721417710693.

Wilson, B., Cockburn, J., Baddeley, A., Hiorns, R., 1989. The development and validation of a test battery for detecting and monitoring everyday memory problems. J. Clin. Exp. Neuropsychol. 11 (6), 855–870.

Yin, R.K., 1969. Looking at upside-down faces. J. Exp. Psychol. 81 (1), 141–145. https://doi.org/10.1037/h0027474.

Young, A.W., Burton, A.M., 2018. Are we face experts? Trends Cognit. Sci. 22 (2), 100–110. https://doi.org/10.1016/j.tics.2017.11.007.

Young, A.W., Hellawell, D., Hay, D.C., 1987. Configural information in face perception. Perception 16, 747–759.