Similarities in Autistic and Neurotypical Visual–Haptic Perception When Making Judgements About Conflicting Sensory Stimuli

Daniel Poole^{1,*}, Ellen Poliakoff¹, Emma Gowen¹, Samuel Couth²,
Rebecca A. Champion¹ and Paul A. Warren¹

¹ Division of Neuroscience and Experimental Psychology,
University of Manchester, Oxford Road, M139PL, UK

² Division of Human Communication, Development and Hearing,
University of Manchester, UK

Received 28 July 2016; accepted 7 February 2017

Abstract

A number of studies have shown that multisensory performance is well predicted by a statistically optimal maximum likelihood estimation (MLE) model. Under this model unisensory estimates are combined additively and weighted according to relative reliability. Recent theories have proposed that atypical sensation and perception commonly reported in autism spectrum condition (ASC) may result from differences in the use of reliability information. Furthermore, experimental studies have indicated that multisensory processing is less effective in those with the condition in comparison to neurotypical (NT) controls. In the present study, adults with ASC ($n = 13$) and a matched NT group ($n = 13$) completed a visual–haptic size judgement task (cf. Gori *et al.*, 2008) in which participants compared the height of wooden blocks using either vision or haptics, and in a dual modality condition in which visual–haptic stimuli were presented in size conflict. Participants with ASC tended to produce more reliable estimates than the NT group. However, dual modality performance was not well predicted by the MLE model for either group. Performance was subsequently compared to alternative models in which the participant either switched between modalities trial to trial (rather than integrating) and a model of non-optimal integration. Performance of both groups was statistically comparable to the cue-switching model. These findings suggest that adults with ASC adopted a similar strategy to NTs when processing conflicting visual–haptic information. Findings are discussed in relation to multisensory perception in ASC and methodological considerations associated with multisensory conflict paradigms.

* To whom correspondence should be addressed. E-mail: daniel.poole@manchester.ac.uk

Keywords

Visual–haptic processing, autism spectrum condition, maximum likelihood estimate, cue-switching, size judgement, multisensory processing

1. Introduction

Autism Spectrum Condition (ASC) is diagnostically characterised by the presence of social and communication deficits, and restricted repetitive behaviours (DSM-V: American Psychiatric Association, 2013). Although ASC is diagnosed behaviourally, the condition is understood to be the result of atypical neurodevelopment including abnormal brain growth (Courchesne, 2004; Courchesne *et al.*, 2003), differences in connectivity between specialised neural regions (Anderson *et al.*, 2011; Brock *et al.*, 2002; Cherkassky *et al.*, 2006; Itahashi *et al.*, 2015; Just *et al.*, 2007), and increased variability in neural response (Dinstein *et al.*, 2012, 2015; Rubenstein and Merzenich, 2003). Sensory symptoms have been included as a sub-category of restricted and repetitive behaviours on the DSM-V, although these are currently considered non-pathognomonic as they are difficult to distinguish from sensory differences reported in other conditions (Grapel *et al.*, 2015; Rubenstein and Merzenich, 2003). However, it is estimated that there is up to 90% prevalence of these sensory symptoms in ASC (Crane *et al.*, 2009; Leekam *et al.*, 2007; Tomchek and Dunn, 2007), which are likely to impact on everyday functioning (Little *et al.*, 2015).

To enable effective perception, the nervous system must integrate information about properties of the environment from the different sensory modalities. For example, to estimate the location of an approaching vehicle, both visual and auditory information need to be combined to provide a coherent and accurate percept of the vehicle's position in 3D space. There is evidence to suggest that ASC affects how sensory information is combined. For instance, ERP recordings have revealed less automatic integration of multisensory information in ASC; the super-additive neural response typically generated by integrated multisensory information is markedly delayed in participants with ASC compared to neurotypical (NT) controls (Brandwein *et al.*, 2013; Russo *et al.*, 2010). In addition, individuals with ASC are less likely to produce illusory responses when presented with conflicting visual and auditory information (Irwin *et al.*, 2011; Mongillo *et al.*, 2008; Stevenson *et al.*, 2014). Furthermore, the processing of task irrelevant multisensory stimuli is atypical in ASC. For instance, participants with ASC produce responses which are influenced by task irrelevant stimuli in a different modality to the target which are presented at a greater temporal (Foss-Feig *et al.*, 2010; Kwakye *et al.*, 2011) and spatial disparity (Poole *et al.*, 2015) in comparison to controls.

Although the circumstances under which individuals with ASC process multisensory information has been considered, less work has examined *how* multisensory information is combined. In several tasks requiring estimation of some perceptual property, NT adults have been found to combine multisensory information to produce performance close to that of the statistically optimal model (see Rohde *et al.*, 2015 for a recent review). It can be shown that the optimal way to combine estimates from multiple sensory modalities relies upon taking into account the relative reliability of each estimate (Landy *et al.*, 1995). Under the Maximum Likelihood Estimate (MLE) model of cue combination, sensory estimates, E_i , of actual properties of the world, s_i (e.g., speed, depth, size) are assumed to be unbiased but corrupted by Gaussian noise:

$$E_i = s_i + N(0, \sigma_i) \quad (1)$$

The combined estimate (E_{12}) based on multiple sensory sources is then obtained as a weighted sum of the estimates (E_1, E_2) obtained from the individual stimuli (see Ernst and Banks, 2002).

$$E_{12} = w_1 E_1 + w_2 E_2 \quad (2)$$

Under MLE the weight (w_i) for each estimate depends on its reliability (inverse variance — i.e., $1/\sigma_i$).

$$\begin{aligned} w_1 &= \frac{1/\sigma_1^2}{1/\sigma_1^2 + 1/\sigma_2^2} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \\ w_2 &= \frac{1/\sigma_2^2}{1/\sigma_2^2 + 1/\sigma_1^2} = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \end{aligned} \quad (3)$$

Examining eqn. (3) we see that under MLE more reliable estimates are given increased weight relative to those that are less reliable. This combination rule ensures that the combined estimate is more reliable than either individual estimate [i.e., $\sigma_{12} < \min(\sigma_1, \sigma_2)$] and is optimal in providing the combined estimate with the highest possible combined reliability (lowest variance):

$$\sigma_{12} = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} \quad (4)$$

The MLE model has been shown to provide a good approximation of human multisensory estimates as measured using cue-conflict paradigms. In a seminal study, Ernst and Banks (2002) asked participants to make comparative height judgements of virtual reality bars. Performance was measured in separate visual and haptic conditions, and a dual modality condition in which participants were presented with visual and haptic stimuli together. Based on eqn. (4) the MLE model predicts that variance in estimated height in the multisensory conditions can be predicted from the measured individual modality

variances. Specifically, if σ_v^2 and σ_h^2 denote the visual and haptic variance the multisensory variance is predicted as:

$$\sigma_{vh}^2 = \frac{\sigma_v^2 \sigma_h^2}{\sigma_h^2 + \sigma_v^2} \quad (5)$$

Furthermore, the dual modality condition in Ernst and Banks' experiment included trials in which visual–haptic stimuli were presented in height conflict. Using this data it is possible to measure whether the participant's point of subjective equality (PSE) is biased towards the size of the visual or haptic stimulus, thereby indicating which cue was given greater weight in the combined estimate. It is then possible to compare performance in each conflict condition with an MLE prediction of the PSE. Using eqns (2) and (3) under MLE the PSE is predicted as E_{vh} :

$$E_{vh} = w_v E_v + w_h E_h \quad (6)$$

where the weights for each modality are calculated based on the unisensory variance estimates:

$$\begin{aligned} w_h &= \frac{\sigma_v^2}{\sigma_h^2 + \sigma_v^2} \\ w_v &= \frac{\sigma_h^2}{\sigma_h^2 + \sigma_v^2} \end{aligned} \quad (7)$$

Ernst and Banks found that participant data was indeed consistent with the MLE model. Dual modality variance was lower than that measured for either unisensory condition and approximated the predictions derived from eqn. (5). Variance was lower for visual estimates than haptic, and in accordance with eqns (6) and (7), performance in each dual modality conflict condition showed a bias towards visual information.

The MLE model has been shown to robustly predict multisensory performance across a number of different task types including visual–haptic estimates of size (Gepshtein and Banks, 2003) and shape (Helbig and Ernst, 2007) as well as visual–auditory estimates of location (Alais and Burr, 2004). Furthermore, the consistency of behaviour with MLE for visual–haptic integration appears to be spatially modulated. When participants completed a size judgement task for virtual reality bars in apparent spatial coincidence, dual modality estimates were well predicted by the MLE model (Gepshtein *et al.*, 2005). However, when the visual–haptic stimuli appeared at a horizontal offset of 60 mm or more, performance in the multisensory condition approximated that of the unisensory condition. A similar task presented a 'tool', visually linking the position of the participant's fingers to the visual bar across a spatial offset (Takahashi *et al.*, 2009). If the tool was consistent with natural use

(appearing to be manipulated by the front of the participants fingers), performance approximated the predictions of the MLE model. This suggests that the optimal weighting of visual–haptic cues occurs when both stimuli are presented within the representation of peripersonal space around the hand.

In recent research the developmental trajectory of MLE has been studied. It has been reported that young children do not weight multisensory cues optimally, with the process maturing across development (Gori *et al.*, 2008). Children aged 5, 6, 8 and 10 years old completed a real life version of the visual–haptic size judgement paradigm developed by Ernst and Banks (2002). The performance of adults and children aged 10 years old was well predicted by the MLE model. Children aged 8 years old weighted their estimates according to relative reliabilities, but performance was not well predicted by the model. The estimates of children aged 5 and 6 were counter to the predictions of the MLE model. As with the adult participants, variance was lower in the unimodal visual trials than haptic trials. However, performance in the dual modality conflict conditions was biased towards a haptic standard. This suggests that, for younger children, the less reliable haptic estimate was given greater weight in the multisensory judgement. The authors proposed that cue-integration does not occur in young children as it is more important for the perceptual systems to recalibrate with the growing body.

Since optimal multisensory cue integration appears to mature across development, we hypothesised that atypical neuronal development in ASC may impact on the maturation of this process. For instance, increased neural excitatory activity in sensory systems in ASC (Rubenstein and Merzenich, 2003) may require recalibration beyond early childhood. The perceptual benefits of optimal cue integration (i.e. less variable estimates) may be traded for a more plastic interaction between the sensory modalities into adulthood. However, if a less reliable sensory modality is given greater weight in a multisensory estimate, then the percept will be noisier. It is likely that noisy perceptual estimates would result in atypical sensory experiences as perception of the environment would be less predictable. Indeed, recent theories have suggested that atypical sensory experiences in ASC are the result of differences in how reliability information is integrated in the brain (Lawson *et al.*, 2014; Pellicano, 2013; Pellicano and Burr, 2012). A recent study of visual-vestibular integration on a self-motion estimation task indicated that participants with ASC performed in a manner consistent with the MLE model (Zaidel *et al.*, 2015). However, participants with ASC utilised a less flexible strategy in adapting to increasing noise in the visual stimuli across trials than controls, suggesting divergent strategies in integrating reliability information. There is also evidence that children with ASC may up-weight less reliable proprioceptive information on multisensory tasks, including a motor co-ordination task (Haswell *et al.*, 2009) and the rubber hand illusion (Cascio *et al.*, 2012).

In the present study, we investigated visual–haptic processing in adults with ASC and matched neurotypical (NT) controls using a real life size discrimination task similar to that of Gori *et al.* (2008). Performance consistent with the MLE model has been robustly replicated for visual–haptic processing in NT adults (e.g., Ernst and Banks, 2002; Gepshtein *et al.*, 2005; Gori *et al.*, 2008; Helbig and Ernst, 2007); and validated using this task. We investigated whether there were differences in the performance of participants with ASC and NT controls when compared to the MLE predictions of combined variance and of PSE shifts in the conflict conditions.

2. Method

2.1. Participants

Thirteen participants with ASC and 13 NT controls matched for age, IQ, gender and handedness took part in the present study. Participants were recruited through adverts in the local community, the Autistic Society Greater Manchester and Tameside Autism Network. Diagnosis was confirmed using module 4 of the Autism Diagnostic Observation Schedule (ADOS; Lord *et al.*, 2000) by a certified assessor. Participant characteristics are included in Table 1. IQ was measured using the Weschler Adult Intelligence Scale (WASI-IV; Weschler, 1999) in a separate session. Two participants with ASC and two controls were identified as left-handed using the Edinburgh Handedness Inventory (EHI, Oldfield, 1971). All participants had normal or corrected-to-normal vision (6/6 vision in both eyes as measured using the Snellen test of visual acuity), and 60 s arc stereo acuity as measured using the TNO test. No participant reported a history of a visual condition. To confirm sensory symptoms in our ASC sample all participants completed the Glasgow Sensory Quotient (GSQ; Robertson and Simmons, 2013). The GSQ is a 42-item Likert-style questionnaire designed to assess sensory differences in adults with ASC. The

Table 1.
Characteristics of participants with ASC and NT, including *t*-test statistics. Mean ± standard deviation values are presented. Significant difference denoted by asterisk

	ASC	NT	<i>t</i> (24)	<i>p</i>
Age	30.46 ± 10.22	31 ± 11.34	0.13	0.900
FSIQ	117.77 ± 10.02	112.61 ± 8.26	1.43	0.165
GSQ score (Bonferroni-corrected, $\alpha = 0.016$)				
Total sensory score	64.92 ± 8.36	42.92 ± 24.92	2.10	0.046
Total hyper score	32.62 ± 12.70	20.54 ± 10.06	2.69	0.018
Total hypo score	32.03 ± 16.98	18.69 ± 9.40	2.53	0.013*

GSQ investigates both hyper- and hypo-functioning across a number of sensory modalities (see Table 1). All participants gave written consent and the study was approved by the University of Manchester ethics committee in accordance with the Declaration of Helsinki.

2.2. Apparatus

The apparatus, design and procedure were modelled on that of Gori *et al.* (2008). Wooden blocks (the stimuli) were permanently affixed to each side of eight opaque Perspex cartridges (220 mm × 180 mm). The centre of each block was 110 mm from the top and 50 mm from the closest edge of the cartridge. The wooden blocks had a width of 150 mm and were 10 mm deep. Ten of the cartridges held blocks that ranged in height between 48–62 mm in 2 mm increments plus blocks 53 mm and 57 mm in height (the *comparison* stimuli). An eleventh cartridge held identical blocks of 55 mm (the *standard*). Two cartridges had blocks of different heights affixed to the front and back: one cartridge held a block 58 mm in height on the front and 52 mm in height on the back (referred to herein as *conflict 1*). The other held a 52 mm block on the front and a 58 mm block on the back (referred to herein as *conflict 2*). The average height of the blocks on both of these cartridges was thus 55 mm (see Fig. 1).

The cartridges were inserted into a Perspex screen which remained at approximately 40 cm viewing distance from the participant. The blocks were presented close to the participant's eye level (similar to Gori *et al.*, 2008), although viewing height was not adjusted for each participant. The nearest edge of the top facing side of each block was approximately 7° below the participant's eye level with the edge closest to the cartridge at approximately 5°. A laptop running E-Prime (Psychology Software Tools Inc., USA) was controlled by the experimenter and enabled the use of PLATO visual occlusion

Figure 1. The cartridges used in each conflict condition. (A) displays conflict 1 and (B) displays conflict 2.

Figure 2. Schematic of the experimental setup. (A) On a dual modality trial the participant views the visual stimuli on the front of a removable cartridge while exploring the haptic stimuli on the back of the cartridge using the thumb and forefinger of their dominant hand. (B) During the inter-stimulus interval, the participant returns their hand to the arm rest and the PLATO glasses were closed while the experimenter changed the cartridge. See main text for details on the geometry of the experimental setup.

spectacles (Translucent Technologies Inc., Canada) to control the visual presentation of stimuli. The laptop was also used to generate the stimulus orders, and for the experimenter to record participant's responses. See Fig. 2 for the experimental setup.

2.3. Design and Procedure

In each trial participants were asked to compare the heights of two blocks, a standard (55 mm in height) and a comparison stimulus of variable height (48, 50, 52, 53, 54, 56, 57, 58, 60, 62 mm), in a two-interval forced-choice procedure (2IFC) using the method of constant stimuli.

Trials were presented in three different conditions: visual, haptic and dual modality. In the visual condition, participants viewed the block presented on the front of the screen. To ensure the viewing distance remained consistent between the visual and dual modality conditions (described below) the participant's dominant arm remained outstretched on an arm rest. At the beginning of each trial the PLATO spectacles remained closed to allow the experimenter to position the first cartridge. Once the first cartridge was in place, the experimenter pressed a button on the laptop which triggered the spectacles to open and begin the trial. The spectacles opened for 1000 ms to allow the participant to view the stimulus, and closed during the inter-stimulus interval. The second cartridge was positioned and the spectacles opened automatically after 2000 ms to allow the participant to view the second stimulus. A sound file saying 'One' and 'Two' played simultaneously with the spectacles opening to remind the participant that they were viewing the first and second stimulus.

In the haptic conditions, participants were instructed to reach behind the screen to feel each stimulus with their thumb and forefinger. The PLATO spectacles remained closed throughout the haptic trials. After positioning the first cartridge, the experimenter pressed a button which began the trial. The participant was instructed to reach towards the stimuli on hearing the same sound file that was used in the visual condition. They were instructed to hold the stimulus for approximately 1000 ms before returning their hand to the arm support during the inter-stimulus interval. The experimenter positioned the second cartridge and the sound file instructed the participant to reach for the second stimuli after 2000 ms.

In the dual modality condition, the visual and haptic conditions were combined so that participants used their sense of vision and touch together. The standard block was always in cue-conflict in the dual modality condition and the comparisons were always cue-consistent. To increase the likelihood of the visual and haptic stimuli being processed simultaneously, the sound file preceded the opening of the PLATO spectacles by 100 ms. This allowed participants to begin to reach for the stimulus on the back of the cartridge before the spectacles opened to view the block on the front of the cartridge. Note that there were no cue-consistent presentations in the dual modality condition. Theoretically there is no reason why the cue-consistent condition is required to assess consistency with MLE. We chose to exclude this condition in order to limit the duration of our experiment, which was close to the upper limit of what is feasible for many ASC participants (~3 h).

The time profile of the stimulus presentation is represented in Fig. 3.

Participants responded verbally in all conditions which the experimenter recorded using the laptop. Breaks were included every 44 trials. The average trial length was 15.82 s (± 2.67).

There were 88 trials in each condition (visual, haptic and dual modality [conflict 1 and conflict 2]) giving a total of 352 trials. In each condition participants were presented with the 48 and 62 mm blocks six times, the 50 and 60 mm blocks eight times, and the 52, 53, 54, 56, 57 and 58 mm ten times. Trials were presented in a random order and the order of presentation (standard or comparison first) was counterbalanced for each comparison size. The visual, haptic and dual modality conditions were presented in separate blocks and the order of testing was randomised across participants.

At the start of each condition, participants completed a practice procedure of at least five trials for that condition (visual, haptic, dual modality). Participants were presented with the 48 mm or 62 mm comparison stimuli, and a no-conflict 55 mm standard on these trials. In the dual modality condition participants were specifically instructed to imagine that the front and back blocks were a single block passing through the cartridge to increase the likelihood of perceived correspondence between the two stimuli (see Ernst, 2008; Helbig

Figure 3. Profile of the stimulus presentation in visual, haptic and dual modality trials. The stimulus presentation and inter-stimulus interval were kept constant across all conditions.

and Ernst, 2007). The experimental procedure began once task instructions were understood for each condition and participants performed at 80% accuracy.

2.4. Analysis

The proportion of trials in which the participant reported the comparison was larger than the standard was calculated for each condition. Individual participant data was fitted by cumulative Gaussian functions using the Palamedes toolbox for MATLAB (Kingdom and Prins, 2009). Each psychometric function was then characterised by two parameters, μ and σ reflecting the mean and SD of the underlying Gaussian distribution. The first parameter (μ) provides a measure of the PSE, i.e., the comparison stimulus size that was perceived as the same size as the standard. In unimodal conditions this parameter should be unbiased (i.e., close to actual reference block size). However, in cue-conflict conditions the PSE should be biased towards whichever of the sensory information sources is up-weighted. The second parameter (σ) reflects the sensitivity of participant judgements to changes in comparison block size and is used as a measure of the reliability of the unisensory and dual modality estimates.

A measure of goodness of psychometric function fit, pDev, was estimated for each fit and these are included in Table 2. Larger values of pDev correspond to better psychometric function fits to the measured data. When $pDev < 0.05$ the function does not provide an acceptable fit of the data and is unlikely to

Table 2.

Psychometric function goodness of fit (mean pDev $\pm$ SD). Values closer to 1 represent better fits (see Kingdom and Prins, 2009)

	Visual	Haptic	Conflict 1	Conflict 2
ASC	0.51 $\pm$ 0.30	0.55 $\pm$ 0.32	0.47 $\pm$ 0.25	0.49 $\pm$ 0.26
NT	0.38 $\pm$ 0.28	0.52 $\pm$ 0.34	0.43 $\pm$ 0.23	0.53 $\pm$ 0.29

provide a reasonable estimate of the participants underlying sensory mechanisms (see Kingdom and Prins, 2009). Functions of three participants with ASC and three NTs were poorly fitted (pDev $<$ 0.05) and were refitted with the lapse rate left free to vary. Functions of one participant with ASC and one NT function were still poorly fitted and these participants were removed from analysis.

2.4.1. Analysis of SD Data

Data for each participant were compared with the predictions of the MLE model to assess optimal cue-integration. The average of the σ_{vh} parameters fitted in the conflict 1 and conflict 2 conditions was calculated to give the dual modality estimate $\sigma_{vh(avg)}$. Before calculating this average we conducted a paired sample t -test to confirm that there was no difference in SD between the conflict conditions, [$t(23) = 1.90$, $p = 0.071$, $d = 0.39$]. The fitted sensory variance parameters σ_v^2 and σ_h^2 were input into eqn. (5) to derive an MLE prediction $\hat{\sigma}_{vh}^2$. The square root of this value $\hat{\sigma}_{vh}$ was compared to $\sigma_{vh(avg)}$ using a mixed ANOVA. Significant interactions were followed up with t -tests. To compare the relative strength of evidence for the null hypothesis (the SD in the dual modality condition and the MLE prediction were drawn from the same distribution), and the alternative (the SD in the dual and MLE conditions were drawn from different distributions), Bayesian within-samples t -tests with default priors were conducted for each group using JASP (JASP Team, 2016). A Bayes factor (BF_{10}) $<$ 1 suggests increasing evidence for the null over the alternative hypothesis whereas a $BF_{10} >$ 1 suggests increasing evidence for the alternative over the null. A $BF_{10} >$ 3 is considered as strong evidence for the alternative hypothesis whereas a $BF_{10} <$ $\frac{1}{3}$ is considered as strong evidence for the null (Dienes, 2014). To reduce the influence of outliers on statistical analysis, data points which were more than 2.5 median absolute deviations from the condition median ($n = 3$, all in the NT group) were replaced with the maximum/minimum value $\pm$ 0.1 prior to analysis (Leys *et al.*, 2013).

2.4.2. Analysis of PSE Data

Every participant's performance in each of the conflict conditions was compared with an MLE prediction of the PSE in each condition from eqn. (5)

and the weights were calculated from the variance data according to eqn. (6). Participant PSEs were averaged and compared between the groups using independent samples *t*-test to explore any differences in shifts towards visual or haptic information in the conflict. Pearson’s correlation coefficients were calculated for each group to explore any relationships between the PSE measured in the dual modality conflict conditions and the MLE predictions of the PSE.

3. Results

Sample psychometric functions for one participant across the unisensory and dual modality conditions are provided in Fig. 4. Note that for this participant there is little bias in the unisensory conditions (PSE close to 0, as predicted). However, in the dual modality conflict conditions the PSE shifts (in this case to up-weight visual information). Note also that, consistent with the PSE data, the slopes of the unisensory data suggest that this participant has a more reliable sensory estimate of visual than haptic information.

3.1. SD Data

Mean standard deviation (σ) parameters are plotted in Fig. 5. If participants combined the dual modality inputs in a statistically optimal manner then

Figure 4. Sample psychometric functions for a single participant. (A) Data from the visual unisensory (red plus signs) and haptic unisensory (blue squares) conditions. (B) Data from conflict 1 (green upward triangles) and conflict 2 (magenta downward triangles). If the curve for conflict 1 shifts rightward and the curve for conflict 2 leftward, it suggests that the participant is up-weighting vision. The reverse pattern suggests that the participant is up-weighting haptics.

Figure 5. Standard Deviation (SD) measured in each condition and the prediction from the MLE model, the SCS model and the suboptimal model. The dual modality data was estimated from an average of the SD in conflict 1 and conflict 2. Participants with ASC are represented in bold blue font and NT participants in black (with crosshatch). The edges of each box represent the upper and lower quartile; the whiskers represents the most extreme decile. The line within each box denotes the condition median and the cross is the mean. Open dots represent extreme data points.

the SD measured in the dual modality condition will be comparable to the prediction from the MLE model (eqn. (5)). A multisensory performance enhancement for either group would be indicated by a reduced SD in the dual modality condition in comparison to the unisensory conditions. No multisensory performance enhancement was observed for either group and dual modality performance did not approximate the statistically optimal model. This was confirmed using a mixed ANOVA (Greenhouse–Geisser adjustment) with group as the between subject factor to compare the size of the SD between the conditions. A significant effect of condition ($F(2.07, 45.43) = 36.61$, $p < 0.001$, $\eta_p^2 = 0.625$), indicated the size of the SD differed between the conditions. This effect was followed up with paired sample t -tests (Bonferonni-corrected, $\alpha = 0.008$). As mentioned above, there were two critical tests: firstly if participants were combining the unisensory cues a multisensory performance enhancement would be indicated by a reduction in the SD in the

dual modality condition in comparison to either unisensory cues. As can be observed in Fig. 5, the dual modality SD was significantly larger than the visual SD [$t(23) = 4.39$, $p < 0.001$, $d = 0.90$]. There was a trend towards the dual modality SD being lower than the haptic SD [$t(23) = 1.99$, $p = 0.058$, $d = 0.41$]. Secondly, the critical test of optimal integration of these cues would be given by comparable SD between the MLE prediction and performance measured in the dual modality condition. However, the dual modality SD was significantly larger than the MLE prediction of SD [$t(23) = 7.68$, $p < 0.001$, $d = 1.57$]. The visual SD was significantly lower than the haptic SD [$t(23) = 5.70$, $p < 0.001$, $d = 1.16$], and as would be expected the MLE prediction of SD was significantly lower than the visual SD [$t(23) = 6.24$, $p < 0.001$, $d = 1.27$] and haptic SD [$t(23) = 10.13$, $p < 0.001$, $d = 2.07$]. There was a trend towards a main effect of group [$F(1, 22) = 4.21$, $p = 0.052$, $\eta_p^2 = 0.161$], indicating that the ASC group tended to have lower SD across the conditions compared to NT participants. Finally, there was no group $\times$ condition interaction [$F(2.07, 45.43) = 0.77$, $p = 0.518$, $\eta_p^2 = 0.003$], indicating that participants with ASC and NT participants produced similar SDs across conditions.

Bayes Factors (BF_{10}) were calculated based on the SD measured in the dual modality condition and the predictions of the MLE model. For participants with ASC, $BF_{10} = 87.69$ meaning that the alternative hypothesis (the SD in the dual modality condition and the MLE prediction were drawn from different distributions) was 87.69 times more likely than the null given the data. For NTs, $BF_{10} = 244.30$ meaning that the alternative hypothesis was 244.30 times more likely given the data.

3.2. PSE Data

PSEs in each of the dual modality conditions (conflict 1, conflict 2) are plotted as a function of the size of the conflict (see Fig. 6). The black dotted line represents the line of best fit. If the participant was up-weighting more reliable sensory information in a manner consistent with the MLE model then the PSEs measured in each conflict condition would fall closely to the MLE predictions of PSE as calculated from eqns (6) and (7) (given by the solid black line in Fig. 6). Alternatively, if the participant was relying entirely on either the visual or haptic cue in the conflict, then the measured data points would fall closely to a model of complete visual or haptic dominance (given by the red and blue dotted lines in Fig. 6 respectively).

Figure 7 provides a further summary of the PSE data. Each participant's PSEs were averaged across the conflict conditions after reversing the polarity of the data in the conflict 2 condition, so that more positive PSEs indicated more visual capture and more negative PSEs more haptic capture (PSE_{VH}).

Figure 6. PSEs plotted as a function of cue-conflict in each of the dual modality conditions for participants with ASC and NTs. Conflict 1 (visual block of 58 mm and a haptic block of 52 mm) is represented by a star and conflict 2 (visual block of 52 mm and a haptic block of 58 mm) by a circle. The dashed black line is the best linear fit of the measured data points for each participant. The solid black line is the MLE prediction for that participant, the red dashed line is the prediction for complete visual dominance and the bold dotted blue line for complete haptic dominance. If the fit of the measured data points falls close to that of the model predictions then the participant is performing in a manner which approximates that particular model.

Figure 7. Participants with ASC and NT average PSE across the dual modality conflict conditions as a function of the MLE predictions. The PSEs in the conflict 2 condition were reversed in polarity so that a positive value indicated that the participant was displaying more visual capture and a negative value more haptic capture. If the data points fall closely to the solid $x = y$ line then it suggests that they are performing in a manner well predicted by the MLE model. The dashed line indicates where data points would fall for complete visual dominance and the dotted line for complete haptic dominance.

In Fig. 7 these values are plotted as a function of the averaged MLE prediction of PSE in each conflict condition (PSE_{MLE}). If the data points fell close to the $x = y$ line then participants were performing in a manner consistent with the MLE model, while data points falling close to the dashed line were consistent with complete visual capture, and points close to the dotted line complete haptic capture. It can be observed from Fig. 7 that participants with ASC tended to approach visual capture while NT PSEs were more variable and included a number performing close to complete haptic capture. The average dual modality conflict PSEs were then compared between the groups using an independent samples t -test. Participants with ASC produced a larger PSE (mean = 1.08 ± 1.18) than NTs (mean = 0.28 ± 1.91), but this difference was not statistically significant [$t(22) = 1.22$, $p = 0.234$, $d = 0.50$]. The Bayes Factor (BF_{10}) was calculated comparing the averaged PSE for the ASC and NT groups. $BF_{10} = 0.64$ meaning the alternative hypothesis was 0.64 times more likely than the null (or that the null was 1.56 times as likely).

Correlation tests were then conducted to explore the extent to which either group's performance was consistent with the MLE model predictions. If performance was perfectly consistent with the MLE predictions then their data points would fall along the $x = y$ line in Fig. 7. A significant positive correlation between measured and predicted PSEs would indicate performance consistent with the MLE model for either group. However, Pearson's corre-

Table 3.

Pearson's r values for relationship between the difference in unisensory variabilities ($|\sigma_v^2 - \sigma_h^2|$) and the differences between dual modality and MLE predictions of variance ($|\sigma_{vh(avg)}^2 - \hat{\sigma}_{vh}^2|$) and PSE ($|PSE_{VH} - PSE_{MLE}|$). Coefficients were calculated for pooled data and when separated by group

	$ \sigma_v^2 - \sigma_h^2 $	n	r	p
ASC + NT	$ \sigma_{vh(avg)}^2 - \hat{\sigma}_{vh}^2 $	24	0.18	0.422
	$ PSE_{VH} - PSE_{MLE} $		0.11	0.622
ASC	$ \sigma_{vh(avg)}^2 - \hat{\sigma}_{vh}^2 $	12	−0.08	0.817
	$ PSE_{VH} - PSE_{MLE} $		−0.24	0.451
NT	$ \sigma_{vh(avg)}^2 - \hat{\sigma}_{vh}^2 $	12	0.14	0.655
	$ PSE_{VH} - PSE_{MLE} $		0.15	0.640

Table 4.

Pearson's r values for relationship between differences between dual modality and MLE predictions of variance ($|\sigma_{vh(avg)}^2 - \hat{\sigma}_{vh}^2|$) and PSE ($|PSE_{VH} - PSE_{MLE}|$). Mahalanobis distance revealed no significant multivariate outliers at χ^2 , $\alpha = 0.001$. The difference between the groups correlation coefficients were compared using the Fisher r to z transformation

	n	r	p	Z	p
	24	0.53	0.008		
ASC	12	0.41	0.182		
NT	12	0.55	0.065	0.39	0.697

lation tests revealed that there was no significant relationship for the ASC [$r(12) = -0.43$, $p = 0.163$] or NT group [$r(12) = -0.03$, $p = 0.926$].

In previous experiments performance is typically closer to MLE predictions when the unisensory weights are closer to being equal (i.e., when there is no strongly dominant cue; e.g., Ernst and Banks, 2002). To investigate this tendency in our data we calculated the correlation (using Pearson's r) between the difference in unisensory weights and the distance of the dual modality estimates from the MLE predictions. In other words we correlated $|\sigma_v^2 - \sigma_h^2|$ against $|\sigma_{vh(avg)}^2 - \hat{\sigma}_{vh}^2|$ and $|\sigma_v^2 - \sigma_h^2|$ against $|PSE_{VH} - PSE_{MLE}|$. Results from these analyses are summarised in Table 3, however, no significant correlations were found.

Finally, we correlated $|PSE_{VH} - PSE_{MLE}|$ against $|\sigma_{vh(avg)}^2 - \hat{\sigma}_{vh}^2|$ to explore whether the participants who produced the greatest multisensory performance enhancement also up-weighted the more reliable stimuli in the conflict conditions. Correlation coefficients are included in Table 4.

Figure 8. Proximity of participant performance measured in the dual modality conditions to predictions of the MLE model. That is, $|\sigma_{vh(avg)}^2 - \hat{\sigma}_{vh}^2|$ as a function of $|PSE_{vh} - PSE_{MLE}|$. ASC participants are represented by blue circles, NT black stars.

There was a significant positive correlation when the group data were pooled, indicating that participants with variance data closer to the MLE predictions also produced PSEs which were closer to the MLE predictions (Fig. 8).

3.3. Non-Optimal Visual–Haptic Processing

Neither group performed in a manner consistent with the statistically optimal MLE model of visual–haptic integration. In fact, neither group produced a multisensory performance enhancement. The SD data in the dual modality condition was not reduced in comparison to the unisensory conditions. We considered two further possible accounts of the data; cue-switching and sub-optimal integration of the cues.

The first alternative account rests upon the idea of stochastic cue-switching (SCS; Kuschel *et al.*, 2010; Landy and Kojima, 2001; Nardini *et al.*, 2008). This model predicts that the participant does not combine the stimuli when presented in conditions of conflict, but switches between them. The participant does not switch randomly, but is likely to consider the more reliable stimuli more frequently (Serwe and Trommershäuser, 2009). Under the SCS model, variance is not reduced compared with the most reliable unisensory estimate (Kuschel *et al.*, 2010) because additional variance may be added in trials in which the participant up-weights the less reliable cue. Furthermore, switching between cues which are presented in size conflict may lead to a further increase in variance relative to relying on only a single cue.

The SCS model predicts that the dual modality variance should be a weighted average based on the unisensory estimates plus a component due to the discrepancy between the cues:

$$\sigma_{vh}^2 = w_v \sigma_v^2 + w_h \sigma_h^2 + w_v w_h (\mu_v - \mu_h)^2 \quad (8)$$

where the weights relate to the probability of following either stimuli. The weights are calculated as the perturbation of the participant's PSE from a given cue in the conflict as a ratio of the overall conflict (see Landy *et al.*, 1995; Young *et al.*, 1993). For extracting the visual weight from conflict 1:

$$w_v = \frac{\text{cue}_1 - \mu_{\text{conflict 1}}}{\Delta \text{cue}} \quad (9)$$

where Δ_{cue} refers to the size of the conflict between the stimuli (in this instance 6 mm) and cue_1 refers to the physical properties of the haptic block in the conflict (52 mm in conflict 1). W_v therefore represents the extent to which the participant is following the visual stimuli when making size estimates in that conflict condition. For each participant, an average of the weights calculated from conflict 1 and conflict 2 were input into eqn. (8).

An alternative model under which weights were derived empirically was also considered. According to this model unisensory cues are integrated, but not weighted according to relative reliability. This model thus describes sub-optimal integration of the cues (Kuschel *et al.*, 2010). The unisensory cues are combined additively and weighted using the weights as calculated in eqn. (9) (which do not necessarily minimise variance) to give a prediction of dual modality variance according to eqn. (10).

$$\sigma_{vh}^2 = (w_v \sigma_v)^2 + (w_h \sigma_h)^2 \quad (10)$$

The dual modality SD was compared with the prediction derived from the SCS and suboptimal model (see Fig. 5).

Paired-sample *t*-tests (Bonferonni-corrected, $\alpha = 0.013$) and Bayes Factor (BF_{10}) were conducted to compare the SD predicted by the SCS model with the SD measured in the dual modality condition. The SCS prediction was lower than the dual modality SD, but not significantly so for participants with ASC [$t(11) = 0.06$, $p = 0.952$, $d = 0.02$], or for NTs [$t(11) = 0.20$, $p = 0.848$, $d = 0.06$]. For participants with ASC the comparison with the SCS model gave $\text{BF}_{10} = 0.29$, meaning the alternative was 0.29 times more likely than the null given the data (or alternatively that the null was 3.45 times more likely than the alternative). For NTs, the comparison between dual modality SD and the predictions of the SCS model $\text{BF}_{10} = 0.29$, meaning the alternative was 0.29 times more likely given the data (or that the null was 3.45 times more likely).

The suboptimal prediction was significantly lower than the dual modality SD for participants with ASC [$t(11) = 3.96$, $p = 0.002$, $d = 1.14$] and for NT

Table 5.

Measures of model fit, residual sum of squares, comparing measured dual modality SD with predictions derived from the MLE, SCS and suboptimal models. Models of visual and haptic capture were approximated by comparing performance with visual and haptic SD. Lower values indicate that the model was a closer fit to the dual modality data

Dual vs.	MLE	SCS	Suboptimal	Visual	Haptic
ASC	9.51	7.66	8.02	8.18	8.92
NT	22.85	10.65	12.32	35.69	15.32

participants [$t(11) = 4.12, p = 0.002, d = 1.19$]. For participants with ASC, the comparison with the suboptimal model gave $BF_{10} = 20.42$ meaning that the alternative was 20.42 times more likely than the alternative given the data. For NTs, the comparison between the dual and the suboptimal model $BF_{10} = 25.75$, meaning the alternative was 25.75 times more likely given the data.

To compare model fits (MLE, SCS, suboptimal integration) to the measured data the residual sum of squares (RSS) was calculated for comparisons of the predicted and measured SDs. Smaller values of RSS indicate that a model provided a closer fit to the measured data points. In addition, visual and haptic SD were also compared with dual modality performance as models of complete visual and haptic capture. RSS values are included in Table 5. For both groups the SCS model provided the best fit of the data.

Individual participant data for each group is presented in Fig. 9 with each participant’s MLE, SCS and suboptimal predictions plotted as a function of their dual modality SD. It can be seen that for both groups measured performance tended to fall closer to the SCS and suboptimal model predictions, although two participants with ASC and one NT also approached MLE performance.

4. Discussion

The aim of the present experiment was to investigate visual–haptic cue integration in ASC and compare this to NT participants. This was motivated primarily by two previous findings in the literature: firstly, NT participants have been shown consistently to integrate visual–haptic information in an optimal manner (Ernst and Banks, 2002; Gori *et al.*, 2008). Secondly, it has been proposed that individuals with ASC may not integrate reliability information in an optimal manner (Lawson *et al.*, 2014; Pellicano and Burr, 2012) and aspects of multisensory processing are disrupted in this group (Poole *et al.*, 2015; Russo *et al.*, 2010).

Figure 9. SD measured in the dual modality condition as a function of each participant's SD predicted by the MLE model, the SCS model and suboptimal model. Where data falls close to the $x = y$ line that participant is performing in a manner which is consistent with that given model.

To address these issues both ASC and NT participants completed a visual–haptic size judgement task based on the paradigm of Gori *et al.* (2008). MLE predictions of SD and PSE were derived from unisensory data for each participant. There was a non-significant trend towards reduced SD in the ASC group across the conditions. However, performance in the dual modality condition was not consistent with the MLE model for either ASC or NT participants. The dual modality SD was greater than the predictions of the MLE model for both groups and the MLE predictions of the PSE in each conflict condition were divergent from participants' measured performance.

The SD in the visual condition was significantly lower than in the haptic condition, which is consistent with previous studies which have investigated visual–haptic height judgements in NT adults (e.g., Ernst and Banks, 2002; Gori *et al.*, 2008). Interestingly, there was a trend towards greater reliability (reduced SD) in ASC participant size judgements in comparison to controls. There is little experimental research investigating haptic processing in ASC. In a previous study, adult participants with ASC performed comparably to controls when using haptics to make comparative length judgements about bars presented at different orientations (Nakano *et al.*, 2012). However, when asked to identify curvilinear and rectangular shapes using only haptic information which they had previously been shown visually, participants with ASC performed more accurately than controls. This suggests that aspects of visual–haptic processing may be superior in adults with ASC compared with NTs.

Multisensory performance enhancements were not observed for either group. The SD in the unisensory visual condition was significantly lower than the dual modality condition. A number of studies have suggested that individuals with ASC do not produce a performance enhancement on audio-visual tasks (Brandwein *et al.*, 2013; Collignon *et al.*, 2013). However, this finding was unexpected in NT participants. A multisensory performance enhancement is critical to the MLE model as each sense is combined so that the multisensory estimate is more reliable than either of the unisensory estimates. Accordingly, both ASC and NT group performance was not well predicted by the model. Additionally, participants in both groups did not up-weight the more reliable cue when making judgements in the dual modality condition as the PSEs deviated from the MLE predictions. However, there were individual differences in the tendency of participants to perform consistently with the MLE model. Indeed, there was a significant positive correlation between PSE and measures of how close the participants measured performance fell to the MLE predictions of reliability. Across the groups, participants who produced the greatest multisensory performance enhancement also tended to up-weight the more reliable stimuli in the conflict conditions, while those who deviated furthest from MLE predictions of reliability also tended to up-weight the less reliable cue.

Although a large body of research has indicated that performance on multisensory tasks is well predicted by the MLE model, there are studies which have failed to replicate the optimal integration of multisensory cues in NT adults. Adults and children aged 6 years old completed a visual–haptic task in which they judged the lengths of physical stimuli (Jovanovic and Drewing, 2014). This study presented participants with a visual, haptic or dual modality standard in comparison to unisensory haptic stimuli. Conflict was introduced in the dual modality condition by magnifying the visual stimuli. Performance in the combined condition was significantly larger than the predictions of the MLE, but smaller than the SCS prediction. Haptic alone performance was as good, or better, than performance in the combined condition in adults indicating no multisensory performance enhancement. However, as all comparison stimuli were unisensory (haptic), this task required potentially effortful visual-to-haptic matching in the dual modality condition which may eliminate the benefit of processing multiple cues. A virtual study of visual–haptic softness perception indicated that adult performance was not well predicted by the MLE model (Cellini *et al.*, 2013). Unisensory haptic judgements were less variable than visual. Participants produced a greater weighting of vision in the combined estimate than predicted, and performance deviated from the predictions of the MLE model. Moreover, where the stimuli were perceived as hard, haptic judgements were more reliable than multisensory judgements.

There may be a number of reasons why performance deviates from optimality (see Rahnev and Denison, 2016 for a review). For instance, the participant may have existing expectations about whether the cues should be combined, or the demands of a given task may encourage the participant to adopt an alternative strategy when processing the cues. A previous study of visual and proprioceptive cue-combination has indicated that the SCS model provided a better prediction of performance than the MLE model in NT adults (Serwe and Trommershäuser, 2009). Participants made judgements regarding the direction of moving radial lines presented visually and a proprioceptive cue which was presented using a haptic force feedback device. In the dual modality condition, the participant was instructed to focus on the proprioceptive cue while ignoring the visual stimulus. Participant SD data were closer to the predictions of the SCS model than that derived from the optimal MLE model. In the present study, the performance of participants in both groups was also comparable to the predictions of the SCS model. The possible role of the demands of the task in encouraging participants to adopt this strategy is discussed further below.

Although the present experiment closely replicates Gori *et al.* there are some methodological differences which may have contributed to the failure to replicate the optimal combination of visual–haptic cues in the NT group. Firstly, in order to reduce the testing time, the present study only included conflicting dual modality cues (with a relatively large conflict of 6 mm). Gori

et al. included a no-conflict condition in which the visual and haptic cues were consistent in size, together with a 4 mm conflict condition and a 6 mm conflict. Although a 6 mm conflict is not unusual in previous similar studies (Ernst and Banks, 2002; Gori *et al.*, 2008), it may be that omitting the cue-consistent conditions encouraged participants to become aware of the conflict. When a conflict is noticed, it is likely that the participant will adopt a conscious decision-making strategy when making judgements about the conflicting stimuli (De Gelder and Bertelson, 2003). This may have encouraged participants to switch between the visual and haptic cues from trial to trial in the dual modality conditions. In the online Supplementary Materials, we report an additional experiment in which a different group of NT adults completed a similar experiment including a dual modality condition with no-conflict. Although the MLE model had some predictive power in this additional experiment, SD in the dual modality condition was greater than the MLE predictions and the measured PSEs of five participants deviated from the MLE predictions of PSE. The SD in the dual modality condition was significantly lower than the predictions of the cue-switching model, but was indistinguishable from the suboptimal model predictions. As this additional experiment did not provide evidence for the MLE model, this suggests that the exclusion of the no-conflict condition alone is unlikely to explain the failure to fully replicate optimal cue integration in NT adults.

It is worth noting that none of the participants in the present study were experienced psychophysical observers. Many previous studies which have reported optimal cue integration between the senses have explored performance in experienced observers and participants who were not naïve to the design of the experiment (Ernst and Banks, 2002; Gepshtein *et al.*, 2005). The performance of non-experienced observers on 2IFC tasks can be more variable and may be influenced by non-sensory factors (Jäkel and Wichmann, 2006). For instance, given the demands of the psychophysical procedure, it may be that the performance in the dual modality condition was influenced by the decision making strategy selected by the participant. Indeed, the tendency of individuals to perform consistently with the MLE, SCS and suboptimal models may reflect the different strategies adopted by participants.

An additional consideration is the spatial discrepancy between the visual and haptic cues. In both the Gori *et al.* (2008) study and the present study, visual–haptic stimuli were presented either side of an occluding screen. The blocks used in the Gori *et al.* (2008) study were 60 mm deep meaning the visual stimuli were separated from where the participant explored the haptic stimuli by *at least* 60 mm. It has previously been observed that visual–haptic cue integration is progressively reduced with increased spatial disparity between the cues, with performance deviating from the MLE prediction to cues separated by comparable distances (Gepshtein *et al.*, 2005). It has been

noted elsewhere that the younger children in the Gori *et al.* study may have been adopting a cue-switching strategy (Nardini *et al.*, 2008). However, the adult and the older children's performance approximated the MLE predictions, which was not captured in the present study. Arguably this difference should lead to greater chances of observing integration in our study relative to the Gori study. Gori and colleagues also adjusted participant eye height so that the stimulus was always at eye level. While we did not control this in the present experiment, eye height was generally close to the level of the stimulus. It is possible that awareness of the depth of the visual stimulus may have alerted the participant to the separation between the stimuli. However since the block depth (10 mm) was considerably smaller than Gori *et al.* (2008) and viewing angle (the furthest point was 5° below eye level) was relatively small, it seems unlikely that this explains the observed differences in results. Considering this issue of spatial discrepancy between the cues, it is interesting that participants with ASC and NT adopted a similar strategy in the present study. In a previous study, visual-tactile interactions were observed at a greater spatial discrepancy in ASC (Poole *et al.*, 2015) which could reflect an extended representation of peripersonal space in ASC and suggest that participants with ASC would be more likely to integrate spatially discrepant cues.

In conclusion, it was not possible to draw strong inferences about the optimal combination of multisensory cues in ASC from the present study because optimal cue integration was not replicated in the adult NT group in either the SD or PSE data. This finding may be due to a number of important experimental factors (e.g., absence of no-conflict condition, use of non-naïve participants). We found that participants with ASC tended to produce more reliable height judgments than controls on this task. The performance of both groups was statistically comparable to a cue-switching model in which cues were not integrated on any trial. As differences in multisensory processing have previously been observed in ASC it is informative that both participants with ASC and NT controls adopt a similar strategy to processing visual–haptic cues in these task conditions.

References

- Alais, D. and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration, *Curr. Biol.* **14**, 257–262.
- American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th edn. American Psychiatric Publishing, Arlington, VA, USA.
- Anderson, J. S., Druzgal, T. J., Froehlich, A., Dubray, M. B., Lange, N., Alexander, A. L., Abildskov, T., Cariello, A. N., Cooperrider, J. R., Bigler, E. D. and Lainhart, J. E. (2011). Decreased interhemispheric functional connectivity in autism, *Cereb. Cortex* **21**, 1134–1146.
- Brandwein, A. B., Foxe, J. J., Butler, J. S., Russo, N. N., Altschuler, T. S., Gomes, H. and Molholm, S. (2013). The development of multisensory integration in high-functioning autism:

- high-density electrical mapping and psychophysical measures reveal impairments in the processing of audiovisual inputs, *Cereb. Cortex* **23**, 1329–1341.
- Brock, J., Brown, C. C., Boucher, J. and Rippon, G. (2002). The temporal binding deficit hypothesis of autism, *Dev. Psychopathol.* **14**, 209–224.
- Cascio, C. J., Foss-Feig, J. H., Burnette, C. P., Heacock, J. L. and Cosby, A. A. (2012). The rubber hand illusion in children with autism spectrum disorders: delayed influence of combined tactile and visual input on proprioception, *Autism* **16**, 406–419.
- Cellini, C., Kaim, L. and Drewing, K. (2013). Visual and haptic integration in the estimation of softness of deformable objects, *I-Perception* **4**, 516–531.
- Cherkassky, V. L., Kana, R. K., Keller, T. A. and Just, M. A. (2006). Functional connectivity in a baseline resting-state network in autism, *Neuroreport* **17**, 1687–1690.
- Collignon, O., Charbonneau, G., Peters, F., Nassim, M., Lassonde, M., Lepore, F., Mottron, L. and Bertone, A. (2013). Reduced multisensory facilitation in persons with autism, *Cortex* **49**, 1704–1710.
- Courchesne, E. (2004). Brain development in autism: early overgrowth followed by premature arrest of growth, *Ment. Retard. Dev. Disabil. Res. Rev.* **10**, 106–111.
- Courchesne, E., Carper, R. and Akshoomoff, N. (2003). Evidence of brain overgrowth in the first year of life in autism, *JAMA* **290**(3), 337–344.
- Crane, L., Goddard, L. and Pring, L. (2009). Sensory processing in adults with autism spectrum disorders, *Autism* **13**, 215–228.
- De Gelder, B. and Bertelson, P. (2003). Multisensory integration, perception and ecological validity, *Trends Cogn. Sci.* **7**, 460–467.
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results, *Front. Psychol.* **5**, 1–17. <http://doi.org/10.3389/fpsyg.2014.00781>.
- Dinstein, I., Heeger, D. J., Lorenzi, L., Minshew, N. J., Malach, R. and Behrmann, M. (2012). Unreliable evoked responses in autism, *Neuron* **75**, 981–991.
- Dinstein, I., Heeger, D. J. and Behrmann, M. (2015). Neural variability: friend or foe? *Trends Cogn. Sci.* **19**, 322–328.
- Ernst, M. O. (2008). Multisensory integration: a late bloomer? *Curr. Biol.* **18**, 519–521.
- Ernst, M. O. and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion, *Nature* **415**(6870), 429–433.
- Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L. and Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders, *Exp. Brain Res.* **203**, 381–389.
- Gepshtein, S. and Banks, M. S. (2003). Viewing geometry determines how vision and haptics combine in size perception, *Curr. Biol.* **13**, 483–488.
- Gepshtein, S., Burge, J., Ernst, M. O. and Banks, M. S. (2005). The combination of vision and touch depends on spatial proximity, *J. Vis.* **5**, 1013–1023.
- Gori, M., Del Viva, M., Sandini, G. and Burr, D. C. (2008). Young children do not integrate visual and haptic form information, *Curr. Biol.* **18**, 694–698.
- Grapel, J. N., Cicchetti, D. V. and Volkmar, F. R. (2015). Sensory features as diagnostic criteria for autism: sensory features in autism, *Yale J. Biol. Med.* **88**, 69–71.
- Haswell, C. C., Izawa, J., Dowell, L. R., Mostofsky, S. H. and Shadmehr, R. (2009). Representation of internal models of action in the autistic brain, *Nat. Neurosci.* **12**, 970–972.
- Helbig, H. B. and Ernst, M. O. (2007). Knowledge about a common source can promote visual–haptic integration, *Perception* **36**, 1523–1533.

- Irwin, J. R., Tornatore, L. A., Brancazio, L. and Whalen, D. H. (2011). Can children with autism spectrum disorders “hear” a speaking face? *Child Dev.* **82**, 1397–1403.
- Itahashi, T., Yamada, T., Watanabe, H., Nakamura, M., Ohta, H., Kanai, C., Iwanami, A., Kato, N. and Hashimoto, R. I. (2015). Alterations of local spontaneous brain activity and connectivity in adults with high-functioning autism spectrum disorder, *Mol. Autism* **6**, 30. <http://doi.org/10.1186/s13229-015-0026-z>.
- Jäkel, F. and Wichmann, F. A. (2006). Spatial four-alternative forced-choice method is the preferred psychophysical method for naïve observers, *J. Vis.* **6**, 1307–1322.
- JASP Team (2016). JASP (Version 0.7.5.5) [Computer software].
- Jovanovic, B. and Drewing, K. (2014). The influence of intersensory discrepancy on visuo-haptic integration is similar in 6-year-old children and adults, *Front. Psychol.* **5**, 57. <http://doi.org/10.3389/fpsyg.2014.00057>.
- Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K. and Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: evidence from an fmri study of an executive function task and corpus callosum morphometry, *Cereb. Cortex* **17**, 951–961.
- Kingdom, F. A. A. and Prins, N. (2009). *Psychophysics: a Practical Introduction*. Academic Press, London, UK.
- Kuschel, M., Di Luca, M., Buss, M. and Klatzky, R. L. (2010). Combination and integration in the perception of visual–haptic compliance information, *IEEE Transn Haptics* **3**, 234–244.
- Kwakye, L. D., Foss-Feig, J. H., Cascio, C. J., Stone, W. L. and Wallace, M. T. (2011). Altered auditory and multisensory temporal processing in autism spectrum disorders, *Front. Integr. Neurosci.* **4**, 129. <http://doi.org/10.3389/fnint.2010.00129>.
- Landy, M. S. and Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges, *J. Opt. Soc. Am.* **18**, 2307–2322.
- Landy, M., Maloney, L., Johnston, E. and Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion, *Vis. Res.* **35**, 389–412.
- Lawson, R. P., Rees, G. and Friston, K. J. (2014). An aberrant precision account of autism, *Front. Hum. Neurosci.* **8**, 302. <http://doi.org/10.3389/fnhum.2014.00302>.
- Leekam, S. R., Nieto, C., Libby, S. J., Wing, L. and Gould, J. (2007). Describing the sensory abnormalities of children and adults with autism, *J. Autism Dev. Disord.* **37**, 894–910.
- Leys, C., Ley, C., Klein, O., Bernard, P. and Licata, L. (2013). Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median, *J. Exp. Soc. Psychol.* **49**, 764–766.
- Little, L. M., Ausderau, K., Sideris, J. and Baranek, G. T. (2015). Activity participation and sensory features among children with autism spectrum disorders, *J. Autism Dev. Disord.* **45**, 2981–2990.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A. and Rutter, M. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism, *J. Autism Dev. Disord.* **30**, 205–223.
- Mongillo, E. A., Irwin, J. R., Whalen, D. H., Klaiman, C., Carter, A. S. and Schultz, R. T. (2008). Audiovisual processing in children with and without autism spectrum disorders, *J. Autism Dev. Disord.* **38**, 1349–1358.
- Nakano, T., Kato, N. and Kitazawa, S. (2012). Superior haptic-to-visual shape matching in autism spectrum disorders, *Neuropsychologia* **50**, 696–703.

- Nardini, M., Jones, P., Bedford, R. and Braddick, O. (2008). Development of cue integration in human navigation, *Curr. Biol.* **18**, 689–693.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory, *Neuropsychologia* **9**, 97–113.
- Pellicano, E. (2013). Sensory symptoms in autism: a blooming, buzzing confusion? *Child Dev. Perspect.* **7**, 143–148.
- Pellicano, E. and Burr, D. (2012). When the world becomes “too real”: a Bayesian explanation of autistic perception, *Trends Cogn. Sci.* **16**, 504–510.
- Poole, D., Gowen, E., Warren, P. A. and Poliakoff, E. (2015). Investigating visual–tactile interactions over time and space in adults with autism, *J. Autism Dev. Disord.* **45**, 3316–3326.
- Rahnev, D. and Denison, D. (2016). Suboptimality in perception, *bioRxiv*. <http://doi.org/10.1101/060194>.
- Robertson, A. E. and Simmons, D. R. (2013). The relationship between sensory sensitivity and autistic traits in the general population, *J. Autism Dev. Disord.* **43**, 775–784.
- Rohde, M., Van Dam, L. C. J. and Ernst, M. (2015). Statistically optimal multisensory cue integration: a practical tutorial, *Multisens. Res.* **29**, 1–39.
- Rubenstein, J. L. R. and Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems, *Genes Brain Behav.* **2**, 255–267.
- Russo, N., Foxe, J. J., Brandwein, A. B., Altschuler, T., Gomes, H. and Molholm, S. (2010). Multisensory processing in children with autism: high-density electrical mapping of auditory–somatosensory integration, *Autism Res.* **3**, 253–267.
- Serwe, S. and Trommershäuser, J. (2009). Combination of noisy directional visual and proprioceptive information, *J. Vis.* **9**, 1–14.
- Stevenson, R. A., Siemann, J. K., Woynaroski, T. G., Schneider, B. C., Eberly, H. E., Camarata, S. M. and Wallace, M. T. (2014). Evidence for diminished multisensory integration in autism spectrum disorder, *J. Autism Dev. Disord.* **44**, 3161–3167.
- Takahashi, C., Diedrichsen, J. and Watt, S. J. (2009). Integration of vision and haptics during tool use, *J. Vis.* **9**, 1–13.
- Tomchek, S. D. and Dunn, W. (2007). Sensory processing in children with and without autism: a comparative study using the short sensory profile, *Am. J. Occup. Ther.* **61**, 190–200.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. The Psychological Corporation, San Antonio, TX, USA.
- Young, M. J., Landy, M. S. and Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues, *Vis. Res.* **33**, 2685–2696.
- Zaidel, A., Goin-Kochel, R. P. and Angelaki, D. E. (2015). Self-motion perception in autism is compromised by visual noise but integrated optimally across multiple senses, *Proc. Natl Acad. Sci. USA* **112**, 6461–6466.