Developmental Recovery of Impaired Multisensory Processing in Autism and the Cost of Switching Sensory Modality

Michael J. **Crosse**[1,2,*], John J. **Foxe**[1,2,3] and Sophie **Molholm**[1,2,3]

[1]The Cognitive Neurophysiology Laboratory, Department of Pediatrics, Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, NY 10461, USA
[2]The Dominick P. Purpura Department of Neuroscience, Rose F. Kennedy Intellectual and Developmental Disabilities Research Center, Albert Einstein College of Medicine, Bronx, NY 10461, USA
[3]The Cognitive Neurophysiology Laboratory, Department of Neuroscience, The Ernest J. Del Monte Institute for Neuroscience, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642, USA

Abstract

Children with autism spectrum disorder (ASD) are often impaired in their ability to cope with and process multisensory information, which may contribute to some of the social and communicative deficits that are prevalent in this population. Amelioration of such deficits in adolescence has been observed for ecologically-relevant stimuli such as speech. However, it is not yet known if this recovery generalizes to the processing of nonsocial stimuli such as more basic beeps and flashes, typically used in cognitive neuroscience research. We hypothesize that engagement of different neural processes and lack of environmental exposure to such artificial stimuli leads to protracted developmental trajectories in both neurotypical (NT) individuals and individuals with ASD, thus delaying the age at which we observe this "catch up". Here, we test this hypothesis using a bisensory detection task by measuring human response times to randomly presented auditory, visual and audiovisual stimuli. By measuring the behavioral gain afforded by an audiovisual signal, we show that the multisensory deficit previously reported in children with ASD recovers in adulthood by the mid-twenties. In addition, we examine the effects of switching between sensory modalities and show that teenagers with ASD incur less of a behavioral cost than their NT peers. Computational modelling reveals that multisensory information interacts according to different rules in children and adults, and that sensory evidence is weighted differently too. In ASD, weighting of sensory information and allocation of attention during multisensory processing differs to that of NT individuals. Based on our findings, we propose a theoretical framework of multisensory development in NT and ASD individuals.

Keywords: bisensory detection, audiovisual integration, reaction time, redundant signals effect, race model, neurodevelopmental disorder, ASD.

**Corresponding author*
Email addresses: michael.crosse@einstein.yu.edu (MJC), john_foxe@urmc.rochester.edu (JJF), sophie.molholm@einstein.yu.edu (SM)

Introduction

Biological events tend to be multisensory, emanating or reflecting multiple forms of energy (e.g. photons, airborne vibrations, volatilized molecules, etc.). Humans have evolved a highly-specialized set of sensory receptors that enable us to sample these different forms of energy concurrently, optimizing how we perceive ecologically-relevant information. For instance, processing redundant multisensory signals often leads to faster reaction times (RTs) than processing the same information separately, a phenomenon known as the *redundant signals effect* (RSE; Todd, 1912; Hershenson, 1962; Kinchla, 1974). While a *race model* account of the RSE predicts that a response is triggered independently by the faster modality (Raab, 1962), the RSE typically exceeds the benefit predicted by statistical facilitation (Miller, 1982). Violation of the race model has been demonstrated using bisensory detection tasks for several decades and is widely interpreted as reflecting the multisensory gain due to pooled or integrated information processing (Gielen et al., 1983; Miller, 1986; Diederich and Colonius, 1987; Harrington and Peck, 1998; Molholm et al., 2002; Murray et al., 2004; Mégevand et al., 2013; Mahoney et al., 2015).

Whereas multisensory processing clearly influences how we perceive most biological events, particularly in instances when sensory evidence is ambiguous (Sumby and Pollack, 1954; Ross et al., 2007; Crosse et al., 2016), individuals with autism spectrum disorder (ASD) often do not benefit from the availability of multisensory information to the same extent as their neurotypical (NT) peers (de Gelder et al., 1991; Smith and Bennetto, 2007; Silverman et al., 2010; Irwin et al., 2011; Bebko et al., 2014; Foxe et al., 2015). We and others have suggested that impaired multisensory processing in ASD contributes to some of the commonly associated phenotypes such as atypical responses to sensory stimulation, and may even have detrimental effects on higher-order processes such as social interaction and communication (Ayres and Tickle, 1980; Martineau et al., 1992; Iarocci and McDonald, 2006; Foxe and Molholm, 2009; Beker et al., 2017; Stevenson et al., 2017).

In previous work by our lab, we demonstrated that multisensory gain increases steadily over the course of development for both speeded audiovisual (AV) detection (Brandwein et al., 2011) and AV speech identification (Ross et al., 2011). This is supported by animal neurophysiology that has shown that the ability to integrate multiple sensory inputs emerges with exposure to multisensory experiences (Wallace et al., 2004; Wallace and Stein, 2007; Stein et al., 2014). Whereas multisensory processing was significantly impaired in children with ASD for both of these tasks (Brandwein et al., 2013; Foxe et al., 2015), we and others have shown that neurotypical levels of AV speech integration are achieved by the time that individuals with ASD reach adolescence (Taylor et al., 2010; Foxe et al., 2015). In contrast, high-functioning teenagers with ASD failed to show reliable multisensory gain when performing a simple AV detection task (Brandwein et al., 2013). Recent theoretical (Beker et al., 2017) and computational (Cuppini et al., 2017) perspectives have suggested that the constant exposure to AV speech during maturation may serve to train multisensory speech function, leading to earlier developmental recovery of function in ASD. In support of this, the trajectory of multisensory development in typically-developing individuals reaches full maturity much earlier for speech stimuli (Ross et al., 2011) compared to non-speech stimuli (Brandwein et al., 2011). Here, using the same AV detection task, we tested the hypothesis that recovery of multisensory function in ASD occurs at a later developmental stage for nonsocial stimuli.

When switching from one sensory modality to another, average response times are slower on trials preceded by a different sensory modality (switch trials) compared to trials preceded by the same modality (repeat trials; Wundt, 1893; Sutton et al., 1961; Spence et al., 2001). Modality switch effects (MSEs) are inherent to any bisensory detection task that uses an intermixed stimulus presentation design (Gondan and Minakata, 2016; Otto and Mamassian, 2017) and have been shown to systematically contribute to the RSE because they are typically larger on unisensory trials than on multisensory trials (Gondan et al., 2004; Van der Stoep et al., 2015a; Shaw et al., 2019). Moreover, data suggest that children with high-functioning ASD incur a greater cost when switching from auditory to visual stimuli than their NT peers

(Williams et al., 2013). We therefore considered group differences in MSEs and quantified their contribution to the RSE. Using a computational modelling framework (Otto and Mamassian, 2012), we investigated how attentional resources were spread across sensory channels during speeded bisensory detection, and considered how this in turn could impact MSEs. We discuss the implications of MSEs on the interpretation of the RSE, and how the interplay between multisensory integration and switch effects may contribute differentially over the course of development in NT and ASD individuals.

Recent studies have demonstrated that multisensory behavior in NT adults can be explained by the basic cognitive architecture of the race model (Otto and Mamassian, 2012; Otto et al., 2013; Innes and Otto, 2019). However, it is not yet known if the same cognitive architecture applies to multisensory processing in children and individuals with ASD, or whether they employ an alternative processing strategy and integrate multisensory inputs according to different rules. To test this, we examined whether the race model framework could be used to predict empirical multisensory benefits in each group. This modelling approach was also used to quantify developmental changes in sensory dominance. Based on our empirical findings and computational analysis, we propose a theoretical framework to explain the maturational patterns of multisensory processing in NT and ASD individuals.

Methods

The present study is based on new analyses of a large body of data collected as part of several previously published studies (Brandwein et al., 2011; Brandwein et al., 2013; Brandwein et al., 2015), as well as new unpublished data.

Participants

A total of 400 individuals participated in the experiment. The data of 42 participants (10.5% of the total sample, 29 ASD) were excluded from all analyses based on the following criteria: 1) they did not fall within the desired age range of 6–40 years, 2) their performance IQ was below 80, 3) their detection accuracy

was less than 3 SDs below the sample's mean, 4) they had an excessive number of false alarms, 5) they had a disproportionate number of misses on visual trials (excessive eye-closure) or on audio trials (not listening), or 6) they had less than 20 RTs per condition (this can bias median RT estimates (Miller, 1988, 1991) as well as race model estimates (Kiesel et al., 2007)). Of the remaining 358 participants, 225 met criteria for NT (age range: 6–36 years; 115 females) and 133 had a diagnosis of ASD (age range: 6–39 years; 34 females). For analysis purposes, age was either treated as a continuous variable or participants were cross-sectioned into four developmental subgroups: children (6–9 years), pre-adolescents (10–12 years), adolescents (13–17 years), adults (18–40 years). Mean age was not statistically different between NT and ASD participants in any of the four age groups ($t < 0.91$, $p > 0.38$, $d < 0.21$). Participant demographics are presented in Table 1.

Individuals were excluded from participating in the experiment if they had a history of seizures or head trauma, or a known genetic disorder. Additionally, NT participants were excluded if they had a history of psychiatric, educational, attentional or other developmental difficulties (as assessed by a history questionnaire), a biological first-degree relative with a known developmental disorder, or if they or their legal guardians endorsed six or more items of inattention or hyperactivity on a DSM-IV checklist for attention deficit disorder. For the vast majority of participants, diagnoses of ASD were obtained by a trained clinical psychologist using the Autism Diagnostic Interview-Revised (Lord et al., 1994) and the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000). Diagnoses of the remaining individuals were made by a licensed clinical psychologist external to this study using the Diagnostic Criteria for Autistic Disorder from the DSM-IV TR (APA, 2000). For more details regarding sub-phenotyping, medication and ethnic demographics, please refer to Brandwein et al. (2013) and Brandwein et al. (2015).

IQ quotients for performance (PIQ), verbal (VIQ) and full-scale (FSIQ) intelligence were assessed in the majority of participants using the Wechsler Abbreviated Scales of Intelligence (WASI; Stano, 1999). Note

that mean PIQ was not statistically different between NT and ASD participants in any of the four age groups ($t < 1.2$, $p > 0.23$, $d < 0.27$). The descriptive statistics for each of the subgroups are summarized in Table 1. Participants were formally screened for normal or corrected-to-normal vision using a Snellen eye test chart and audiometric threshold evaluation confirmed that all participants had within-normal-limits hearing. All procedures were approved by the institutional review boards of the City College of New York and the Albert Einstein College of Medicine. All participants or legal guardians of participants provided written informed consent in accordance with the tenets of the 1964 Declaration of Helsinki.

Table 1. Demographic characteristics of participant populations.

	NT				ASD			
	6–9 yrs	10–12 yrs	13–17 yrs	18–40 yrs	6–9 yrs	10–12 yrs	13–17 yrs	18–40 yrs
n	51	46	54	74	44	34	31	23
n_{female}	27	26	24	38	7	6	10	11
n_{IQ}	45	43	48	10	44	33	30	21
Age	8.1 (1.2)	11.5 (1.0)	15.0 (1.3)	25.3 (3.6)	8.1 (1.0)	11.4 (0.7)	14.7 (1.6)	25.8 (5.1)
F_1 score	0.90 (0.07)	0.93 (0.06)	0.95 (0.04)	0.97 (0.02)	0.85 (0.08)	0.87 (0.08)	0.92 (0.07)	0.95 (0.04)
PIQ	106.1 (13.0)	109.7 (10.7)	104.9 (13.3)	109.9 (12.3)	106.2 (17.1)	106.6 (16.2)	107.9 (13.2)	108.5 (13.9)
VIQ	113.0 (10.6)	111.8 (13.0)	113.1 (12.8)	115.1 (16.1)	97.3 (19.8)	99.4 (19.1)	99.9 (18.8)	109.3 (15.8)
FSIQ	111.4 (11.5)	112.2 (11.7)	110.1 (12.5)	114.5 (14.0)	101.7 (17.5)	102.8 (17.4)	104.1 (14.1)	110.0 (14.4)
ADOS	–	–	–	–	7.3 (2.3)	8.0 (0.9)	6.9 (3.3)	–

Note: n_{female} indicates the number of female participants in respective age groups and n_{IQ} indicates the number of participants for whom IQ scores were obtained. The number of participants for whom ADOS scores were obtained is 31, 19, 7 respectively. PIQ: performance IQ; VIQ: verbal IQ; FSIQ: full-scale IQ (assessed using the WASI). F_1 scores indicate participants' detection accuracy, accounting for false alarms (see Methods for details). Values indicate the group mean with standard deviation shown in parentheses.

Stimuli and procedure

The stimulus materials were identical to those described in Brandwein et al. (2011). In brief, visual (V) stimuli consisted of a red disc (diameter: 3.2 cm; duration: 60 ms), located 0.4 cm above a central fixation

crosshair on a black background. The disc subtended visual angles of 1.5° vertically and horizontally and the bottom of the disc subtended 0.9° vertically above the crosshair (Fig. 1A). Auditory (A) stimuli consisted of a 1-kHz pure tone, sampled at 44.1 kHz (duration: 60 ms; rise/fall time: 5 ms). Audiovisual (AV) stimuli consisted of the combined simultaneous pairing of the auditory and visual stimuli described above.

Participants performed a speeded bisensory detection task on a computer and were seated 122 cm from the visual display in a dimly-lit, sound-attenuated booth. RTs were recorded during the simultaneous recording of electrophysiological (EEG) data, however, the EEG data are not reported in this study (for an account of previous EEG analyses, please refer to Brandwein et al., 2011; Brandwein et al., 2013; Brandwein et al., 2015). To reduce predictability, the stimuli were presented in a completely randomized order with equal probability and the interstimulus interval (ISI) was randomly jittered between 1000–3000 ms according to a uniform, square-wave distribution (see Fig. 1A). Stimulus presentation was controlled using Presentation® software (Neurobehavioral Systems, Inc., Berkeley, CA). Auditory stimuli were delivered binaurally at an intensity of 75 dB SPL via a single, centrally-located loudspeaker (JBL Duet Speaker System, Harman Multimedia). Visual stimuli were presented at a resolution of 1280 × 1024 pixels on a 17-inch Flat Panel LCD monitor (Dell Ultrasharp 1704FTP). The auditory and visual stimuli were presented in close spatial proximity, with the speaker placed atop the monitor and aligned vertically to the visual stimulus. Participants were instructed to press a button on a response pad (Logitech Wingman Precision Gamepad) with their right thumb as soon as they perceived any of the three stimuli. Analogue triggers indicating the latencies of stimulus onsets and button presses were sent to the acquisition PC via Presentation® and stored digitally at a sampling rate of 512 Hz in a separate channel of the EEG data file using ActiView software (BioSemi™, Amsterdam, The Netherlands). Stimuli were presented in blocks of ~100 trials and participants typically completed 6–10 blocks in total.

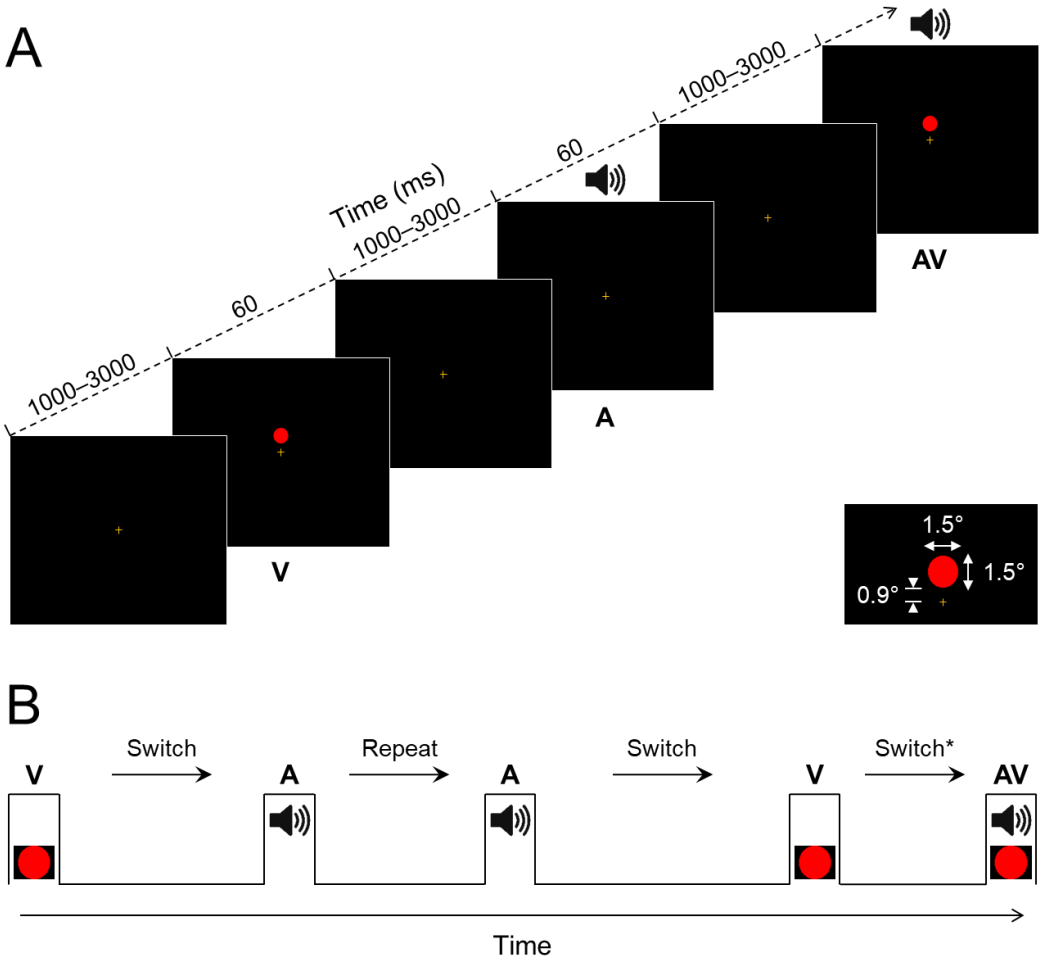


Figure 1. Bisensory detection task. **A**, Auditory (A), visual (V) and audiovisual (AV) stimuli (60-ms duration) were presented in a randomized order every 1000–3000 ms. Participants responded to each stimulus with a button press as fast as possible. **B**, Stimuli were categorized as either switch or repeat trials based on the modality of the preceding stimulus (repeat trials: AV→AV, A→A, V→V; switch trials: V→AV*, A→AV*, A→V, V→A). Asterisks indicate trials that are only partial switches. Trials AV→A and AV→V were excluded from the analysis as they were considered neither switches nor repeats.

Data analysis

Detection accuracy was assessed in order to identify participants that did not attend adequately to the stimuli. To account for false alarms and excessive button pressing, F_1 scores were computed as the harmonic mean of precision and recall (Van Rijsbergen, 1979):

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (1)$$

where precision = hits/(hits + false alarms) and recall = hits/(hits + misses). Responses were considered as false alarms if they occurred earlier than 100 ms post stimulus onset, or if they occurred after another response but before the next stimulus. Responses were considered as misses if they occurred later than 2000 ms post stimulus onset, or if there was no response at all to a given stimulus.

Response times were measured relative to the onset time of the preceding stimulus and analyzed separately for each participant in MATLAB (The MathWorks, Inc., Natick, MA). Responses were excluded from all analyses if there was more than one response within a given trial (double-presses), they occurred within the first 3 trials of a block (considered training) or the preceding ISI was not between 1000–3000 ms (due to system errors). An outlier correction procedure was performed before the main RT analyses. First, RTs that did not fall within 100–2000 ms post-stimulus were removed. On average, fast outliers (<100 ms, considered anticipatory responses) made up 0.7% (±0.9) of trials and slow outliers (>2000 ms, considered misses) made up 0.4% (±0.6) of trials. Second, RTs outside the middle 95th percentile (2.5–97.5) of their respective conditions were removed. While not all RTs outside of this range are necessarily outliers, those within this range are most likely to come from the cognitive processes under consideration (Ratcliff, 1993). This approach minimizes the impact of outliers with only negligible truncation biases (Ulrich and Miller, 1994) and captures the range of RTs at an individual-participant level, an important factor when dealing with significant inter-subject variability.

Analysis of RT data was conducted on the whole RT distribution by splitting it into discrete quantiles (Ratcliff, 1979). RTs were organized into 20 linearly-spaced quantiles between the 2.5–97.5 cutoffs used for outlier correction. Because outlier correction was performed separately for each condition, the lowest 2.5 and highest 97.5 percentiles were used for all three conditions in order to maintain the relationship between them. Cumulative distribution functions (CDFs) were obtained by calculating the cumulative

probability of RTs occurring below time t given a signal X, $P(RT_X \leq t|X)$. CDFs were averaged or "Vincentized" across participants at each corresponding quantile (Vincent, 1912). Note, this approach does not require there to be an equal number of RTs in each condition (Ulrich et al., 2007).

Race model analysis

To obtain quantitative predictions of statistical facilitation, we used Raab's race model (Raab, 1962). Race models predict that the response to a redundant signal is triggered by the faster of the two sensory modalities. Let $P(RT_A \leq t|AV)$ and $P(RT_V \leq t|AV)$ represent the CDFs of the A and V components of an AV stimulus, respectively. Assuming the RT distributions of the A and V components overlap, the probability of either triggering a response can be represented using probability summation:

$$P(\mathrm{RT}_{\mathrm{A}\cup\mathrm{V}} \leq t|\mathrm{AV}) = P(\mathrm{RT}_{\mathrm{A}} \leq t|\mathrm{AV}) + P(\mathrm{RT}_{\mathrm{V}} \leq t|\mathrm{AV}) - P(\mathrm{RT}_{\mathrm{A}\cap\mathrm{V}} \leq t|\mathrm{AV}) \quad (2)$$

where $P(RT_{A\cap V} \leq t|AV)$ is the probability of the A and V signals triggering a response at the same time. To solve this analytically, we need to make two assumptions: 1) RTs to the A and V components of the AV signal follow the same distributions as the RTs to the unisensory A and V signals, such that $P(RT_A \leq t|AV) = P(RT_A \leq t|A)$ and $P(RT_V \leq t|AV) = P(RT_V \leq t|V)$, an assumption known as context invariance (Ashby and Townsend, 1986; Luce, 1986; Miller, 2016); 2) RTs to the A and V components of the AV signal are statistically independent, such that their joint probability $P(RT_{A\cap V} \leq t|AV)$ can be calculated by the product of $P(RT_A \leq t|AV)$ and $P(RT_V \leq t|AV)$ (Meijers and Eijkman, 1977). Simplifying $P(RT_{AUV} \leq t|AV)$ to $F_{AUV}(t)$, $P(RT_A \leq t|A)$ to $F_A(t)$ and $P(RT_V \leq t|V)$ to $F_V(t)$, equation 2 can be represented as:

$$F_{\mathrm{A}\cup\mathrm{V}}(t) = F_{\mathrm{A}}(t) + F_{\mathrm{V}}(t) - F_{\mathrm{A}}(t) \times F_{\mathrm{V}}(t) \quad (3)$$

Note, the joint probability term is often omitted from equation 3 to produce an upper bound known as Miller's bound or the race model inequality (Miller, 1982), as the assumption of statistical independence is poorly motivated; it is likely that responses to signals on different channels compete for resources

(Miller, 1978, 1982; Colonius, 1986, 1990; Gondan and Minakata, 2016). Assuming that the allocation of attentional resources to each channel is partially determined by the modality of the previous trial (Miller, 1982), we separated the unisensory RTs by their preceding sensory modality and computed individual race models before averaging across them:

$$\bar{F}_{\mathrm{A\cup V}}(t) = \frac{1}{3}\sum_{m=1}^{3} F_{\mathrm{A\cup V}}(m,t) \tag{4}$$

where *m* is the preceding modality. This approach captured some of the dependency between RTs to signals on different channels, resulting in an estimate of statistical facilitation that was less conservative at every quantile ($p < 0.025$, two-tailed permutation tests). Note that using Raab's model or Miller's bound typically yields the same outcome qualitatively (Van der Stoep et al., 2015b; Van der Stoep et al., 2015a).

Multisensory benefits were quantified by the area between the CDFs in the multisensory condition and the most effective unisensory condition (Otto et al., 2013). First, we computed the multisensory benefit predicted by the race model (Fig. 2B, left):

$$\mathrm{benefit}_{\mathrm{pred}} = \int_0^1 \bar{F}_{\mathrm{A\cup V}}(t) - \max[F_{\mathrm{A}}(t), F_{\mathrm{V}}(t)]\, dt \tag{5}$$

where the integral is taken over every quantile *t* from 0 to 1. The term $\max[F_A(t), F_V(t)]$ represents a lower bound of facilitation, known as Grice's bound (Grice et al., 1984), whereby no statistical benefit is observed for a redundant signal at any quantile. Similarly, we computed empirical benefits based on the actual multisensory RTs (Fig. 2B, right):

$$\mathrm{benefit}_{\mathrm{emp}} = \int_0^1 F_{\mathrm{AV}}(t) - \max[F_{\mathrm{A}}(t), F_{\mathrm{V}}(t)]\, dt \tag{6}$$

Note that this is not the same as measuring multisensory interactions since Grice's bound does not account for statistical facilitation (see Innes and Otto, 2019). Rather, it quantifies the benefit afforded by a redundant signal relative to that of the most effective unisensory signal.

To determine whether the empirical multisensory benefits exceeded statistical facilitation, we computed the difference between the CDFs of the multisensory condition and the race model at every quantile (Molholm et al., 2002). Positive values indicate quantiles where multisensory RTs were faster than predicted, i.e., violation of the race model. To obtain an overall index of multisensory gain, we calculated the area under the curve (AUC) by taking the integral over every quantile as before (Fig. 4A):

$$\text{gain} = \int_0^1 F_{\text{AV}}(t) - \bar{F}_{\text{AUV}}(t)\, dt \tag{7}$$

While it is common practice to interpret the AUC above zero as an index of multisensory interactions (Miller, 1986; Nozawa et al., 1994; Hughes et al., 1998), equation 6 is equal to the AUC above zero minus the AUC below zero (Colonius and Diederich, 2006; Krueger Fister et al., 2016). This is mathematically equivalent to the difference between predicted benefits (Eq. 5) and empirical benefits (Eq. 6) and represents the overall behavioral gain across the participant's entire RT distribution. Qualitatively, this is equivalent to using only the positive portion (e.g., Nidiffer et al., 2016), because the AUC below the x-axis is negatively correlated with the AUC above the x-axis (see Fig. 4B). Moreover, the majority of younger participants in this study did not exceed statistical facilitation, rendering a statistical analysis based on the positive AUC less powerful. All race model analyses were conducted using the RaceModel open-source toolbox (https://github.com/mickcrosse/RaceModel).

Modality switch effects

When testing the race model, randomly interleaving sensory modalities is necessary to minimize the opportunity for different processing strategies to be deployed under unisensory and multisensory

conditions and hence satisfy the assumption of context invariance (Gondan and Minakata, 2016; Miller, 2016; Otto and Mamassian, 2017). While modality switch effects are inherent to such task conditions, their size and contribution to processes such as the RSE are rarely if ever quantified. It has been suggested that reporting the size of MSEs should become a routine procedure in RSE studies and that failure to do so would render such studies incomplete (see Otto and Mamassian, 2017). Accordingly, we assessed MSEs in NT and ASD individuals and whether or not they were likely to account for the observed RSE.

To examine MSEs, RTs were separated into those preceded by the same modality (repeat trials) and those preceded by a different modality (switch trials). Unisensory trials preceded by multisensory trials (AV→A, AV→V) were excluded from this analysis as they were considered neither switches nor repeats (repeat trials: A→A, V→V, AV→AV; switch trials: V→A, A→V, V→AV, A→AV). Separate CDFs were obtained for switch and repeat trials within each condition. The CDFs of the two multisensory switch conditions (A→AV, V→AV) were averaged to produce one multisensory switch condition (V/A→AV). MSEs were quantified by the area between the CDFs of the switch and repeat trials:

$$\mathrm{MSE} = \int_0^1 F_{\mathrm{repeat}}(t) - F_{\mathrm{switch}}(t)\, dt \tag{8}$$

To examine the impact of switching sensory modality on the observed multisensory gain, separate tests of the race model were performed for switch and repeat trials.

Modelling channel dependency and RT variability

It is widely considered that violation of the race model necessitates the rejection of its basic architecture in favor of the so-called coactivation model, whereby multisensory activity is pooled or integrated prior to the formation of a decision (Miller, 1982). Alternatively, sensory evidence could accumulate along separate channels that interact with one another, forming separate decisions that are then coupled by a task-relevant logical operation (Fig. 10; Mordkoff and Yantis, 1991; Townsend and Wenger, 2004; Otto

and Mamassian, 2017). Seminal work by Otto and Mamassian (2012) demonstrated that the basic race architecture can be used to explain empirical multisensory RT data by including two additional parameters to account for the additional variability or noise η, typically observed in empirical multisensory RTs compared to that predicted by probability summation, and the correlation ρ between RTs to signals on different sensory channels. Figure 6A illustrates the effect of trial history on the correlation between RTs on different channels as a function of RT quantile. Conceptually, Miller's and Grice's bounds assume a perfect negative and positive correlation respectively, whereas Raab's model assumes zero correlation (i.e., independence). Otto's context variant race model on the other hand makes no such assumptions, allowing the correlation parameter ρ to vary in a way that optimizes how the model predicts the empirical data.

Applying this modelling approach, we examined the values of ρ and η that optimized the model fit for each participant in order to gain additional insight into the cognitive processes underlying group differences in multisensory processing and modality switching. Using the RSE-box (v1.0) toolbox (https://github.com/tomotto/RSE-box; Otto, 2018), Gaussian functions were fit to the reciprocal of the unisensory RT distributions via the LATER model approach (Noorani and Carpenter, 2016), which assumes that the reciprocals of the RT distributions are normally distributed with mean μ and SD σ (see Fig. 6B). These parameters were then used to generate the probability density function (PDF) of the maximum distribution $f_{AUV}(x) = f_A(-x) + f_V(-x)$, where

$$f_{\mathrm{A}}(x) = \frac{1}{\sigma_{\mathrm{A}} + \eta}\varphi\left(\frac{x + \mu_{\mathrm{A}}}{\sigma_{\mathrm{A}} + \eta}\right) \times \Phi\left(\frac{\rho(x + \mu_{\mathrm{A}})}{(\sigma_{\mathrm{A}} + \eta)\sqrt{1 - \rho^2}} - \frac{x + \mu_{\mathrm{V}}}{(\sigma_{\mathrm{V}} + \eta)\sqrt{1 - \rho^2}}\right) \quad (9)$$

where φ and Φ are the PDF and CDF of the standard normal distribution, respectively. Calculation of $f_V(x)$ was obtained analogously to equation 9. A more detailed description can be found in Otto and Mamassian (2012), supplementary information.

304 **Predicting multisensory benefits**

305 While Raab's race model typically underestimates the amount of multisensory benefit observed in healthy
306 adults, it has been shown to explain much of the variance in empirical benefits across participants and
307 stimulus conditions (Otto et al., 2013; Innes and Otto, 2019). This provides further evidence that the race
308 model could serve as a potential framework for the underlying cognitive architecture (Otto and
309 Mamassian, 2017). We tested whether multisensory behavior in children and individuals with ASD follow
310 the predictions of the race model or some alternative processing strategy. We proposed two alternative
311 strategies to the race model: 1) multisensory responses are biased towards a specific modality, regardless
312 of which is faster; 2) multisensory responses are biased towards the modality of the previous trial,
313 regardless of which modality is faster. Model 1 could be biased towards either the auditory (Model 1A) or
314 the visual (Model 1V) modality and can be expressed as follows:

315 $$F_{1b}(t) = \frac{1}{3}\sum_{m=1}^{3} F_b(m,t) \tag{10}$$

316 where *m* is the preceding modality (A, V, AV) and *b* is the modality that the system is biased towards (A
317 or V). Model 2 was biased towards the previous modality, except on AV trials, where it was biased towards
318 either the auditory (Model 2A) or the visual (Model 2V) modality:

319 $$F_{2b}(t) = \frac{1}{3}\left(F_{\mathrm{A}}(\mathrm{A},t) + F_{\mathrm{V}}(\mathrm{V},t) + F_b(\mathrm{AV},t)\right) \tag{11}$$

320 Each model was used to obtain a new measure of predicted benefits and assessed based on how
321 accurately it could predict empirical benefits. To examine any potential developmental transitions in
322 processing strategy, we parametrically varied the probability of a response being triggered by a race
323 strategy versus the above strategies as follows:

$$\text{benefit}_{ib} = \int_0^1 (1-p)\bar{F}_{\text{A}\cup\text{V}}(t) + pF_{ib}(t) - \max[F_{\text{A}}(t), F_{\text{V}}(t)]\, dt, \qquad \text{for } p = 0, 0.25, \ldots 1 \qquad (12)$$

where $F_{ib}(t)$ is the biased model and p is the probability of it triggering a response. When p = 0, the model constitutes a pure race model and when p = 1, the model constitutes a purely biased model.

Statistical analyses

As an initial inquiry, a linear mixed-effects model was used to determine which parameters influenced RTs. The model was fit using the maximum likelihood criterion. Single-trial RTs were the continuous numeric dependent variable. Diagnosis was a contrast-coded fixed factor (NT, ASD), age was a continuous numeric fixed factor (6–40 years), and condition was a multi-level nominal fixed factor (AV, A, V). Subjects were included as a random factor, along with by-subject slope adjustments for condition (Barr et al., 2013). ISI was included as another random factor, as well as preceding modality with slope adjustments for condition. Subsequent analyses employing standard linear models coded fixed effects as above. A one-way analysis of covariance (ANCOVA) was used to assess the correspondence between empirical and predicted benefits, treating age group as a partialled out categorical variable (Bland and Altman, 1995).

A mediation analysis (Baron and Kenny, 1986) was used to establish whether the relationship between participants' age and multisensory gain was mediated by a direct effect of age on MSE. Age was chosen as the causal variable in the model because of its known effect on race model violation (Brandwein et al., 2011). For this analysis, MSEs were averaged across the two unisensory conditions (V→A, A→V), as we hypothesized that it was a slowing of unisensory RTs that was the cause of the observed RSE. Using the M3 Toolbox (https://github.com/canlab/MediationToolbox), we constructed a three-variable mediation model with age as the causal variable, gain as the outcome variable and MSE as the mediating variable (Fig. 9C). For MSE to be considered a mediator, the following criteria must be met based on three separate regressions: 1) the causal variable must affect the outcome, 2) the causal variable must affect the

346 mediator, and 3) the mediator must affect the outcome but the causal variable must either no longer
347 affect the outcome (full mediation) or at least weaken the effect (partial mediation). Significance and SE
348 of the associated path coefficients were bootstrapped (10,000 samples) and adjusted using the bias-
349 corrected and accelerated percentile method (Wager et al., 2008).

350 All *post hoc* statistical comparisons were conducted using nonparametric permutation tests (10,000
351 permutations) based on the *t*-statistic and adjusted to control for family-wise error rate using the t_{max}
352 correction method (Westfall and Young, 1993; Blair et al., 1994). This method has been shown to control
353 for Type 1 error at a desired level when performing tests of the race model at multiple quantiles and the
354 power of the test is reasonable even for small samples (Gondan, 2010). Equivalence of variance was
355 established prior to all unpaired tests using a permuted *F*-test and the appropriate *t*-statistic was then
356 applied based on the outcome. Effect sizes were calculated using Cohen's *d* and were bias-corrected
357 according to sample size (Hedges and Olkin, 1985). All confidence intervals (CIs) were bootstrapped
358 (10,000 samples) at the 95% confidence level and adjusted using the bias-corrected and accelerated
359 percentile method (Davison and Hinkley, 1997). Correlation analyses were conducted using permuted
360 Pearson correlation or Spearman rank coefficients (Bishara and Hittner, 2012). All *post hoc* statistical tests
361 and effect size calculations were conducted using the PERMUTOOLS open-source toolbox
362 (https://github.com/mickcrosse/PERMUTOOLS).

363 # Results

364 **Reaction times and multisensory benefits**

365 A linear mixed-effects analysis was used to examine the effect of diagnosis, age and stimulus condition on
366 response times (R^2_{adj} = 0.495). Subjects, ISI and preceding modality were included as random factors, along
367 with slope adjustments for condition (see Methods for details). Participants with ASD responded more
368 slowly to stimuli than their NT peers (β = 47.6, *SE* = 12.3, *p* = 0.0001; Fig. 2A). There was an effect of

369 maturation, with older participants responding faster than younger participants (β = −9.1, *SE* = 0.84, *p* =
370 2×10^{-27}). Responses to multisensory stimuli were faster than those to both audio (β = 55.2, *SE* = 9.96, *p* =
371 3×10^{-8}) and visual (β = 67.1, *SE* = 6.7, *p* = 6×10^{-24}) stimuli, indicating the presence of an RSE. There was an
372 interaction between age and RSE (RSE_A: β = −0.6, *SE* = 0.22, *p* = 0.006; RSE_V: β = −0.48, *SE* = 0.18, *p* = 0.008).

373 To examine the RSE in detail, a general linear model was constructed to quantify the effects of diagnosis
374 and age on predicted (R^2_{adj} = 0.086) and empirical benefits (R^2_{adj} = 0.255). Predicted benefits decreased as
375 a function of age (β = −0.07, *SE* = 0.01, *p* = 1×10^{-7}) and were not significantly different in NT and ASD
376 individuals (β = 0.3, *SE* = 0.2, *p* = 0.12). Conversely, empirical benefits increased with age (β = 0.18, *SE* =
377 0.02, *p* = 2×10^{-17}) and were smaller in ASD individuals (β = −1.5, *SE* = 0.3, *p* = 7×10^{-7}). This suggests that
378 the race model over-predicts empirical benefits for younger individuals and under-predicts them for older
379 individuals (see Fig. 2C). Moreover, the race model does not predict the group differences in empirical
380 multisensory benefits, suggesting an integrative deficit.

381

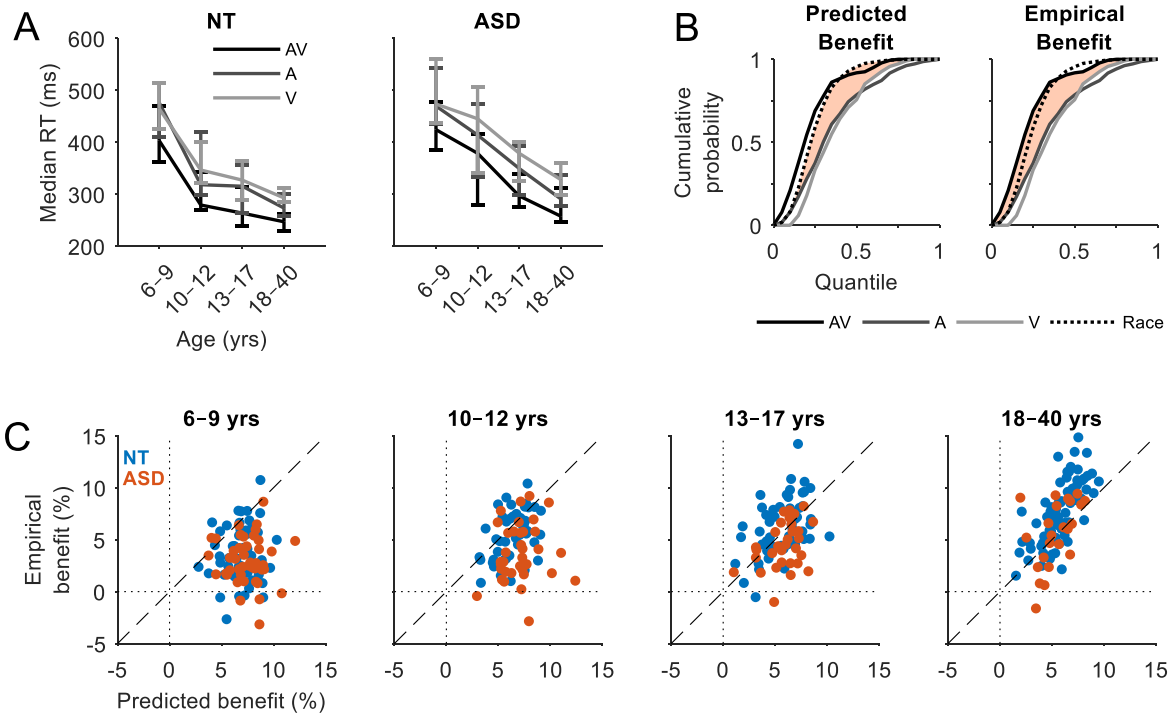


Figure 2. Reaction times and multisensory benefits. **A**, Group median RTs for NT (left panel) and ASD (right panel) individuals as a function of age group. Error bars indicate 95% CIs (bootstrapped). **B**, RT cumulative probability for each of the three stimulus conditions and the race model (Eq. 4). Predicted benefits (left panel) are quantified by the area between the CDFs of the race model and the faster of the unisensory conditions (Eq. 5). Empirical benefits (right panel) are quantified by the area between the CDFs of the multisensory condition and the faster of the unisensory conditions (Eq. 6). Data from an example NT adult participant. **C**, Predicted benefits versus empirical benefits by age group. Each datapoint represents an individual participant (blue = NT, red = ASD).

Testing the race model

To determine whether the RSE exceeded statistical facilitation, we compared the multisensory CDFs to the race model at each of the first 7 quantiles (maximum number of quantiles violated by any group). Violation of the race model was assessed using right-tailed permutation tests with t_{max} correction (Gondan, 2010). NT participants showed evidence of violation at one or more quantiles in every age group,

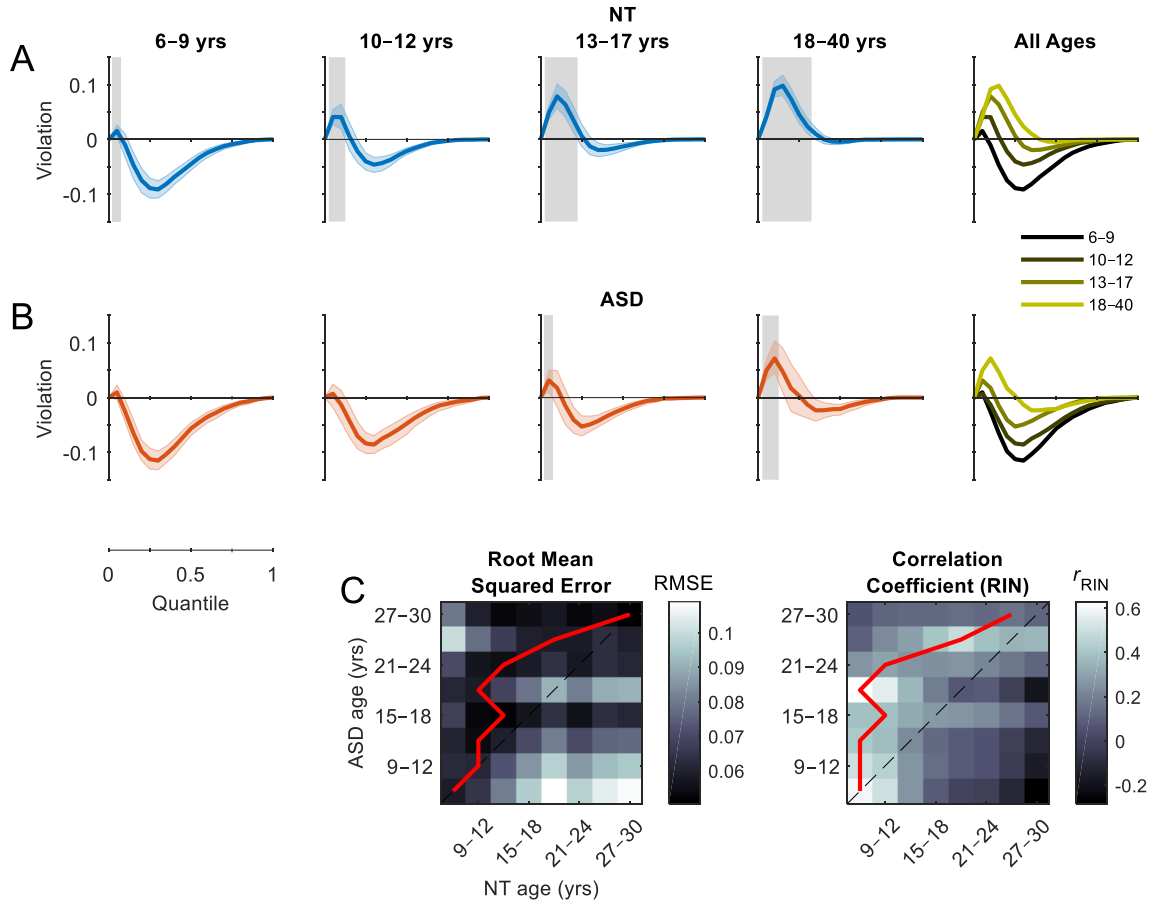
396 the number of quantiles increasing as a function of age ($p < 0.05$, shaded area, Fig. 3A). The percentage
397 of participants that exceeded statistical facilitation at each quantile is illustrated in Figure S1. Individuals
398 with ASD showed no evidence of violation between the ages of 6–12 years (Fig. 3B). However, evidence
399 of violation emerges in adolescence (first quantile) and becomes more evident in adulthood (first 2
400 quantiles; see Table 2 for the statistics of each race model test). Note, these results were replicated
401 qualitatively using the more conservative Miller's bound, albeit at less quantiles (see Table S1).

402 **Table 2.** Test statistics comparing CDFs of multisensory RTs with the race model. Values shown indicate
403 effect sizes (Cohen's *d* corrected for sample size) and 95% CIs (bootstrapped) in brackets. Asterisks
404 indicate significant race model violation ($p < 0.05$, right-tailed permutation tests, t_{max} corrected).

	NT				ASD			
Q	6–9 yrs	10–12 yrs	13–17 yrs	18–40 yrs	6–9 yrs	10–12 yrs	13–17 yrs	18–40 yrs
1	0.18[0.1,0.3]*	0.44[0.3,0.7]*	0.54[0.4,0.9]*	0.83[0.6,1.1]*	0.12[-0.0,0.3]	0.07[-0.1,0.3]	0.30[0.1,0.6]*	0.62[0.4,1.1]*
2	-0.06[-0.2,0.1]	0.21[0.1,0.4]*	0.41[0.3,0.6]*	0.73[0.6,0.9]*	-0.14[-0.2,-0.0]	-0.06[-0.2,0.1]	0.09[-0.1,0.3]	0.38[0.2,0.6]*
3	-0.21[-0.3,-0.1]	0.03[-0.1,0.2]	0.29[0.2,0.4]*	0.59[0.4,0.8]*	-0.29[-0.4,-0.2]	-0.19[-0.3,-0.1]	-0.08[-0.3,0.1]	0.19[0.0,0.4]
4	-0.35[-0.5,-0.2]	-0.11[-0.2,-0.0]	0.17[0.1,0.3]*	0.44[0.3,0.6]*	-0.48[-0.7,-0.3]	-0.34[-0.6,-0.2]	-0.29[-0.6,-0.1]	0.07[-0.1,0.3]
5	-0.53[-0.8,-0.4]	-0.27[-0.4,-0.2]	0.03[-0.1,0.2]	0.31[0.2,0.4]*	-0.66[-1.1,-0.5]	-0.50[-0.8,-0.3]	-0.54[-0.9,-0.3]	0.01[-0.2,0.2]
6	-0.80[-1.1,-0.6]	-0.43[-0.7,-0.3]	-0.09[-0.2,0.0]	0.21[0.1,0.3]*	-0.89[-1.4,-0.6]	-0.70[-1.2,-0.5]	-0.75[-1.3,-0.5]	-0.07[-0.3,0.1]
7	-1.00[-1.4,-0.8]	-0.55[-0.8,-0.4]	-0.20[-0.4,-0.1]	0.12[-0.0,0.2]	-1.12[-1.7,-0.8]	-0.75[-1.3,-0.6]	-0.97[-1.5,-0.6]	-0.17[-0.4,-0.0]

405 To compare race model violation between NT and ASD individuals of different ages, we computed the
406 root-mean-square error (RMSE) and correlation coefficient between each participant's violation function
407 and that of every other participant. Because the violation functions are typically non-normal, we applied
408 a rank-based inverse normal (RIN) transformation (Bliss, 1967), prior to assessing the Pearson correlation.
409 Participants were split into 8 age groups separated by 3 years between the ages of 6–30 years (there were
410 too few participants above 30 years of age). Matrices containing RMSE and correlation values were
411 obtained by averaging over the values within each age group (Fig. 3C). The red line in Figure 3C indicates
412 the age groups that are most similar, and its divergence above the dotted midline suggests that
413 multisensory behavior in ASD participants corresponded more closely to that of younger NT participants,

414 i.e., a developmental delay. Convergence of the red and dotted lines suggests that this delay may recover
415 in adulthood, in line with our original hypothesis. This is further examined in the following section.



416

417 **Figure 3.** Testing the race model. **A**, **B**, Violation of the race model is quantified by the difference between
418 the CDFs of the multisensory condition and the race model. Positive values reflect quantiles where
419 multisensory RTs were faster than predicted by the race model. Gray shaded regions indicate significant
420 differences ($p < 0.05$, right-tailed permutation tests, t_{max} corrected). Colored error bounds indicate 95%
421 CIs (bootstrapped). **C**, Root mean squared error (left panel) and RIN-transformed Pearson correlation
422 coefficient (right panel) between the violation functions for NT and ASD participants of different ages
423 (range: 6–30 years, increment: 3 years). Red lines indicate the minimum (left panel) and maximum (right
424 panel) values of each row (i.e., the age groups that were most similar). Divergence of the red line above
425 the dotted midline indicates a developmental delay in ASD participants.

Delayed multisensory development in autism

We constructed a linear model to evaluate the effects of diagnosis and age on multisensory gain (R^2_{adj} = 0.388). Multisensory gain, as indexed by the AUC (Eq. 6, Fig. 4A), increased as a function of age (β = 0.25, *SE* = 0.02, $p = 2\times10^{-21}$) but was significantly reduced in participants with ASD compared to NT individuals (β = −1.98, *SE* = 0.68, p = 0.004). The absence of an interaction suggests that this maturation effect was present in both groups (β = 0.01, *SE* = 0.04, p = 0.77). *Post hoc* comparisons were conducted within each of the four age groups. For this analysis, NT participants were sex- and age- matched to each of the ASD participants and compared at every quantile using two-tailed (unpaired) permutation tests. Group differences were observed in the adolescent group at quantiles 4 and 5 ($p < 0.05$, shaded area, Fig. 4C). To compare the overall multisensory gain, we conducted permutation tests on the AUC (Fig. 4D), revealing differences in participants aged 10–12 years ($t_{(50)}$ = 2.22, p = 0.031, d = 0.61, 95CI [0.1, 1.15]) and 13–17 years ($t_{(60)}$ = 2.57, p = 0.014, d = 0.65, 95CI [0.18, 1.15]), but not 6–9 years ($t_{(60)}$ = 0.88, p = 0.39, d = 0.21, 95CI [−0.28, 0.72]) or 18–40 years ($t_{(44)}$ = 1.81, p = 0.077, d = 0.52, 95CI [−0.03, 1.19]). The moderate effect size in the adult group suggests that individuals with ASD might not have "caught up" entirely by 18 years of age.

The effect of maturation can be seen more clearly by charting multisensory gain as a function of age (Fig. 4E). Age was highly predictive of multisensory gain between 6–17 years (NT: R^2 = 0.34, p = 0; ASD: R^2 = 0.21, p = 0) but not between 18–40 years (NT: R^2 = 0.005, p = 0.56; ASD: R^2 = 0.052, p = 0.296), suggesting that maturation of this process ceases in adulthood. To characterize this developmental trajectory more precisely, we calculated the mean multisensory gain with a moving window k of 7 years in increments of 1 year (Fig. 4F). Controls were sex- and age- matched to ASD individuals within each 7-year window and compared using two-tailed permutation tests (FDR corrected). In NT participants, multisensory gain increased steadily between 6–18 years of age. In individuals with ASD, the rate of increase was more gradual and was significantly lower than that of their NT peers between the ages of 11–21 years ($p < 0.05$,

shaded area, Fig. 4F). However, by the mid-twenties, multisensory gain was commensurate with that of NT individuals suggesting that this deficit recovers in adulthood, confirming our original hypothesis. Given that maturation appears to continue well into adulthood, a *post hoc* analysis was conducted whereby the adult group was subdivided into participants aged 18–23 years (n = 12) and 24–40 years (n = 11) to examine multisensory gain before and after this "catch up" point. As expected, there were significant group differences in adults aged 18–23 years ($t_{(20)}$ = 2.24, p = 0.039; d = 0.92, 95CI [0.18, 1.98]; Fig. 4H, left) but not in adults aged 24–40 years ($t_{(22)}$ = 0.36, p = 0.72; d = 0.14, 95CI [−0.64, 1.0]; Fig. 4H, right). Group average violation functions were almost identical at every quantile between NT and ASD adults aged 24–40 years (Fig. 4G, right), suggesting both qualitative and quantitative recovery.

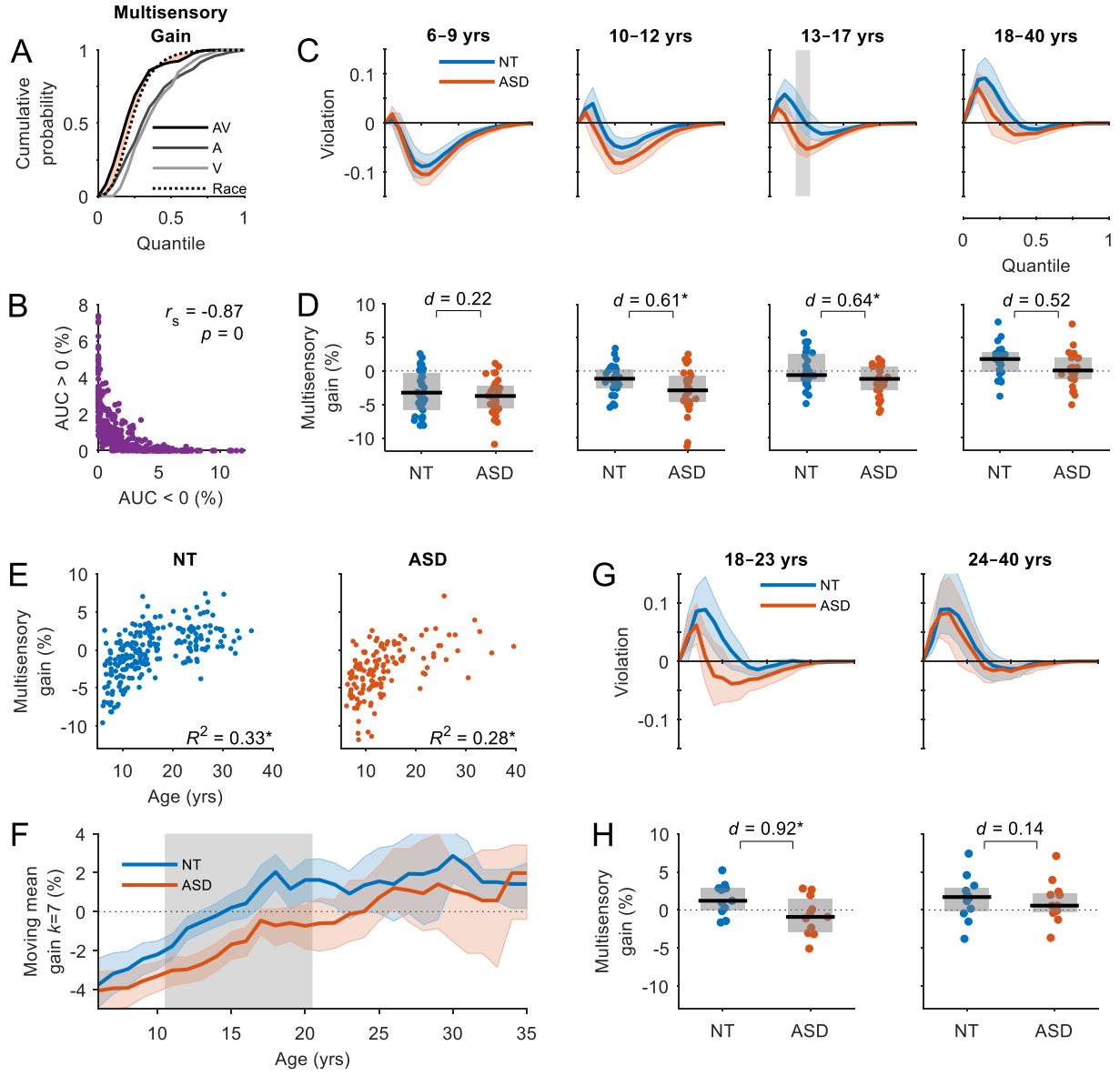


Figure 4. Developmental course of multisensory gain. **A**, RT cumulative probability for each of the three stimulus conditions and the race model. Multisensory gain is quantified by the area between the CDFs of the multisensory condition and the race model (Eq. 7). Data from an example NT adult participant. **B**, The area under the curve (AUC) below zero is negatively correlated with the AUC above zero, providing information about participants that do not exceed statistical facilitation. **C**, Race model violation for ASD (red trace) and sex- and age- matched NT (blue trace) participants by age group. Colored error bounds indicate 95% CIs (bootstrapped). Gray shaded regions indicate significant group differences ($p < 0.05$, two-

tailed permutation tests, t_{max} corrected). **D**, Multisensory gain by age group. Boxplots indicate the median value (black line) and interquartile range (grey box). Each datapoint represents an individual participant (blue = NT, red = ASD). Brackets indicate unpaired statistical comparisons (*p < 0.05, two-tailed permutation tests, FDR corrected). **E**, Multisensory gain as a function of age for NT (left) and ASD (right) individuals. Each datapoint represents an individual participant. **F**, Mean multisensory gain calculated with a moving window k of 7 years in increments of 1 year from 6–35 years for NT (blue trace) and ASD (red trace) participants. Colored error bounds indicate 95% CIs (bootstrapped). Gray shaded regions indicate significant group differences (p < 0.05, two-tailed permutation tests, FDR corrected). **G**, Race model violation for ASD and sex- and age- matched NT adults separated into 18–23 years (n = 12, left) and 24–40 years (n = 11, right). **H**, Multisensory gain for the same adult groups.

Modality switch effects

To quantify MSEs, we derived separate CDFs for switch and repeat trials and computed the area between them (Eq. 8). We modelled the effects of diagnosis, age and condition on MSEs using a linear model (R^2_{adj} = 0.303). MSEs increased with age (β = 0.17, SE = 0.02, $p = 7\times10^{-24}$) and were reduced in individuals with ASD compared to NT individuals (β = −1.18, SE = 0.24, $p = 7\times10^{-7}$). Compared to multisensory trials, MSEs were larger on both auditory trials (β = 4.67, SE = 0.03, $p = 6\times10^{-57}$) and visual trials (β = 3.66, SE = 0.03, $p = 3\times10^{-37}$). Follow-up permutation tests revealed that MSEs were only reduced in the adolescent ASD group, and only when switching from auditory to visual stimuli ($t_{(60)}$ = 2.76, p = 0.021, d = 0.69, 95CI [0.22, 1.19]; Fig. 5A). A more detailed examination using a moving mean estimate of MSE showed that group differences emerged between the ages of 10–16 years (p < 0.05, shaded area, Fig. 5B, right). The maturational course of visual to auditory MSEs appears to continue later into development than that of auditory to visual switches in both groups (Fig. 5B, left).

Contrary to our results, a study by Williams et al. (2013) found that individuals with ASD between the ages of 8–15 years exhibited a greater cost to switching from auditory to visual stimuli than their age-matched NT peers. To make a more direct comparison with their study, we performed a two-tailed permutation test on a group of sex- and age- matched participants between the ages of 8–15 years (n = 72) and used a similar measure of MSE based on mean RT values. This approach yielded the same outcome as before, with ASD individuals exhibiting smaller MSEs (NT: 30.5 ± 27.4 ms, ASD: 19.7 ± 38.7 ms; $t_{(142)}$ = 1.93, p = 0.049, d = 0.32, 95CI [−0.004, 0.66]), confirming the discrepancy was not the result of how MSE was quantified. The only remaining difference between our two studies was that Williams et al. (2013) used longer ISIs (3–5 s versus 1–3 s). Thus, we repeated the test focusing on RTs with preceding ISIs between 2.5–3 s. Limiting the analysis to longer ISIs caused a significant drop in MSE for NT individuals (16.2 ± 42.8 ms) but not so much for individuals with ASD (16.8 ± 51.5 ms). Moreover, this modification revealed no group differences ($t_{(142)}$ = −0.07, p = 0.954, d = −0.01, 95CI [−0.34, 0.31]), suggesting invocation of disparate mechanisms underlying MSEs at shorter versus longer ISIs.

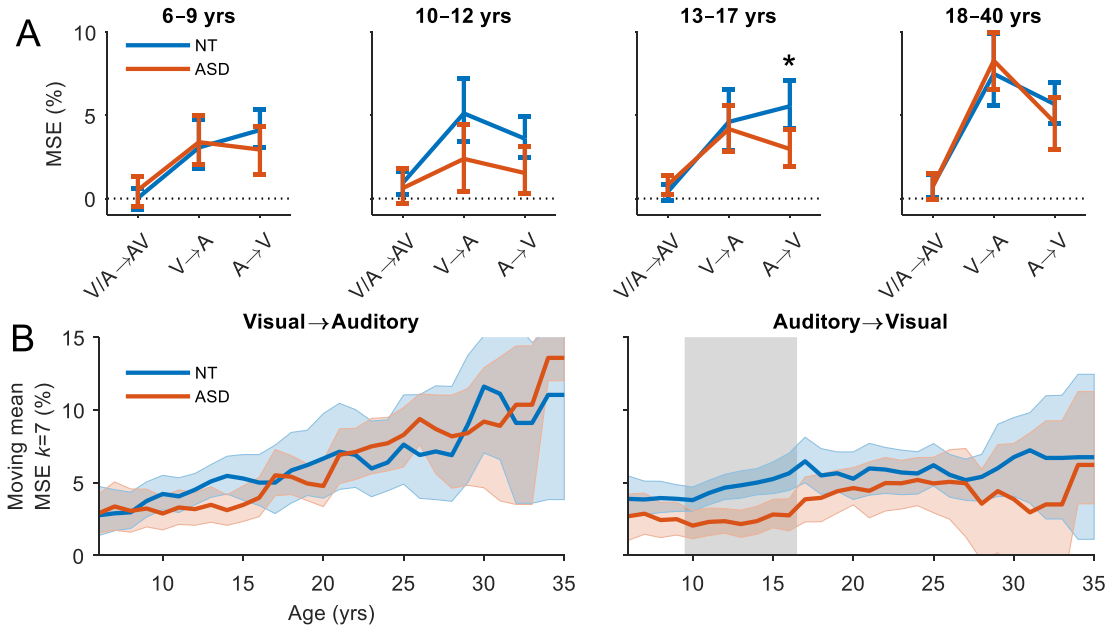


Figure 5. Modality switch effects. **A**, Mean MSE for each condition by age group. MSEs were quantified by the area between the CDFs of the switch and repeat trials (Eq. 8). Error bars indicate 95% CIs

(bootstrapped). Asterisks indicate significant group differences ($p < 0.05$, two-tailed permutation tests, t_{max} corrected). **B**, Mean MSE for visual to auditory (left panel) and auditory to visual (right panel) switches calculated with a moving window k of 7 years in increments of 1 year from 6–35 years for NT (blue trace) and ASD (red trace) participants. Colored error bounds indicate 95% CIs (bootstrapped). Gray shaded regions indicate significant group differences ($p < 0.05$, two-tailed permutation tests, FDR corrected).

Divided attention in autism

To gain a better understanding of what aspects of multisensory processing led to differences in behavior, we adopted a computational framework based on the race model (Otto and Mamassian, 2012). The inclusion of 2 additional free parameters in the race model allowed us to quantify the additional variability or noise η in empirical multisensory RTs, as well as the correlation ρ between RTs on different sensory channels, giving us insight into how attention is divided between them (see Methods for details). We hypothesized that the increase in RT variability would be larger for individuals with higher multisensory gain, and that channel dependency would be lower or more negatively correlated for individuals with greater MSEs. The best-fitting estimates of the noise parameter η increased with age ($\beta = 0.0045$, $SE = 0.0008$, $p = 4\times10^{-8}$) but was not statistically different between NT and ASD participants ($\beta = -0.016$, $SE = 0.012$, $p = 0.17$; $R^2_{adj} = 0.0899$; Fig. 6C). The best-fitting estimates of the correlation parameter ρ decreased with age ($\beta = -0.035$, $SE = 0.003$, $p = 1\times10^{-29}$) and were lower (and sometimes more negative) for NT individuals ($\beta = -0.25$, $SE = 0.04$, $p = 2\times10^{-9}$; $R^2_{adj} = 0.38$; Fig. 6D). *Post hoc* permutation tests revealed moderate group differences in participants aged 10–12 years ($t_{(50)} = 1.97$, $p = 0.05$, $d = 0.54$, 95CI [0.01, 1.18]) and 13–17 years ($t_{(60)} = 2.15$, $p = 0.036$, $d = 0.54$, 95CI [0.06, 1.06]). This greater (more positive) channel dependency in ASD suggests a more even spread of attention across sensory systems.

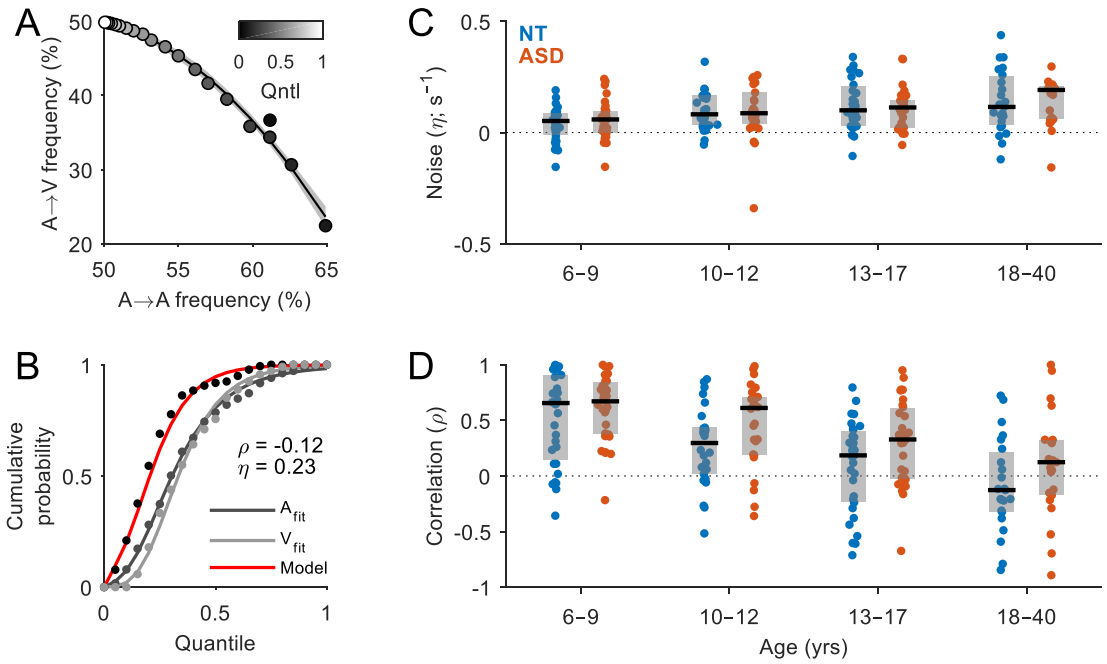


Figure 6. Modelling channel dependency and RT variability. **A**, Frequency of visual and auditory trials preceded by auditory trials in each quantile (i.e., switch versus repeat trials). Quantiles are indicated by a grayscale, graduating from black (fastest quantile) to white (slowest quantile). Example data averaged over all NT adult participants. **B**, CDFs were fit to the unisensory RT data and used to predict empirical multisensory RT data via Otto's context variant race model (Otto and Mamassian, 2012). Free parameters ρ and η account for the correlation between RTs on different channels and increased RT variability or noise, respectively. Data from an example NT adult participant. **C**, **D**, Best-fitting model parameters ρ and η by diagnosis and age group. Boxplots indicate the median value (black line) and interquartile range (grey box). Each datapoint represents an individual participant (blue = NT, red = ASD).

Modelling multisensory development

To assess whether the race model could predict multisensory benefits in children and individuals with ASD, one-way ANCOVAs were used to measure the correlation between predicted and empirical benefits in each age group. Predicted benefits were correlated with empirical benefits in both NT ($F_{(1,217)} = 62.86$, $p = 1\times10^{-13}$, $R^2 = 0.23$) and ASD ($F_{(1,125)} = 5.25$, $p = 0.024$ $R^2 = 0.04$) individuals but an interaction suggested

that this relationship was dependent on age group (NT: $F_{(3,217)} = 5.4$, $p = 0.0013$, $R^2 = 0.07$; ASD: $F_{(3,125)} = 2.58$, $p = 0.057$, $R^2 = 0.06$). Figure 7A, B shows that the ability of the race model to predict empirical benefits increases significantly over the course of development. While a race model account predicts a significant proportion of the variance in the adult groups (NT: $R^2 = 0.49$, $p = 0$; ASD: $R^2 = 0.25$, $p = 0.017$), it accounted for almost none of the variance in the youngest (6–9 years) groups (NT: $R^2 = 0.007$, $p = 0.55$; ASD: $R^2 = 0.002$, $p = 0.78$). If multisensory benefits in young children cannot be explained by probability summation, then how do we model the underlying cognitive architecture?

To address this question we proposed two alternative models: 1) multisensory RTs are biased towards a specific modality (Model 1A, V), 2) multisensory RTs are biased towards the modality of the preceding trial (Model 2A, V; see Methods for details). We parametrically varied the probability p of a multisensory response being triggered by a race strategy or one of the above bias-driven strategies and assessed how well each model could predict empirical benefits. Figure 7C shows that Model 1A was most accurate at predicting the variance in empirical benefits across children with ASD aged 6–9 years, suggesting that their responses were mostly triggered by the previous modality, with a bias towards the auditory modality. In their NT counterparts, none of the models provided significant improvement beyond the race model, although there was evidence for a bias towards the auditory modality as well (Fig. 7C). In children with ASD aged 10–12 years, Model 1V provided a significant improvement in performance, suggesting that RTs were largely determined by the previous modality, but this time, with a bias towards the visual modality (Fig. 7D). In NT children aged 10–12 years, there was no major improvement beyond the race model again, but there was a slight bias towards the visual modality as well. In teenagers and adults, none of the models outperformed the race model suggesting that individuals with ASD begin to adopt a race strategy by adolescence (Fig. S2). Before this stage, it appears that they are biased towards the preceding modality regardless of which sensory modality is faster.

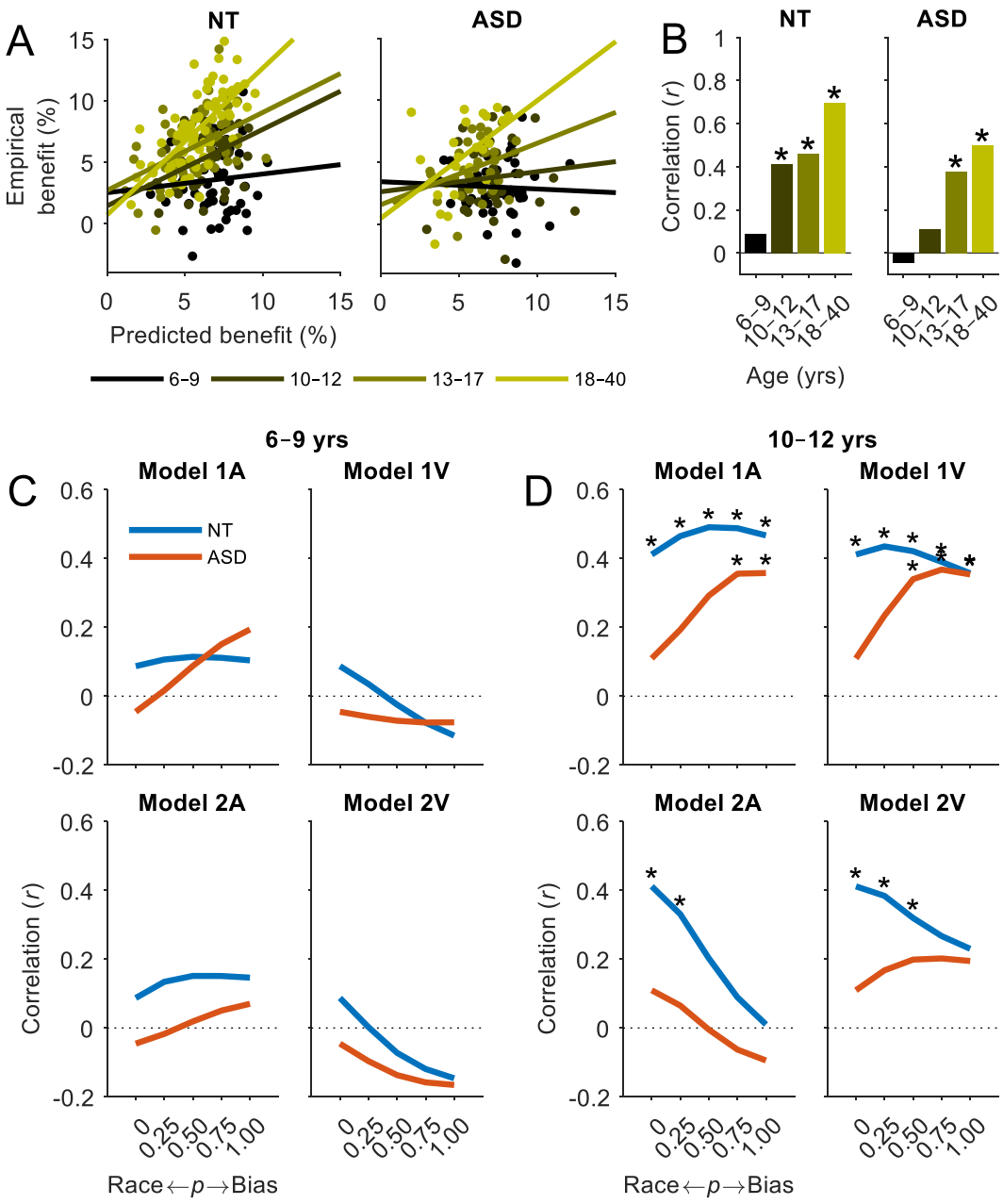


Figure 7. Predicting multisensory benefits. **A**, Predicted benefits versus empirical benefits for NT (left panel) and ASD (right panel) participants. Each datapoint represents an individual participant and age group is indicated by color. Solid lines represent linear fits to the data by age group. **B**, Pearson correlation coefficient (*r*) of the regression fits in panel A. Asterisks indicate significant correlations ($p < 0.05$, two-tailed permutation tests). **C**, **D**, Four alternative models of multisensory processing were tested. Model 1A was biased towards the auditory modality and Model 1V towards the visual modality. Model 2A was

biased towards the preceding modality and the A modality when preceded by an AV trial, and Model 1V was biased towards the preceding modality and the V modality when preceded by an AV trial. The probability *p* of a response being triggered by a race strategy or a biased strategy was parametrically varied between 0 and 1 in increments of 0.25. The ability of each model to predict the variance in empirical benefits was assessed within each age group based on the Pearson correlation coefficient as in panel B. Data presented are the two younger age groups. See supplementary material for the two older age groups.

Developmental changes in sensory dominance

To further examine developmental patterns in sensory dominance, we tested the same models but with the probability of a sensory-specific bias set to 1 (Fig. 8A). Evaluating model performance as before, we noticed an auditory dominance in both groups at 6–9 years of age that shifted to a visual dominance by 10–12 years of age. In the NT group, this visual dominance appears to continue into adulthood in accordance with the well-known Colavita visual dominance effect (Colavita, 1974). However, in the ASD group, this sensory weighting appears to shift once again in adolescence, leading to an auditory dominance in adulthood.

If such sensory dominances genuinely exist when processing multisensory information, one would expect to see a greater MSE when switching from the less dominant modality to an AV stimulus. To test this hypothesis, we examined MSEs on AV switch trials, this time separating trials preceded by A and V stimuli (Fig. 8B). MSEs were normalized by MSEs for the grouped V/A→AV trials to allow for meaningful comparison across age groups (this did not change the results qualitatively). Based on our modelling analysis, we expected to see greater MSEs on V→AV trials for TD children (6–9 years) and on A→AV trials for older TD children and adults (10–40 years). We expected something similar for ASD individuals with another shift in adolescence. The data in Figure 8B suggest that, as predicted, MSEs were greater on V→AV

trials for TD children (6–9 years) and on A→AV trials for older TD children and adults (10–40 years). For ASD individuals, the data suggest the reverse, with greater MSEs on A→AV trials in children and teenagers (6–17 years) and on V→AV trials for adults (18–40 years).

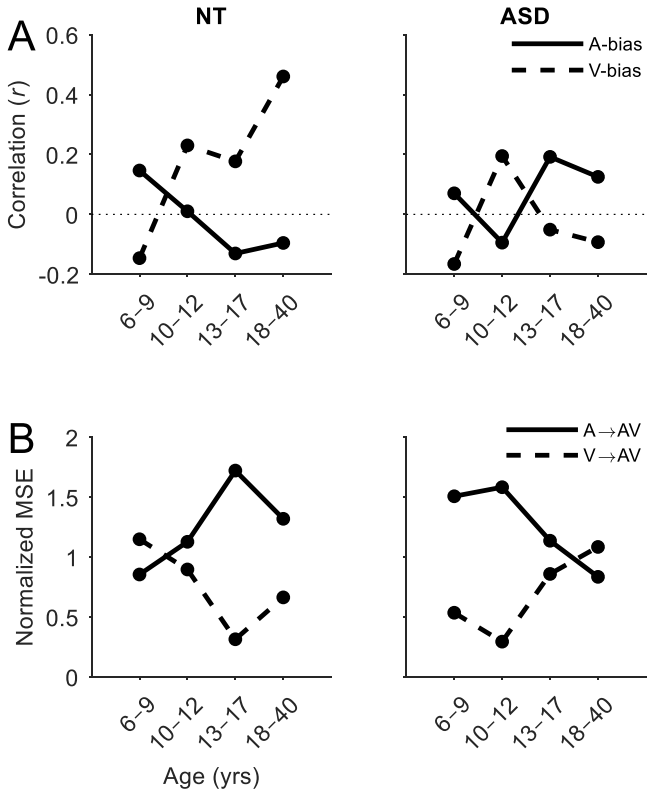


Figure 8. Sensory dominance during audiovisual processing. **A**, Developmental changes in sensory dominance were examined by measuring the performance of models 2A (A-bias, solid trace) and 2V (V-bias, dotted trace) with the probability of a sensory bias *p* set to 1. The ability of each model to predict the variance in empirical benefits was assessed within each age group based on the Pearson correlation coefficient. **B**, Modality switch effects for AV trials separated by trials preceded by A-stimuli (solid trace) and V-stimuli (dotted trace). MSEs were quantified by the area between the CDFs of the switch and repeat trials and normalized by the grouped V/A→AV MSEs.

Linking modality switch effects and redundant signals effects

To examine the relationship between MSEs and multisensory gain, we performed a series of partial correlations across participants, controlling for age (Table 3). As one might predict, there was a strong positive correlation between the average multisensory gain on switch trials and the average MSE on unisensory trials (but not on multisensory trials). However, there was no significant correlation between multisensory gain on repeat trials and MSEs on unisensory trials, whereas there was a strong positive correlation with MSEs on multisensory trials. This pattern, which was identical in both groups (see Fig. S3), confirms that MSEs on unisensory trials are more likely to contribute to multisensory gain. Figure 9A, B illustrates the impact of switching sensory modality on race model violation and multisensory gain, respectively. 87% of NT individuals exhibited a larger multisensory gain on switch trials than on repeat trials ($t_{(224)} = 15.62$, $p = 0$, $d = 0.84$, 95CI [0.73, 0.96]), with 82% of individuals with ASD showing the same ($t_{(132)} = 6.74$, $p = 0$, $d = 0.51$, 95CI [0.35, 0.68]). Nevertheless, when we submitted RTs from the repeat trials to a race model test, every group violated the race model as before except the adolescent ASD group (Table S2), even when using a more conservative test based on Miller's bound (Table S3).

Table 3. Partial correlations between multisensory gain and MSEs, controlling for age. Multisensory gain was computed separately for switch trials (left columns) and repeat trials (right columns). Values indicate coefficients of determination (R^2) and significance of the correlation (p).

	Gain on Switch Trials			Gain on Repeat Trials		
	V/A→AV	V→A	A→V	V/A→AV	V→A	A→V
NT	$R^2 = 0.001$	$R^2 = 0.3$	$R^2 = 0.18$	$R^2 = 0.25$	$R^2 = 2\times10^{-6}$	$R^2 = 0.003$
	$p = 0.7$	$p = 3\times10^{-19}$	$p = 3\times10^{-11}$	$p = 1\times10^{-15}$	$p = 0.98$	$p = 0.43$
ASD	$R^2 = 0.035$	$R^2 = 0.29$	$R^2 = 0.21$	$R^2 = 0.31$	$R^2 = 0.026$	$R^2 = 0.001$
	$p = 0.03$	$p = 2\times10^{-11}$	$p = 4\times10^{-8}$	$p = 3\times10^{-12}$	$p = 0.06$	$p = 0.75$

Having established the relationship between MSEs and multisensory gain, we wished to determine whether the contribution of the former was a full or partial. To do this, we submitted the data to a mediation analysis (Wager et al., 2008). Specifically, we tested whether MSEs mediated the relationship between participant age and multisensory gain (Fig. 9C, D). First, we established that age was a reliable

predictor of both MSE (NT: β = 0.25, *SE* = 0.03, *p* = 0.0002; ASD: β = 0.22, *SE* = 0.05, *p* = 0.001) and multisensory gain (NT: β = 0.25, *SE* = 0.02, *p* = 0.001; ASD: β = 0.27, *SE* = 0.04, *p* = 0.0002), meeting the first two criteria for mediation (see Methods for details). MSE affected gain, controlling for age (NT: β = 0.23, *SE* = 0.05, *p* = 0.0002; ASD: β = 0.33 *SE* = 0.08, *p* = 0.0001) and the mediation effect was significant for both groups (NT: β = 0.06, *SE* = 0.01, *p* = 0.0002; ASD: β = 0.07 *SE* = 0.02, *p* = 0.0001). However, there was still a significant direct path between age and gain when controlling for MSE (NT: β = 0.19, *SE* = 0.03, *p* = 0.0002; ASD: β = 0.19 *SE* = 0.04, *p* = 0.0004), indicating that MSE only partially mediated the observed relationship between age and multisensory gain.

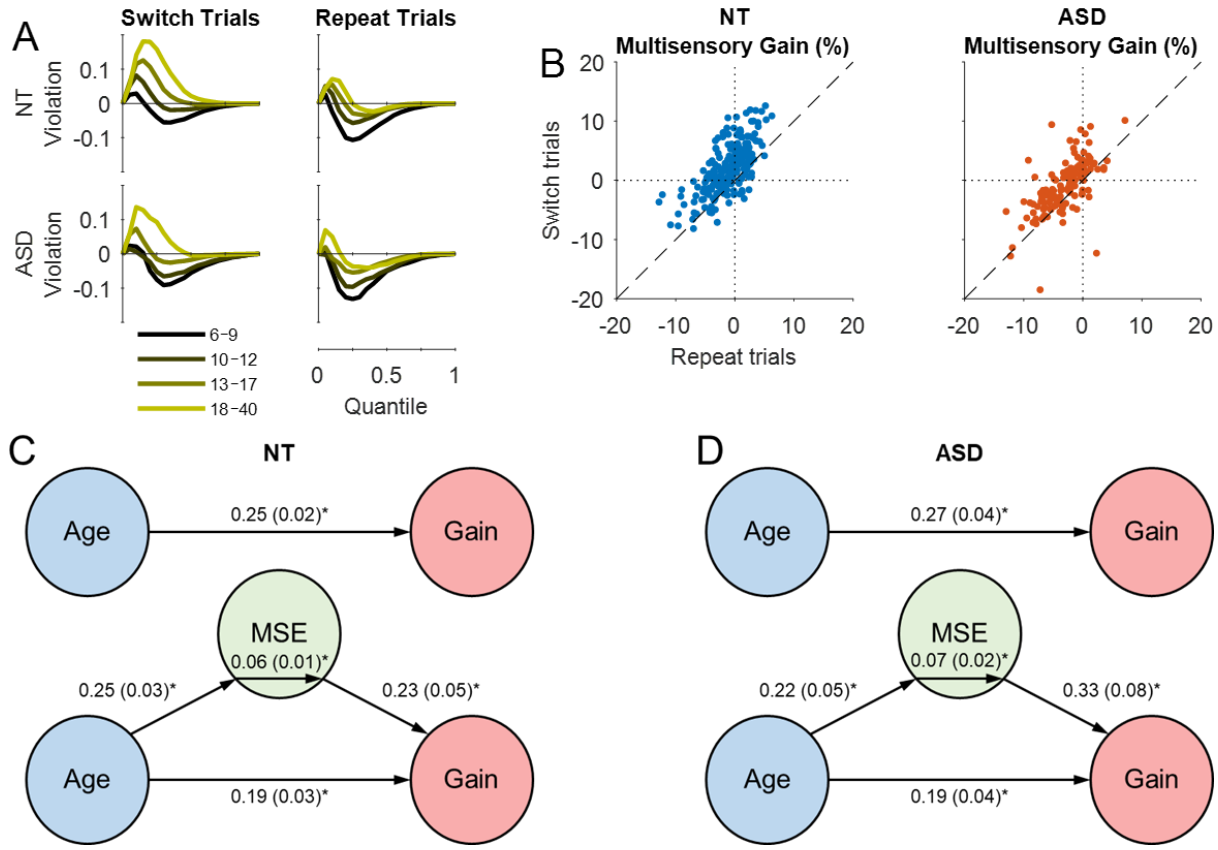


Figure 9. Linking modality switch effects and redundant signals effects. **A**, Race model violation by diagnosis and age for switch trials (left panel) and repeat trials (right panel). **B**, Multisensory gain on switch trials versus repeat trials for NT (left panel) and ASD (right panel) individuals. Each datapoint represents

638 an individual participant. **C**, **D**, Mediation model that tested whether modality switch effects (MSEs)
639 mediated the effect of age on multisensory gain. Paths between nodes are labeled with regression
640 coefficients, with SE in parentheses (*$p < 0.001$, bootstrapped). In both groups, age predicted gain (top
641 path), and predicted MSE controlling for gain (lower left path). The middle coefficients indicate formal
642 mediation effects but the significant direct paths between age and gain controlling for MSE (bottom path)
643 suggest only partial mediation, i.e., MSE did not explain all of the shared variance between age and gain.

644 ## Discussion

645 Our data suggest that the amelioration of multisensory deficits in ASD generalizes to the case of nonsocial
646 AV stimuli, but that the developmental trajectory of this recovery is protracted compared to that observed
647 in AV speech studies (e.g., Taylor et al., 2010; Foxe et al., 2015). We hypothesized that this delay may be
648 due to lack of environmental exposure to such ecologically-irrelevant stimuli (Beker et al., 2017; Cuppini
649 et al., 2017), or engagement of neural processes with longer developmental trajectories. Indeed,
650 multisensory gain in NT individuals has been shown to reach full maturity much later for simple AV stimuli
651 such as those used here (Brandwein et al., 2011) compared to AV speech stimuli (Ross et al., 2011). This
652 undoubtedly effects the average age at which individuals with ASD catch up to their NT peers, suggesting
653 that it is important to consider the maturational course in typically-developing individuals within different
654 contexts when examining developmental recovery in any clinical population.

655 The disparity in multisensory development for speech and non-speech stimuli likely reflects the fact that
656 multisensory processing occurs across distributed networks and that different stimuli and tasks tap into
657 unique processes with varying maturational courses (Chandrasekaran, 2017). The task employed in the
658 current study required the speeded detection of simple AV stimuli, without discrimination, identification
659 or any higher-order cognitive processing. Integration of such simple AV stimuli likely consists of early
660 cross-sensory activation of visual and auditory cortical regions, enhancing detection of the incoming visual

and auditory inputs, respectively (Molholm et al., 2002; Mercier et al., 2013; Mercier et al., 2015). In contrast, identification of AV speech engages an extensive network of hierarchically-organized brain areas (Hickok and Poeppel, 2007; Peelle, 2019), projecting the spectrotemporal dynamics to a phonetic representation and from there to a lexical-semantic one. Moreover, integration of auditory and visual speech cues may act through multiple integrative mechanisms (see Peelle and Sommers, 2015); 1) an early mechanism that provides information about the timing of the incoming acoustic input, activating auditory cortex and increasing perceptual sensitivity (Megevand et al., 2018), 2) a later mechanism that provides information about the content of a vocal utterance (i.e., place and/or manner of articulation), reducing the density of phonemic and lexical neighborhoods (Tye-Murray et al., 2007). Clearly, task demands and stimuli play a major role in the patterns of multisensory deficits and recovery functions that are observed for any given experimental paradigm.

Alternatively, differences in maturational patterns could be caused by influences from task-specific, extraneous neural processes. Phenomena such as modality switch effects, which contribute significantly to multisensory gain in a bisensory detection task but not in an AV speech identification task, could prolong the perceived maturational course of multisensory processing. While this is consistent with the fact that maturation of MSEs (visual to auditory) extended well into adulthood (Fig. 5B, left), the developmental trajectory of multisensory gain was qualitatively unchanged when the contribution of MSEs was diminished by focusing on the repeat trials (Fig. S4). This, and the results of the mediation analysis, suggest that MSEs are not the sole driving factor behind our measure of multisensory gain and its prolonged maturational course. Another extraneous factor that could impact the developmental trajectory of multisensory gain in the context of a bisensory detection task is the underlying processing strategy used to couple decisions on separate sensory channels and trigger a speeded response (e.g., race strategy versus sensory-specific bias). This factor would likely have no influence in the context of an AV speech identification task, thus potentially contributing to the disparity in multisensory development.

685 **Modality switch effects**

686 One of the unexpected findings to emerge from our analysis was the reduced cost of switching sensory
687 modality (auditory to visual) in ASD participants between the ages of 10–16 years. This ran contrary to a
688 recent study (Williams et al., 2013) that reported larger switch costs in individuals with ASD of
689 approximately the same age. Interestingly, a *post hoc* analysis of our data that focused on trials with
690 longer ISIs closer to that of Williams et al. (2013) led to a significant reduction in MSEs in NT individuals
691 and only a slight reduction in individuals with ASD. This modification revealed no group difference,
692 suggesting an interaction between group and ISI. A possible explanation for this interaction comes from
693 work investigating MSEs in schizophrenia patients and the so-called "trace theory" (Zubin, 1975). This
694 theory suggests that sensory information leaves traces of residual activity in different neuronal
695 populations, facilitating the processing of subsequent stimuli of the same sensory modality and inhibiting
696 the processing of stimuli of other modalities. Zubin (1975) predicted that these traces attenuate over time
697 but persist longer in individuals with schizophrenia. If such an inhibitory cross-sensory mechanism were
698 weaker in individuals with ASD, but persisted longer over time, it would explain the interaction that we
699 observe here and the findings of Williams et al. (2013). Evidence in support of this theory comes from a
700 recent study that demonstrated that individuals with ASD weight recent stimuli less heavily than NT
701 individuals and that their perception is dominated by longer-term statistics (Lieder et al., 2019).

702 Reduced cross-sensory inhibition would undoubtedly make it easier to process subsequent sensory
703 information in other modalities, leading to lower MSEs. However, it would also likely result in lower
704 attentional engagement with task/sensory-relevant information. This is consistent with the fact that
705 individuals with ASD had slower RTs and lower F_1 scores across all conditions and age groups. Moreover,
706 by modelling the correlation between RTs on separate sensory channels, we demonstrated a higher (more
707 positive) channel dependency in ASD, suggesting a greater spread of attentional resources across sensory
708 modalities. Neurophysiological evidence of such disengagement comes from previous work by our lab

709 that demonstrated increased susceptibility to distraction in children with ASD compared to NT children
710 (Murphy et al., 2014). This behavioral deficit was accompanied by a reduced suppression of sensory-
711 irrelevant information, as indexed by EEG recordings of alpha-band oscillatory activity. While individuals
712 with ASD appear to utilize longer-term statistics to make predictions about their sensory environment
713 (Lieder et al., 2019), other work suggest that they tend to overestimate the volatility of their environment
714 at the expense of learning to build stable predictions (Lawson et al., 2017). In the current study, stimuli
715 were presented in a random order with equal probability, meaning there was a 66.6% chance of the same
716 unisensory input occurring on the next trial (including the AV condition). Based on these statistics, it is
717 more efficient to predict the reoccurrence of same signal (or part of it) on the next trial and to direct
718 attention therein. If these statistics are not being actively used to build predictions about the modality of
719 an upcoming stimulus, as may be the case in ASD, then the participant may be less likely to prepare for it
720 and less averse to switching sensory modality. This fits well with the notion that individuals with autism
721 rely more on bottom-up over top-down processing (Maekawa et al., 2011).

722 **Multisensory integration or modality switch effects?**

723 It is well established that MSEs systematically contribute to multisensory facilitation in a bisensory
724 detection task (Gondan et al., 2004; Van der Stoep et al., 2015a; Shaw et al., 2019). To determine the role
725 of MSEs, we performed separate tests of the race model using switch and repeat trials. While we found
726 that multisensory gain was much greater on switch trials than on repeat trials, there was still evidence of
727 race model violation on repeat trials. However, it is important to consider that in the context of a mixed
728 block design, responses on repeat trials are likely subject to residual switch effects from earlier trials ($n-2$,
729 $n-3$, etc.). Furthermore, if we consider the impact that switching modality has on RTs, a mixed block design
730 could be said to violate the assumption of context invariance. While it is unlikely that it would present the
731 opportunity to change strategy from trial to trial in a top-down manner, it is conceivable that the
732 continuously changing context (from switch to repeat conditions) could invoke disparate processing

mechanisms in a bottom-up manner (for a detailed discussion, see Shaw et al., 2019). We also measured the correlation between multisensory gain and MSEs on unisensory and multisensory trials, partialling out the effects of age. There was a strong positive correlation for unisensory (but not multisensory) stimuli, as would be expected if MSEs were to impact multisensory gain systematically. This was followed up with a mediation analyses to determine whether MSEs mediated the observed relationship between age and multisensory gain. This analysis indicated only partial mediation, suggesting that neural processes other than MSEs (e.g., cross-sensory interactions) were contributing to the observed multisensory gain. Differences in the developmental trajectories of MSEs and multisensory gain lend further support to the notion that neural processes unrelated to switching modality contribute to the RSE (Gondan et al., 2004).

Another way to examine the contribution of MSEs is to remove the presence of switch trials by using a blocked design. In another study by our lab (Shaw et al., 2019), we demonstrated that RTs to simple AV stimuli do not violate the race model when the three conditions are presented in entirely separate blocks. Comparing the median RTs between blocked and mixed conditions revealed a slowing of the unisensory but not the multisensory RTs in the mixed condition that could be largely accounted for by increased RTs on switch trials. Another study that used a block design (Otto and Mamassian, 2012) did in fact report evidence of violation, but importantly, presented AV stimuli in background noise which are more likely to recruit integrative mechanisms during bisensory detection (Wallace et al., 1996; Senkowski et al., 2011; Stevenson et al., 2012). Thus, violation of the race model may involve an interplay between integrative and switching processes that carry different weights in different contexts (mixed versus block presentations) and under different stimulus conditions (clean versus noisy). However, the race model test relies on the assumption of context invariance, such that the stimulus conditions are presented in an intermixed and unpredictable fashion (Miller, 1982; Gondan and Minakata, 2016; Miller, 2016). By interleaving the conditions, the participant does not know which stimulus to expect and presumably processes, say, an auditory signal in the same way under unisensory and multisensory conditions. Thus,

757 violation of the race model is assumed to be due to multisensory interactions rather than differences in
758 processing strategies. In contrast, when unisensory and multisensory stimuli are presented in separate
759 blocks, there may be opportunity to employ different processing strategies in order to optimize
760 performance.

761 **Modelling multisensory behavior and sensory dominance**

762 The race model provides a plausible framework for the underlying cognitive architecture implemented
763 during speeded bisensory detection (Otto and Mamassian, 2017). Theoretically, it is the most efficient
764 processing strategy in the context of a bisensory detection task, as an OR logic gate (which is the operation
765 implemented by probability summation) perfectly matches the task demands placed on the participant
766 (i.e., respond to A *or* V *or* AV). It has been demonstrated that the same architecture can be adapted to
767 meet the demands of different tasks by changing this logical coupling accordingly (Otto and Mamassian,
768 2012). This suggests that adults have the ability to flexible couple multisensory decisions according to
769 these task demands. Our data revealed that the race model under-estimated empirical benefits in the
770 majority of adults, and over-estimated empirical benefits in the majority of children (see Fig. 2C). In other
771 words, most children did not even achieve the basic level of statistical facilitation expected from
772 multisensory processing. Similarly, the race model was predictive of the variance in empirical benefits
773 across adults, but less so in children. These results suggest that the ability to flexibly couple multisensory
774 decisions according to the task demands may only be formed later in development. To gain insight into
775 the rules or strategies that might govern multisensory processing prior to the formation of this flexible
776 coupling, we tested two alternative processing strategies. We examined fits between the empirical data
777 and model behavior that was based on a parametric weighting of the race model and each alternative
778 strategy (see Methods for details). In children with ASD aged 6–12 years, model fits suggested that the
779 response to an AV stimulus was mostly triggered by the modality corresponding to the previous trial, as
780 opposed to the fastest modality. This is a less efficient processing strategy because the slower modality

could sometimes trigger a response, but explains why younger participants did not achieve statistical facilitation. In NT children aged 6–12 years, the same model offered only marginal improvements beyond probability summation, suggesting that they begin to adopt a race strategy much earlier in development than their ASD peers.

A mechanistic explanation for such sub-optimal multisensory behavior in children comes from the idea of multisensory competition; before the brain develops the ability to optimally integrate multisensory information in a way that facilitates behavior, multisensory signals are thought to compete with one another for cognitive resources, inhibiting effective processing of redundant stimuli (Sinnett et al., 2008; Cuppini et al., 2010). By applying Hebbian learning rules of potentiation and depression (Hebb, 1949), computational studies have successfully modelled this developmental transition from a default state in which multisensory inputs interact competitively to one in which they interact cooperatively (Cuppini et al., 2011; Cuppini et al., 2018). Indeed, if a competition scenario were set up, it would likely favor the most effective modality, which in our case would be the preceding modality as attention has already been directed therein. This idea of multisensory competition also provides an alternative interpretation of our modelling analysis. If the previous modality were to inhibit processing of the other modality, a decision coupling mechanism that triggers a response based on the fastest modality (i.e., race strategy) would yield the same result as one based on the previous modality (i.e., biased strategy). Thus, our data point to two possible explanatory mechanisms: 1) an early cross-sensory interaction that transitions from being inhibitory to facilitatory, 2) a later decision coupling mechanism that transitions from adopting a sensory-biased strategy to a race strategy (see Fig. 10). Given the results of our modelling analysis, it is possible that both of these integrative mechanisms have not yet matured in children with ASD aged 6–12 years, whereas their NT peers may have already developed the ability to flexibly couple multisensory decisions, but not integrate early cross-sensory inputs in a facilitatory manner. Understanding the computational principles of these mechanisms and disentangling their contribution to multisensory behavior is important

because the race model test assumes that statistical facilitation provides a baseline measure of multisensory information processing against which to gauge interaction effects (Miller, 1982; Ulrich et al., 2007). However, if statistical facilitation does not reflect the underlying mode of operation in certain populations, then how can such a model be used to obtain valid measures of multisensory integration?

Another interesting finding to emerge from our modelling analysis was that NT children aged 6–9 years appear to be biased towards the auditory modality during audiovisual processing, but thereafter become biased towards the visual modality. These results were supported by a follow-up analysis based on MSEs on AV trials, as well as previous research that has demonstrated an auditory dominance in infants and young children when they are presented with AV stimuli (Lewkowicz, 1988a, b), and a visual dominance in adults, commonly known as the Colavita visual dominance effect (Colavita, 1974). Several studies have traced the transition from an auditory to a visual dominance over the course of childhood (Robinson and Sloutsky, 2004; Nava and Pavani, 2013) and, in line with our data, suggest that this sensory reweighting occurs at around 9–10 years of age (Nava and Pavani, 2013). Sensory reweighting has also been shown to occur around 8–10 years of age for the visual and haptic modalities (Gori et al., 2008). Our modelling analysis suggests that the same trend appears to emerge in children with ASD between the ages of 6–12 years, but then reverses once more during adolescence, favoring the auditory modality in adulthood. However, our MSE analysis suggests that a visual dominance exists initially in children with ASD, only shifting to an auditory dominance in adulthood. Given the smaller sample sizes in the ASD group, it is possible that the MSE analysis may be a more reliable index of sensory dominance than our modelling analysis which relies on a correlational measure. Indeed, a visual dominance has been previously reported in children with ASD (O'Connor and Hermelin, 1965), but its transition into adulthood has not yet been documented to our knowledge. This finding may suggest that ASD individuals that display neurotypical levels of multisensory gain in adulthood may be doing so by way of an alternative sensory weighting strategy.

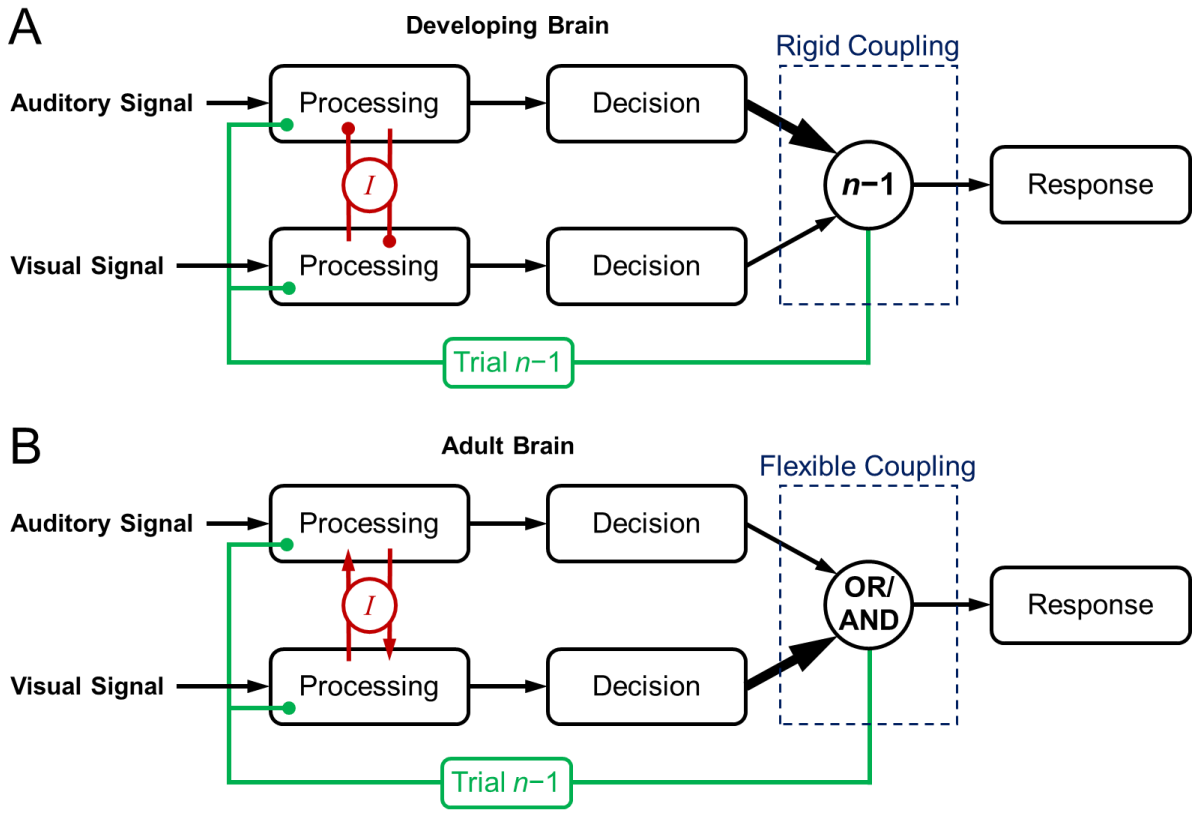


Figure 10. Cognitive architecture describing information processing for speeded bisensory detection. **A**, **B**, Based on our findings, we proposed separate models of multisensory processing in the developing (A) and adult (B) brain. Building on the race model architecture, multisensory information is processed on separate channels but can interact in a cross-sensory manner and is subject to history effects (trial *n*−1). Separate unisensory decisions are made on each channel and are coupled according to modality/context-dependent rules in order to from a multisensory decision. An inherent sensory dominance biases the probability of implementing a processing strategy based on modality versus context. During the early stages of maturation, multisensory signals compete for resources through inhibitory cross-sensory projections. Coupling of unisensory decisions is rigid, with responses being triggered by the preceding modality as opposed to the fastest (but note that the previous modality may typically be the fastest due to earlier inhibitory interactions). Either way, this results in less effective processing of multisensory information. In adulthood, individuals have developed the ability to integrate multisensory information in

a way that facilitates information processing as well as the ability to flexibly couple unisensory decisions according to specific task demands. Sensory dominance transitions from auditory to visual (Colavita effect) during maturation but may be weighted differently in individuals with ASD.

Neural mechanisms underlying impaired multisensory behavior in autism

Prior work by our lab suggests that the neural processes underlying multisensory integration are impaired in children with autism (Brandwein et al., 2013). Specifically, we found that EEG correlates of integration were weaker (of lower amplitude) and occurred later in the information processing hierarchy. Neural indices of integration over parieto-occipital scalp between 140–160 ms were predictive of race model violation in NT children but not in children with ASD. Using the same paradigm, we recorded intracranial electrophysiology in adults with epilepsy and demonstrated that visual stimulation influenced the phase of ongoing oscillations in auditory cortex (Mercier et al., 2015), and auditory stimulation influenced the phase of ongoing oscillations in visual cortex (Mercier et al., 2013), such that cross-sensory stimulation appears to prime ancillary sensory cortices to make them more receptive to their primary sensory input. The response to the primary sensory input (e.g., visual stimulation of visual cortex) is then enhanced for multisensory trials (Mercier et al., 2013), at least in a bisensory detection task such as the current one. Furthermore, neuro-oscillatory phase alignment across the sensorimotor network was significantly enhanced by multisensory stimulation, and was related to the speed of a response (Mercier et al., 2015). Such phase resetting of ongoing neural oscillations by functionally distinct and distant neuronal ensembles is thought to be fundamental to multisensory integration (Lakatos et al., 2007; Schroeder et al., 2008; Fiebelkorn et al., 2011; Fiebelkorn et al., 2013). Impaired cross-sensory phase-resetting, as might be predicted by reduced subcortical and cortical connectivity, would likely result in impaired integrative abilities. In autism, there is evidence for such disrupted connectivity (Zeng et al., 2017; Arnold Anteraper et al., 2018), although these findings are mixed and somewhat inconclusive (Vasa et al., 2016). Nevertheless, disrupted connectivity could in turn lead to impaired cross-sensory phase-resetting and

hence contribute to impaired multisensory processing in ASD. As we already mentioned, weaker cross-sensory inhibition might account for reduced MSEs in ASD (Murphy et al., 2014), possibly also due to poorer brain connectivity. In contrast, it is possible that cross-sensory connectivity in children with ASD is fully intact, but that integration of multisensory information has not yet transitioned from a state of completion, to one of facilitation, as discussed earlier (Cuppini et al., 2011; Cuppini et al., 2018).

Conclusions

From the current analyses we can draw several conclusions. 1) When assessed using the race model test, multisensory processing in individuals with ASD "normalizes" by the mid-twenties. 2) In younger children, including those with ASD, statistical facilitation does not appear to reflect the underlying mode of operation as it does in adults. This could be caused by early cross-sensory inhibition and/or ineffective coupling of multisensory decision processes. 3) Differences in both MSEs and patterns in sensory dominance indicate fundamental alterations in how the nervous system of children with ASD respond to even the simplest of multisensory environments. 4) Greater channel dependency in ASD suggest a more even spread of attention, possibly due to an aversion to making predictions based on short-term statistics or an impairment in cross-sensory inhibition. The current findings also make clear that there is significant work ahead of us before we truly understand developmental shifts in multisensory integration, decision coupling, sensory weighting and modality switch effects, how these differ in children with ASD and each of their roles within the context of speeded bisensory detection. Individual variance and group differences in such patterns undoubtedly contribute to how the sensory environment is experienced at a given stage of development. Here we set the stage for detailed characterization of these processes and their interactions, to in turn understand potential roadblocks to the typical development of multisensory processing in ASD, and some of the factors that might contribute to sensory reactivity in this group.

Author Contributions

S.M. and J.J.F. designed the original experiment. M.J.C., S.M. and J.J.F. conceived of the current study. M.J.C. analyzed the data and produced all illustrations in consultation with S.M. and J.J.F. M.J.C. wrote the first substantive draft of manuscript. S.M. and J.J.F. provided editorial input to M.J.C. on multiple subsequent drafts.

Acknowledgements

The authors thank Dr. Juliana Bates who administered the majority of the clinical assessments, Dr. Natalie Russo for her contributions to the clinical phenotyping of a portion of our ASD cohort, Dr. Cristiano Cuppini for his valuable discussions, Dr. Ana Francisco for her support with statistical analyses, Douwe Horsthuis, Elise Tavern and Danielle DeMaio for their help with data collection and Dr. Alice Brandwein who conducted the previous studies for which most of these data were collected. We also extend our heartfelt gratitude to the participants and their families that have contributed their time to participate in this research. This work was supported in large part by an RO1 grant from the National Institute of Mental Health (MH085322 to S.M. and J.J.F.) and a U54 Grant (HD090260) to the Human Clinical Phenotyping Core of the Rose F. Kennedy Intellectual and Developmental Disabilities Research Center.

Competing Financial Interests

The authors declare no competing financial interests.

References

APA D-APA (2000) Diagnostic and statistical manual, 4th edn, Text Revision (DSM-IV-TR). American Psychiatric Association, Washington.

Arnold Anteraper S, Guell X, D'Mello A, Joshi N, Whitfield-Gabrieli S, Joshi G (2018) Disrupted Cerebro-cerebellar Intrinsic Functional Connectivity in Young Adults with High-functioning Autism Spectrum Disorder: A Data-driven, Whole-brain, High Temporal Resolution fMRI Study. Brain connectivity.

Ashby FG, Townsend JT (1986) Varieties of perceptual independence. Psychol Rev 93:154.

Ayres AJ, Tickle LS (1980) Hyper-responsivity to touch and vestibular stimuli as a predictor of positive response to sensory integration procedures by autistic children. Am J Occup Ther 34:375-381.

915 Baron RM, Kenny DA (1986) The moderator–mediator variable distinction in social psychological research:
916 Conceptual, strategic, and statistical considerations. J Pers Soc Psychol 51:1173.
917 Barr DJ, Levy R, Scheepers C, Tily HJ (2013) Random effects structure for confirmatory hypothesis testing:
918 Keep it maximal. Journal of memory and language 68:255-278.
919 Bebko JM, Schroeder JH, Weiss JA (2014) The McGurk effect in children with autism and Asperger
920 syndrome. Autism Res 7:50-59.
921 Beker S, Foxe JJ, Molholm S (2017) Ripe for solution: Delayed development of multisensory processing in
922 autism and its remediation. Neurosci Biobehav Rev.
923 Bishara AJ, Hittner JB (2012) Testing the significance of a correlation with nonnormal data: comparison of
924 Pearson, Spearman, transformation, and resampling approaches. Psychol Methods 17:399.
925 Blair RC, Higgins JJ, Karniski W, Kromrey JD (1994) A study of multivariate permutation tests which may
926 replace Hotelling's T2 test in prescribed circumstances. Multivariate Behavioral Research 29:141-
927 163.
928 Bland JM, Altman DG (1995) Calculating correlation coefficients with repeated observations: part 2—
929 correlation between subjects. BMJ 310:633.
930 Bliss CI (1967) Statistics in biology. New York, NY: McGraw-Hill.
931 Brandwein AB, Foxe JJ, Russo NN, Altschuler TS, Gomes H, Molholm S (2011) The development of
932 audiovisual multisensory integration across childhood and early adolescence: a high-density
933 electrical mapping study. Cerebral Cortex 21:1042-1055.
934 Brandwein AB, Foxe JJ, Butler JS, Russo NN, Altschuler TS, Gomes H, Molholm S (2013) The development
935 of multisensory integration in high-functioning autism: high-density electrical mapping and
936 psychophysical measures reveal impairments in the processing of audiovisual inputs. Cerebral
937 Cortex 23.
938 Brandwein AB, Foxe JJ, Butler JS, Frey H-P, Bates JC, Shulman LH, Molholm S (2015) Neurophysiological
939 indices of atypical auditory processing and multisensory integration are associated with symptom
940 severity in autism. J Autism Dev Disord 45:230-244.
941 Chandrasekaran C (2017) Computational principles and models of multisensory integration. Curr Opin
942 Neurobiol 43:25-34.
943 Colavita FB (1974) Human sensory dominance. Percept Psychophys 16:409-412.
944 Colonius H (1986) Measuring channel dependence in separate activation models. Percept Psychophys
945 40:251-255.
946 Colonius H (1990) Possibly dependent probability summation of reaction time. J Math Psychol 34:253-
947 275.
948 Colonius H, Diederich A (2006) The race model inequality: interpreting a geometric measure of the amount
949 of violation. Psychol Rev 113:148–154.
950 Crosse MJ, Di Liberto GM, Lalor EC (2016) Eye can hear clearly now: inverse effectiveness in natural
951 audiovisual speech processing relies on long-term crossmodal temporal integration. J Neurosci
952 36:9888-9895.
953 Cuppini C, Stein BE, Rowland BA (2018) Development of the mechanisms governing midbrain multisensory
954 integration. J Neurosci:2631-2617.
955 Cuppini C, Ursino M, Magosso E, Rowland BA, Stein BE (2010) An emergent model of multisensory
956 integration in superior colliculus neurons. Front Integr Neurosci 4:6.
957 Cuppini C, Stein BE, Rowland BA, Magosso E, Ursino M (2011) A computational study of multisensory
958 maturation in the superior colliculus (SC). Exp Brain Res 213:341-349.
959 Cuppini C, Ursino M, Magosso E, Ross LA, Foxe JJ, Molholm S (2017) A Computational Analysis of Neural
960 Mechanisms Underlying the Maturation of Multisensory Speech Integration in Neurotypical
961 Children and Those on the Autism Spectrum. Front Hum Neurosci 11:518.
962 Davison AC, Hinkley DV (1997) Bootstrap methods and their application: Cambridge university press.

963 de Gelder B, Vroomen J, Van der Heide L (1991) Face recognition and lip-reading in autism. Eur J Cogn
964 Psychol 3:69-86.
965 Diederich A, Colonius H (1987) Intersensory facilitation in the motor component? Psychol Res 49:23-29.
966 Fiebelkorn IC, Foxe JJ, Butler JS, Mercier MR, Snyder AC, Molholm S (2011) Ready, set, reset: stimulus-
967 locked periodicity in behavioral performance demonstrates the consequences of cross-sensory
968 phase reset. J Neurosci 31:9971-9981.
969 Fiebelkorn IC, Snyder AC, Mercier MR, Butler JS, Molholm S, Foxe JJ (2013) Cortical cross-frequency
970 coupling predicts perceptual outcomes. Neuroimage 69:126-137.
971 Foxe JJ, Molholm S (2009) Ten years at the multisensory forum: musings on the evolution of a field. Brain
972 Topogr 21:149-154.
973 Foxe JJ, Molholm S, Del Bene VA, Frey H-P, Russo NN, Blanco D, Saint-Amour D, Ross LA (2015) Severe
974 multisensory speech integration deficits in high-functioning school-aged children with autism
975 spectrum disorder (ASD) and their resolution during early adolescence. Cerebral Cortex 25:298-
976 312.
977 Gielen SC, Schmidt RA, Van Den Heuvel PJ (1983) On the nature of intersensory facilitation of reaction
978 time. Percept Psychophys 34:161-168.
979 Gondan M (2010) A permutation test for the race model inequality. Behav Res Methods 42:23-28.
980 Gondan M, Minakata K (2016) A tutorial on testing the race model inequality. Attention, Perception, &
981 Psychophysics 78:723-735.
982 Gondan M, Lange K, Rösler F, Röder B (2004) The redundant target effect is affected by modality switch
983 costs. Psychonomic bulletin & review 11:307-313.
984 Gori M, Del Viva M, Sandini G, Burr DC (2008) Young children do not integrate visual and haptic form
985 information. Curr Biol 18:694-698.
986 Grice GR, Canham L, Gwynne JW (1984) Absence of a redundant-signals effect in a reaction time task with
987 divided attention. Percept Psychophys 36:565-570.
988 Harrington L, Peck C (1998) Spatial disparity affects visual-auditory interactions in human sensorimotor
989 processing. Exp Brain Res 122:247-252.
990 Hebb DO (1949) The organization of behavior. In: Wiley, New York Reprint Lawrence Erlbaum.
991 Hedges LV, Olkin I (1985) Statistical Methods for Meta Analysis. Orlando, FL: Academic Press.
992 Hershenson M (1962) Reaction time as a measure of intersensory facilitation. J Exp Psychol 63:289.
993 Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nature Reviews Neuroscience
994 8:393-402.
995 Hughes HC, Nelson MD, Aronchick DM (1998) Spatial characteristics of visual-auditory summation in
996 human saccades. Vision Res 38:3955-3963.
997 Iarocci G, McDonald J (2006) Sensory integration and the perceptual experience of persons with autism. J
998 Autism Dev Disord 36:77-90.
999 Innes BR, Otto TU (2019) A comparative analysis of response times shows that multisensory benefits and
1000 interactions are not equivalent. Sci Rep 9.
1001 Irwin JR, Tornatore LA, Brancazio L, Whalen D (2011) Can children with autism spectrum disorders “hear”
1002 a speaking face? Child Dev 82:1397-1403.
1003 Kiesel A, Miller J, Ulrich R (2007) Systematic biases and Type I error accumulation in tests of the race model
1004 inequality. Behav Res Methods 39:539-551.
1005 Kinchla R (1974) Detecting target elements in multielement arrays: A confusability model. Percept
1006 Psychophys 15:149-158.
1007 Krueger Fister J, Stevenson RA, Nidiffer AR, Barnett ZP, Wallace MT (2016) Stimulus intensity modulates
1008 multisensory temporal processing. Neuropsychologia 88:92-100.
1009 Lakatos P, Chen C-M, O'Connell MN, Mills A, Schroeder CE (2007) Neuronal oscillations and multisensory
1010 interaction in primary auditory cortex. Neuron 53:279-292.

Lawson RP, Mathys C, Rees G (2017) Adults with autism overestimate the volatility of the sensory environment. Nat Neurosci 20:1293.

Lewkowicz DJ (1988a) Sensory dominance in infants: I. Six-month-old infants' response to auditory-visual compounds. Dev Psychol 24:155.

Lewkowicz DJ (1988b) Sensory dominance in infants: II. Ten-month-old infants' response to auditory-visual compounds. Dev Psychol 24:172.

Lieder I, Adam V, Frenkel O, Jaffe-Dax S, Sahani M, Ahissar M (2019) Perceptual bias reveals slow-updating in autism and fast-forgetting in dyslexia. Nat Neurosci:1.

Lord C, Rutter M, Le Couteur A (1994) Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 24:659-685.

Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, Pickles A, 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.

Luce RD (1986) Response times: Their role in inferring elementary mental organization: Oxford University Press on Demand.

Maekawa T, Tobimatsu S, Inada N, Oribe N, Onitsuka T, Kanba S, Kamio Y (2011) Top-down and bottom-up visual information processing of non-social stimuli in high-functioning autism spectrum disorder. Res Autism Spectr Disord 5:201-209.

Mahoney JR, Molholm S, Butler JS, Sehatpour P, Gomez-Ramirez M, Ritter W, Foxe JJ (2015) Keeping in touch with the visual system: spatial alignment and multisensory integration of visual-somatosensory inputs. Front Psychol 6:1068.

Martineau J, Roux S, Adrien J, Garreau B, Barthélémy C, Lelord G (1992) Electrophysiological evidence of different abilities to form cross-modal associations in children with autistic behavior. Electroencephalogr Clin Neurophysiol 82:60-66.

Megevand P, Mercier MR, Groppe DM, Golumbic EZ, Mesgarani N, Beauchamp MS, Schroeder CE, Mehta AD (2018) Phase resetting in human auditory cortex to visual speech. bioRxiv:405597.

Mégevand P, Molholm S, Nayak A, Foxe JJ (2013) Recalibration of the multisensory temporal window of integration results from changing task demands. PLoS One 8:e71608.

Meijers L, Eijkman E (1977) Distributions of simple RT with single and double stimuli. Percept Psychophys 22:41-48.

Mercier MR, Foxe JJ, Fiebelkorn IC, Butler JS, Schwartz TH, Molholm S (2013) Auditory-driven phase reset in visual cortex: human electrocorticography reveals mechanisms of early multisensory integration. Neuroimage 79:19-29.

Mercier MR, Molholm S, Fiebelkorn IC, Butler JS, Schwartz TH, Foxe JJ (2015) Neuro-Oscillatory Phase Alignment Drives Speeded Multisensory Response Times: An Electro-Corticographic Investigation. The Journal of Neuroscience 35:8546-8557.

Miller J (1978) Multidimensional same–different judgments: Evidence against independent comparisons of dimensions. J Exp Psychol Hum Percept Perform 4:411.

Miller J (1982) Divided attention: Evidence for coactivation with redundant signals. Cogn Psychol 14:247-279.

Miller J (1986) Timecourse of coactivation in bimodal divided attention. Percept Psychophys 40:331-343.

Miller J (1988) A warning about median reaction time. J Exp Psychol Hum Percept Perform 14:539.

Miller J (1991) Reaction time analysis with outlier exclusion: Bias varies with sample size. The quarterly journal of experimental psychology 43:907-912.

Miller J (2016) Statistical facilitation and the redundant signals effect: What are race and coactivation models? Attention, Perception, & Psychophysics 78:516-519.

1058 Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ (2002) Multisensory auditory-visual
1059 interactions during early sensory processing in humans: a high-density electrical mapping study.
1060 Cognitive Brain Research 14:115-128.
1061 Mordkoff JT, Yantis S (1991) An interactive race model of divided attention. J Exp Psychol Hum Percept
1062 Perform 17:520.
1063 Murphy JW, Foxe JJ, Peters JB, Molholm S (2014) Susceptibility to distraction in autism spectrum disorder:
1064 Probing the integrity of oscillatory alpha-band suppression mechanisms. Autism Res 7:442-458.
1065 Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Schroeder CE, Foxe JJ (2004)
1066 Grabbing your ear: rapid auditory–somatosensory multisensory interactions in low-level sensory
1067 cortices are not constrained by stimulus alignment. Cerebral cortex 15:963-974.
1068 Nava E, Pavani F (2013) Changes in sensory dominance during childhood: Converging evidence from the
1069 Colavita effect and the sound-induced flash illusion. Child Dev 84:604-616.
1070 Nidiffer AR, Stevenson RA, Fister JK, Barnett ZP, Wallace MT (2016) Interactions between space and
1071 effectiveness in human multisensory performance. Neuropsychologia 88:83-91.
1072 Noorani I, Carpenter R (2016) The LATER model of reaction time and decision. Neurosci Biobehav Rev
1073 64:229-251.
1074 Nozawa G, Reuter-Lorenz PA, Hughes H (1994) Parallel and serial processes in the human oculomotor
1075 system: bimodal integration and express saccades. Biol Cybern 72:19-34.
1076 O'Connor N, Hermelin B (1965) Sensory dominance: In autistic imbecile children and controls. Arch Gen
1077 Psychiatry 12:99-103.
1078 Otto TU (2018) An analysis and modelling toolbox to study multisensory response times. In: International
1079 Multisensory Research Forum. Toronto, Canada.
1080 Otto TU, Mamassian P (2012) Noise and correlations in parallel perceptual decision making. Curr Biol
1081 22:1391-1396.
1082 Otto TU, Mamassian P (2017) Multisensory decisions: the test of a race model, its logic, and power.
1083 Multisensory Research 30:1-24.
1084 Otto TU, Dassy B, Mamassian P (2013) Principles of multisensory behavior. J Neurosci 33:7463-7474.
1085 Peelle JE (2019) The neural basis for auditory and audiovisual speech perception. PsyArXiv.
1086 Peelle JE, Sommers MS (2015) Prediction and constraint in audiovisual speech perception. Cortex 68:169-
1087 181.
1088 Raab DH (1962) Division of psychology: Statistical facilitation of simple reaction times. Trans N Y Acad Sci
1089 24:574-590.
1090 Ratcliff R (1979) Group reaction time distributions and an analysis of distribution statistics. Psychol Bull
1091 86:446.
1092 Ratcliff R (1993) Methods for dealing with reaction time outliers. Psychol Bull 114:510.
1093 Robinson CW, Sloutsky VM (2004) Auditory dominance and its change in the course of development. Child
1094 Dev 75:1387-1401.
1095 Ross LA, Saint-Amour D, Leavitt VM, Javitt DC, Foxe JJ (2007) Do you see what I am saying? Exploring visual
1096 enhancement of speech comprehension in noisy environment. Cerebral Cortex 17:1147-1153.
1097 Ross LA, Molholm S, Blanco D, Gomez-Ramirez M, Saint-Amour D, Foxe JJ (2011) The development of
1098 multisensory speech perception continues into the late childhood years. Eur J Neurosci 33:2329-
1099 2337.
1100 Schroeder CE, Lakatos P, Kajikawa Y, Partan S, Puce A (2008) Neuronal oscillations and visual amplification
1101 of speech. Trends Cogn Sci 12:106-113.
1102 Senkowski D, Saint-Amour D, Höfle M, Foxe JJ (2011) Multisensory interactions in early evoked brain
1103 activity follow the principle of inverse effectiveness. Neuroimage 56:2200-2208.
1104 Shaw L, Crosse MJ, Nicholas E, Muhan Chen A, Braiman M, Shoari A, Freedman EG, Molholm S, Foxe JJ
1105 (2019) Evidence for Modality Switch Costs in Redundant Signals Effect Paradigms.

Silverman LB, Bennetto L, Campana E, Tanenhaus MK (2010) Speech-and-gesture integration in high functioning autism. Cognition 115:380-393.

Sinnett S, Soto-Faraco S, Spence C (2008) The co-occurrence of multisensory competition and facilitation. Acta Psychol (Amst) 128:153-161.

Smith EG, Bennetto L (2007) Audiovisual speech integration and lipreading in autism. Journal of Child Psychology and Psychiatry 48:813-821.

Spence C, Nicholls ME, Driver J (2001) The cost of expecting events in the wrong sensory modality. Percept Psychophys 63:330-336.

Stano J (1999) Wechsler abbreviated scale of intelligence. San Antonio, TX: The Psychological Corporation.

Stein BE, Stanford TR, Rowland BA (2014) Development of multisensory integration from the perspective of the individual neuron. Nature Reviews Neuroscience 15:520-535.

Stevenson RA, Bushmakin M, Kim S, Wallace MT, Puce A, James TW (2012) Inverse effectiveness and multisensory interactions in visual event-related potentials with audiovisual speech. Brain Topogr 25:308-326.

Stevenson RA, Segers M, Ncube BL, Black KR, Bebko JM, Ferber S, Barense MD (2017) The cascading influence of multisensory processing on speech perception in autism. Autism:1362361317704413.

Sumby WH, Pollack I (1954) Visual contribution to speech intelligibility in noise. J Acoust Soc Am 26:212-215.

Sutton S, Hakerem G, Zubin J, Portnoy M (1961) The effect of shift of sensory modality on serial reaction-time: A comparison of schizophrenics and normals. The American Journal of Psychology 74:224-232.

Taylor N, Isaac C, Milne E (2010) A comparison of the development of audiovisual integration in children with autism spectrum disorders and typically developing children. J Autism Dev Disord 40:1403-1411.

Todd JW (1912) Reaction to multiple stimuli: Science Press.

Townsend JT, Wenger MJ (2004) A theory of interactive parallel processing: new capacity measures and predictions for a response time inequality series. Psychol Rev 111:1003.

Tye-Murray N, Sommers M, Spehar B (2007) Auditory and visual lexical neighborhoods in audiovisual speech perception. Trends in Amplification 11:233-241.

Ulrich R, Miller J (1994) Effects of truncation on reaction time analysis. J Exp Psychol Gen 123:34.

Ulrich R, Miller J, Schröter H (2007) Testing the race model inequality: an algorithm and computer programs. Behav Res Methods 39:291-302.

Van der Stoep N, Van der Stigchel S, Nijboer T (2015a) Exogenous spatial attention decreases audiovisual integration. Attention, perception, & psychophysics 77:464-482.

Van der Stoep N, Spence C, Nijboer T, Van der Stigchel S (2015b) On the relative contributions of multisensory integration and crossmodal exogenous spatial attention to multisensory response enhancement. Acta Psychol (Amst) 162:20-28.

Van Rijsbergen CJ (1979) Information Retrieval: Butterworth-Heinemann.

Vasa RA, Mostofsky SH, Ewen JB (2016) The disrupted connectivity hypothesis of autism spectrum disorders: time for the next phase in research. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging 1:245-252.

Vincent SB (1912) The function of the viborissae in the behavior of the white rat (Animal Behavior Monographs Vol. 1, No. 5). New York, NY: H Holt.

Wager TD, Davidson ML, Hughes BL, Lindquist MA, Ochsner KN (2008) Prefrontal-subcortical pathways mediating successful emotion regulation. Neuron 59:1037-1050.

Wallace MT, Stein BE (2007) Early experience determines how the senses will interact. J Neurophysiol 97:921-926.

1154 Wallace MT, Wilkinson LK, Stein BE (1996) Representation and integration of multiple sensory inputs in
1155 primate superior colliculus. J Neurophysiol 76:1246-1266.
1156 Wallace MT, Perrault TJ, Hairston WD, Stein BE (2004) Visual experience is necessary for the development
1157 of multisensory integration. J Neurosci 24:9580-9584.
1158 Westfall PH, Young SS (1993) Resampling-Based Multiple Testing: Examples and Methods for p-Value
1159 Adjustment (Wiley Series in Probability and Statistics).
1160 Williams DL, Goldstein G, Minshew NJ (2013) The modality shift experiment in adults and children with
1161 high functioning autism. J Autism Dev Disord 43:794-806.
1162 Wundt WM (1893) Grundzüge der physiologischen Psychologie, 4 Edition. Leipzig: W. Engelmann.
1163 Zeng K, Kang J, Ouyang G, Li J, Han J, Wang Y, Sokhadze EM, Casanova MF, Li X (2017) Disrupted Brain
1164 Network in Children with Autism Spectrum Disorder. Sci Rep 7:16253.
1165 Zubin J (1975) Problem of attention in schizophrenia. Experimental approaches to psychopathology:139-
1166 166.

1167