

Atypical Rapid Audio-Visual Temporal Recalibration in Autism Spectrum Disorders

Jean-Paul Noel, Matthew A. De Nier, Ryan Stevenson, David Alais, and Mark T. Wallace

SCIENTIFIC ABSTRACT

Changes in sensory and multisensory function are increasingly recognized as a common phenotypic characteristic of Autism Spectrum Disorders (ASD). Furthermore, much recent evidence suggests that sensory disturbances likely play an important role in contributing to social communication weaknesses—one of the core diagnostic features of ASD. An established sensory disturbance observed in ASD is reduced audiovisual temporal acuity. In the current study, we substantially extend these explorations of multisensory temporal function within the framework that an inability to rapidly recalibrate to changes in audiovisual temporal relations may play an important and under-recognized role in ASD. In the paradigm, we present ASD and typically developing (TD) children and adolescents with asynchronous audiovisual stimuli of varying levels of complexity and ask them to perform a simultaneity judgment (SJ). In the critical analysis, we test audiovisual temporal processing on trial t as a condition of trial $t - 1$. The results demonstrate that individuals with ASD fail to rapidly recalibrate to audiovisual asynchronies in an equivalent manner to their TD counterparts for simple and non-linguistic stimuli (i.e., flashes and beeps, hand-held tools), but exhibit comparable rapid recalibration for speech stimuli. These results are discussed in terms of prior work showing a speech-specific deficit in audiovisual temporal function in ASD, and in light of current theories of autism focusing on sensory noise and stability of perceptual representations.

LAY ABSTRACT

The integration of information across the different sensory modalities constitutes a fundamental step toward building a cohesive and comprehensive perceptual representation of the world. This integration and perceptual “binding” is highly dependent on the temporal structure of the multisensory cues. In ASD, multisensory temporal acuity has been found to be impaired, most notably for the integration of audiovisual speech stimuli, a finding that is confirmed in the current study. In addition, we show a striking difference in how those with ASD recalibrate their audiovisual temporal judgments based on prior trial history relative to those who are typically-developing. Most notable is the finding that whereas recalibration for speech stimuli fails to differ between ASD and TD participants, those with ASD fail to recalibrate when making judgments concerning non-speech audiovisual stimuli. These results not only expand our understanding of multisensory temporal function in ASD, but also have important implications for models suggesting changes in predictive coding and sensory priors in autism. *Autism Res* 2016, 00: 000–000. © 2016 International Society for Autism Research, Wiley Periodicals, Inc.

Keywords: autism; multisensory; audio-visual; plasticity; recalibration

Introduction

Autism Spectrum Disorder (ASD) is defined by core deficits in social communication and social interaction, as well as repetitive patterns of behavior and/or interests. In addition, the presence of atypical sensory processing is reported to be characteristic of a majority of those suffering from ASD [Marco et al., 2011] and has recently been included as a diagnostic feature for ASD in the

DSM-5 [American Psychiatric Association, 2013]. A number of recent reports suggest that those with ASD exhibit deficits in performance on tasks requiring the utilization of sensory information from multiple sensory modalities [Brandwein et al., 2013; Smith & Bennetto, 2007].

Impaired temporal acuity for audiovisual stimuli, most notably for audiovisual speech stimuli, is a prominently documented deficit of multisensory processing

From the Neuroscience Graduate Program, Vanderbilt Brain Institute, Vanderbilt University Medical School, Vanderbilt University, Nashville, Tennessee (J.-P.N.); Vanderbilt Brain Institute, Vanderbilt University Medical School, Vanderbilt University, Nashville, Tennessee (J.-P.N., M.N., R.S., M.T.W.); Medical Scientist Training Program, Vanderbilt University Medical School, Vanderbilt University, Nashville, Tennessee (M.N.); Department of Psychology, University of Toronto, Toronto, Ontario, Canada (R.S.); Department of Psychology, University of Western Ontario, London, Ontario, Canada (R.S.); Brain and Mind Institute, University of Western Ontario, London, Ontario, Canada (R.S.); School of Psychology, University of Sydney, Sydney, Australia (D.A.); Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, Tennessee (M.T.W.); Department of Psychology, Vanderbilt University, Nashville, Tennessee (M.T.W.)

Received January 13, 2016; accepted for publication March 24, 2016

Address for correspondence and reprints: Mark Wallace, Ph.D., 7110 MRB III BioSci Bldg, 465, 21st Ave South, Nashville, TN 3721. E-mail: mark.wallace@vanderbilt.edu

Published online 00 Month 2016 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/aur.1633

© 2016 International Society for Autism Research, Wiley Periodicals, Inc.

in ASD [Bebko, Weiss, Demark, & Gomez, 2006; Foss-Feig et al., 2010; Kwakye, Foss-Feig, Cascio, Stone, & Wallace, 2011; Stevenson et al., 2014; for review see Stevenson et al., 2014]. The temporal proximity of sensory signals is an important factor in determining whether the separate sensory cues should be “bound” into a singular entity, as events and objects occurring closer in time are more likely to have a common origin [Noel, Wallace, Orchard-Mills, Alais, & Van der Burg, 2015; Stevenson, Zemtsov, & Wallace, 2012]. In ASD, an impaired ability to properly integrate multisensory temporal information may lead to an inability to accurately represent the environment and derive the behavioral benefits normally arising from multisensory integration [Wallace & Stevenson 2014]. For example, in speech perception, integrating vision of a speaker’s lip and mouth movements with the auditory speech signal improves speech comprehension in noisy environments [Sumby & Pollack 1954]. If integration of auditory and visual speech signals is compromised, as seen in ASD [Foss-Feig et al., 2010; Stevenson et al., 2014], then these changes in multisensory function may cascade into deficits in speech comprehension and delays in language development [Stevenson, Segers, Ferber, Barense, & Wallace, 2014].

Complicating temporal integration is the fact that the relative timing of auditory and visual stimuli changes with source distance [Noel, Lukowska, Wallace, & Serino, 2016; Pöppel, 1985; Sugita & Suzuki, 2003]. To deal with this, multisensory temporal function is highly dynamic and can recalibrate adaptively to the changing statistics of the sensory input [for review see Vroomen & Keetels, 2010]. This makes sense in a world where the spatial and temporal relationships between auditory stimuli are constantly changing yet still bring perceptual benefits when integrated into a coherent percept. Initial investigations of temporal recalibration used simultaneity judgment (SJ) tasks with variously asynchronous stimuli to demonstrate that following extensive exposure to asynchronous stimuli, the point of subjective simultaneity (PSS—the asynchrony at which two stimuli are most likely to be judged as simultaneous) shifts toward the repeatedly presented asynchrony [Fujisaki, Shimojo, Kashino, & Nishida, 2004; Vroomen, Keetels, De Gelder, & Bertelson, 2004]. Although these initial studies used long exposures, it has recently been shown that temporal recalibration can occur rapidly following a single asynchronous event [Van der Burg, Alais, & Cass, 2013]. This form of rapid temporal recalibration has been demonstrated for both low-level audiovisual stimuli [Van der Burg et al., 2013] and for perceptually complex audiovisual stimuli such as speech [Van der Burg & Goodbourn, 2015]. In addition, Van der Burg et al. [2013], report a positive relationship between the size of one’s temporal binding window (TBW; a measure of multisensory temporal acuity) and

the degree to which one will incorporate recent sensory history into a perceptual representation. Extending this to ASD, if those with autism have larger TBWs (Stevenson et al., 2014), they may be expected to show a greater tendency toward rapid temporal recalibration.

Such a postulation, that of an altered ability to rapidly (e.g., on a trial-to-trial basis) update and adapt sensory representations, is not only supported by the fact that individuals with ASD have larger TBWs, but also by recent empirical and computational work. Specifically, Zaidel, Goin-Kochel, & Angelaki [2015], have provided evidence that individuals with ASD have attenuated and inflexible Bayesian priors (i.e., sensory expectations) and instead rely more heavily on the actual incoming sensory information (i.e., likelihoods; see Rosenberg, Patterson, & Angelaki [2015] for a discussion on the possible neural computation driving such effect). These individuals, thus, may more faithfully represent the immediate external world and are conceivably less influenced by previous knowledge of the world embodied in a prior. A hypo-prior account of autism would postulate a reduced ability to rapidly adapt to the sensory statistics of the world [Pellicano & Burr, 2012], which we will test here using rapid adaptation to audiovisual asynchrony.

In this study, we postulate that there may be a relationship between poor audiovisual temporal acuity in ASD and the ability to rapidly recalibrate to the statistics of ongoing multisensory stimuli. Indeed, differences in such short-term plasticity (the ability to rapidly adapt to changing sensory statistics of the environment) may play an integral role in the construction of longer-term (i.e., more cumulative) representations of audiovisual temporal acuity. We examine rapid audiovisual temporal recalibration in ASD children and adolescents relative to a matched typically developing (TD) cohort, comparing three kinds of stimuli: simple (flash/beep), complex non-speech and complex speech stimuli. These different tasks are employed in an effort to understand the coding and representational levels at which deficits emerge. One prediction of the Bayesian hypo-prior account of ASD is that rapid recalibration will be diminished across all levels of stimulus complexity, but perhaps more so for more complex stimuli as ASD individuals demonstrate a heightened sensitivity to noise [Rosenberg, Patterson, & Angelaki, 2015]. In contrast, previous work showing changes in multisensory temporal acuity in ASD [Van der Burg et al., 2013] would predict more (and not less) rapid recalibration, specifically for speech stimuli.

Methods and Materials

Participants

Participants were 26 TD (12 males) and 26 ASD (24 males) age-matched individuals (TD, $M = 11.6$ years old,

SD = 3.79, range = 8–17; ASD = 12.3, SD = 3.05, range = 7–17). Participants in the ASD group had been previously diagnosed with ASD from a clinician practitioner according to the diagnostic criteria of the DSM-5 and diagnosis was confirmed by research-reliable clinicians using the Autism Diagnosis Observation Schedule (ADOS) and/or Autism Diagnostic Interview-Revised (ADI-R). Raw TBW data from seven of the ASD participants has been reported in a prior paper by our group [Stevenson et al., 2014]. Exclusion of these participants from analysis did not alter any of the below-mentioned statistical effects, and thus these data were included in the report. Individuals in the TD group did not have a diagnosis of ASD or any other psychiatric disorder. Participants in the ASD and TD group were matched for IQ (Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-2), TD = 111.52, SD = 14.73; ASD, M = 110.23, SD = 14.05), and although they were not for gender, statistical analyses revealed no differences between genders in any of the measures reported below (all $P > 0.42$). Groups (ASD vs. TD) also differed on the vocabulary subtest of the WASE-2 (T-scores: TD, M = 60.73, SD = 7.80; ASD, M = 54.33, SD = 10.08; $t(50) = 2.57$, $P = 0.013$). All participants had self-reported normal visual and auditory acuity. Vanderbilt University Medical Center's Institutional Review Board approved all experimental protocols, and written informed consent was obtained from all participants.

Material and apparatus

Three distinct categories of audiovisual stimuli were presented: flash-beeps (simple), dynamic handheld tools (complex non-speech stimuli), and single syllable utterances (speech stimuli). With regard to the flash-beep stimuli, the visual component consisted of a white ring surrounding a white fixation cross on a black background, and the ring subtended 17.3° of visual angle. Visual stimulus duration was 10 ms. The auditory stimulus was a 3500 Hz pure tone with a duration of 13 ms. In terms of the dynamic handheld tools, the visual stimulus consisted of a full cycle of motion of a hand holding and utilizing a hammer. The video was 1 s in duration, had a resolution of 400×400 pixels, subtended 17.3° of visual angle, and was presented in grayscale. The auditory stimulus was a congruent hammering noise presented monaurally. Lastly, speech stimuli consisted of two audio-visual clips of a female speaker uttering single instance of the syllables/ba/or/ga/. As for the dynamic handheld tools, the visual component of this stimulus had a resolution of 400×400 pixels and subtended 17.3° of visual angle, and was presented in grayscale. Presentations were 2 s in duration, and each presentation included the entire articulation of the syllable, including pre-articulatory gestures (see

Fig. 1 for illustration). The set of audio-visual stimuli onset asynchronies (SOA) utilized were; 0, ± 10 , ± 20 , ± 50 , ± 80 , ± 100 , ± 150 , ± 200 , ± 250 , and ± 300 ms, for the flash beep stimuli, 0, ± 50 , ± 100 , ± 150 , ± 200 , ± 250 , and ± 300 ms, for the complex non-speech stimuli, and 0, ± 50 , ± 100 , ± 150 , ± 200 , ± 250 , ± 300 , and ± 400 ms, for speech stimuli. Negative values indicate SOAs in which the auditory stimulus leads the visual stimulus. Twenty repetitions per condition were performed for both flash-beep and speech stimuli, while 15 were performed for the tool stimulus. Stimuli were presented at a distance of 60 cm from the participants. All stimuli and stimulus control was generated using MATLAB (MathWorks Inc., Natick, MA) software with the Psychophysics Toolbox extension [Brainard, 1997; Pelli, 1997]. The duration and temporal onsets and offsets of the stimuli were confirmed via a Hameg 507 oscilloscope. This stimulus set has been extensively utilized in prior studies [Stevenson & Wallace, 2013; Quinto, Thompson, Russo, & Trehub, 2010].

Procedure

Each participant completed the three separate experiments (SJs using the flash-beep, tool, and speech stimuli) within a larger battery of experiments spread over 4 days. The order of stimulus conditions was randomized across participants, and subjects were instructed to perform a binary (synchronous or not) audio-visual SJ. Participants were seated inside an unlit sound attenuating WhisperRoomTM, and were monitored at all times by a closed circuit infrared camera. Details regarding the experimental procedure and design have been previously reported in Stevenson et al. [2014].

Analysis

All participants completed all trials, and thus there was no difference in the number of trials performed between ASD and TD groups. All trials were included in the analysis and there was no response time restriction. To test for rapid temporal recalibration effects we conducted a one-back analysis (analyzing trial t 's response as a conditional of trial $t-1$'s SOA). Distributions of perceived simultaneity (i.e., report of synchrony) as a function of SOA were compiled for each participant and each stimulus type (flash-beep, tool, and speech) separately for the case where trial $t-1$ had a negative SOA (i.e., audition led) and where trial $t-1$ had a positive SOA (i.e., vision led). These distributions were subsequently fitted with a Gaussian normal distribution (See Fig. 2) whose amplitude, mean, and standard deviation were free to vary (See Eq. (1)). The shape of the normal distribution proved to accurately describe the reports of synchrony (mean $R^2 = 0.932$, see Supporting Information for more detail). The mean of the best-

fitting distribution was taken as the PSS, and the distribution's standard deviation as a measure of the TBW. The PSS is the SOA at which participants are most likely

to categorize a presentation as synchronous and the TBW is the temporal interval over which participants are likely to categorize the presentation as synchronous).

$$P(\text{response}|\text{SOA}) = \text{amp} \times \exp\left(-\frac{(\text{SOA}-\text{PSS})^2}{2\text{SD}^2}\right) \quad (1)$$

The amount of change in PSS as a function of the prior trial ($\Delta\text{PSS} = \text{PSS audio-leading} - \text{PSS visual-leading}$) and TBW ($\Delta\text{TBW} = \text{TBW audio-leading} - \text{TBW visual-leading}$) was calculated for each type of SJ task (flash-beep, tool, speech) and for each participant [see Harvey, Van der Burg, & Alais, 2014; Van der Burg & Goodbourn, 2015; Van der Burg et al., 2013]. The group mean change in PSS is depicted in Figure 2 as the distance between the blue vertical line (prior trial was audio-leading) and the red vertical line (prior trial was visual-leading).

We also attempted to divide our data not only on the nature of the $t-1$ trial (audio-leading vs. visual-leading), but at a finer grain, depending on the particular stimulus onset asynchrony (SOA) presented on trial $t-1$. However, the limited number of repetitions per condition prevented the reliable extraction of PSS and TBW values for every SOA at $t-1$.

Results

ASD and TD individuals recalibrate similarly for speech stimuli but not for simple or complex non-speech stimuli

Mean PSS did not differ between TD and ASD groups across any of the stimuli employed (all $P > 0.6$). In contrast, an analysis based on the previous trial (auditory-leading or visual-leading) showed a different picture (see Fig. 2). A 2 (Group: TD vs. ASD) X 3 (Stimulus Complexity: flash-beep, tool, speech) mixed-model ANOVA was ran on the amount of change PSS values exhibited as a consequence of prior trial history. As illustrated in Figure 3, findings revealed no main effect of Task Complexity ($F(2, 110) = 1.169, P = 0.314, 1 - \beta = 0.25$), and only a trend toward a main effect of Group ($F(1, 55) = 2.751, P = 0.102, 1 - \beta = 0.37$). Nonetheless, results did show a significant Group X Stimulus

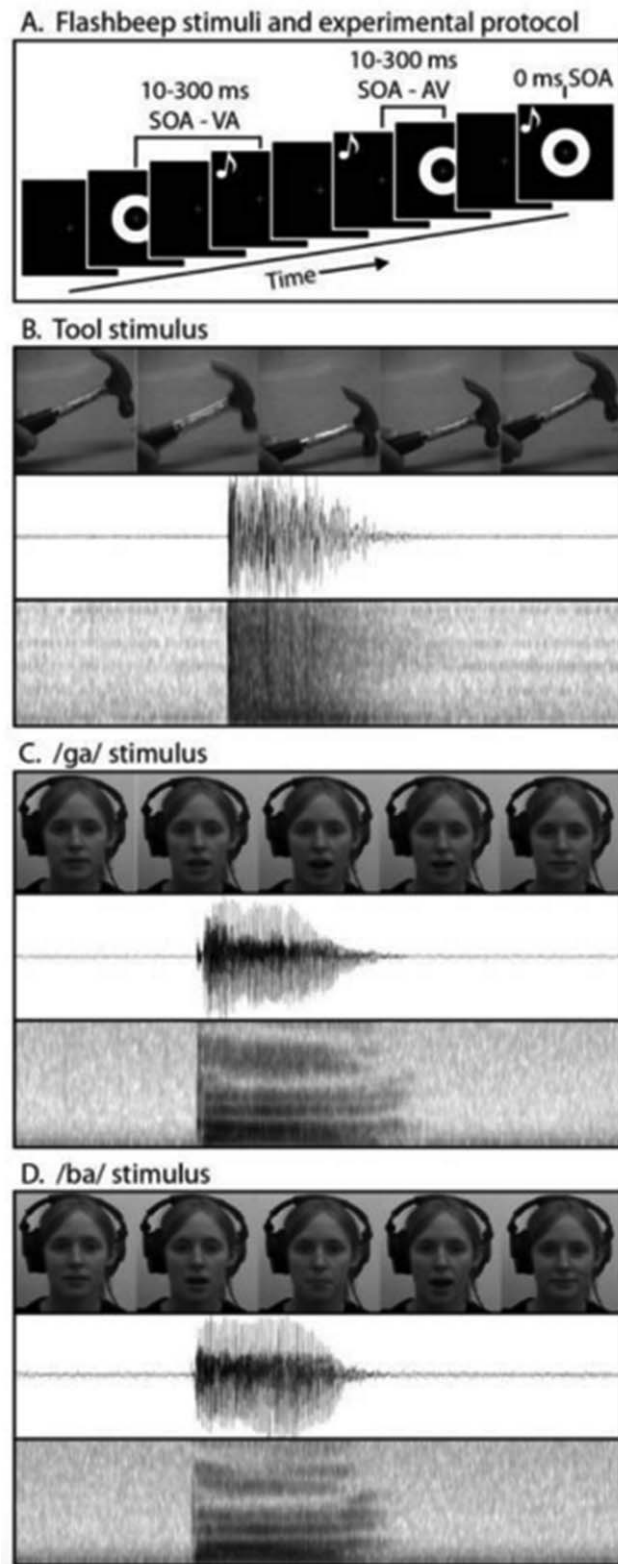


Figure .1.

Figure 1. Methods and Apparatus. flash-beep (A), complex dynamic non-speech (B), and speech (C and D) audio-visual stimuli were presented at a certain stimulus onset asynchrony. Participants were to report whether the stimuli were presented synchronously or not. Top to bottom is represented, the individual frames from dynamic visual stimuli, the auditory waveform, and the auditory spectrogram for the stimuli utilized. Trials began with a 500–1500 ms inter-trial interval, followed by a stimulus presentation. After the stimulus presentation, a response screen appeared, and the next trail began after participants responded.

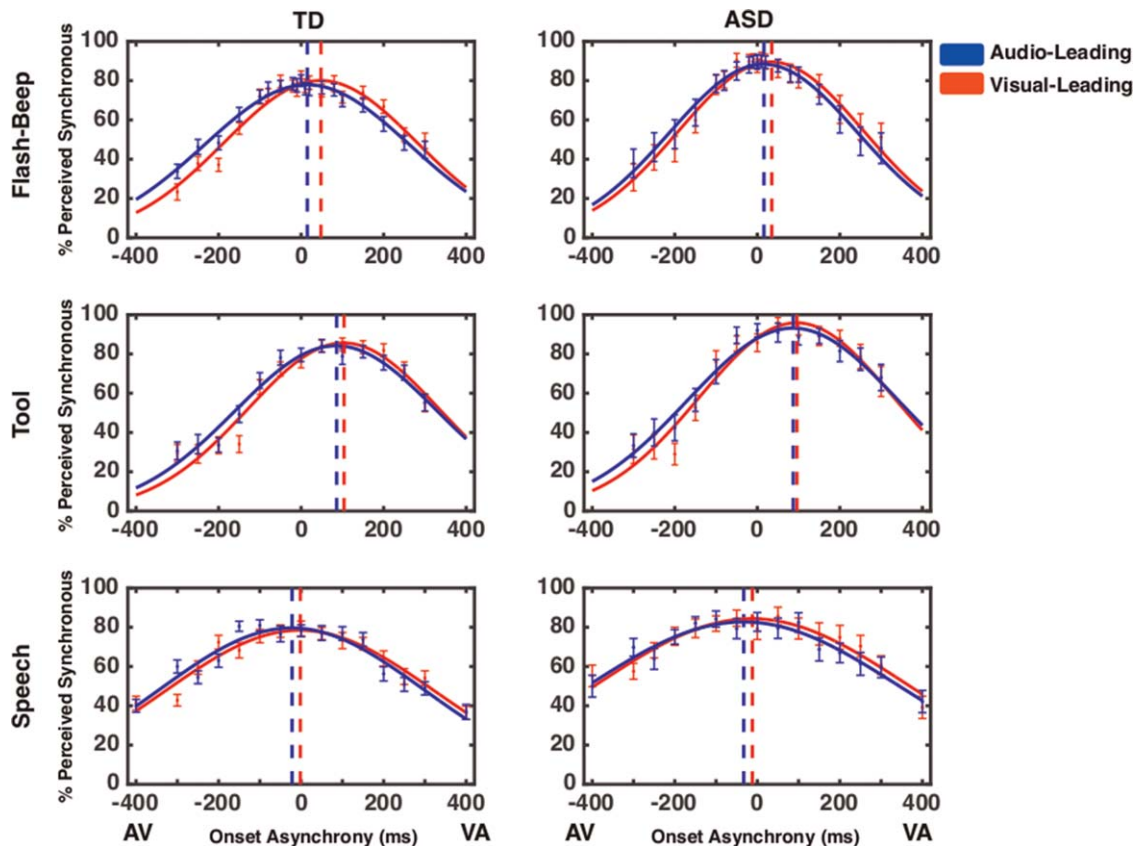


Figure 2. Mean proportion with which participants reported perceiving the audiovisual stimuli as simultaneous (y-axis) as a function of stimulus onset asynchrony (SOA, on x-axis). These functions are shown separately for stimulus complexity (top-to-bottom, from least to most complex stimuli; flash-beep, tool, speech) and for experimental group (left TD, right ASD). Within each plot, two functions are shown based on the timing of the preceding trial (audio-leading in blue, visual-leading in red). The vertical blue (audio-leading) and red (visual-leading) represent, respectively, the mean point of subjective simultaneity (PSS) for each distribution. The distance between the PSS values represents the change in PSS as a consequence of whether the preceding trial was an audio- or visual-leading trial. For pictorial representation, a normal Gaussian distribution is fitted to average data and error bars represent ± 1 S.E.M.

Complexity interaction ($F(2, 75) = 2.108, P = 0.017$, partial $\eta^2 = 0.31$). This interaction was further explored by means of separate independent-samples t -tests (within tasks across groups). These analyses demonstrated that the interaction was driven by significant differences (i.e., changes in the PSS) between TD (mean $\Delta = 33.15$ ms, $SD = 28.47$ ms) and ASD (mean $\Delta = 17.96$ ms, $SD = 16.29$ ms) when performing the flash-beep SJ task ($t(50) = 2.219, P = 0.030$), as well as significant differences when performing the tool SJ task ($t(50) = 2.017, P = 0.048$; TD, mean $\Delta = 18.06$ ms, $SD = 21.93$ ms; ASD, mean $\Delta = 9.12$ ms, $SD = 25.61$ ms). In contrast, no significant differences in the magnitude of shift in PSS were seen between groups for the speech SJ task ($t(50) = 0.044, P = 0.965$; TD, mean $\Delta = 20.86$, $SD = 20.45$; ASD, mean $\Delta = 20.07, SD = 20.36$). An analysis of the effect for stimulus complexity within each group revealed that the degree to which the PSS values changed as a function of prior trial history was depend-

ent on stimulus complexity for TD ($F(2, 24) = 2.947, P = 0.017$), but not for children and adolescents with ASD ($F(2, 24) = 0.314, P = 0.732$). Surprisingly, much as for the TBW size (see below), the change in PSS as a function of the preceding trial was not correlated within-subjects across the various levels of stimulus complexity (all $R^2 < 0.31$, all $P > 0.09$). See Supporting Information online for summary statistics.

Individuals with ASD have larger multisensory TBWs

In close concordance with our prior research [Stevenson et al., 2014], when not taking into account prior trial history, audiovisual temporal acuity as indexed by the size of TBWs was different across the ASD and TD groups. A 2 (Group: TD vs. ASD) X 3 (Stimulus Complexity: flash-beep, tool, speech) mixed-model ANOVA was ran on the width of individual's TBW. Results demonstrated a significant main effect for Stimulus Complexity ($F(2, 75) = 5.604, P = 0.005$), a trend toward

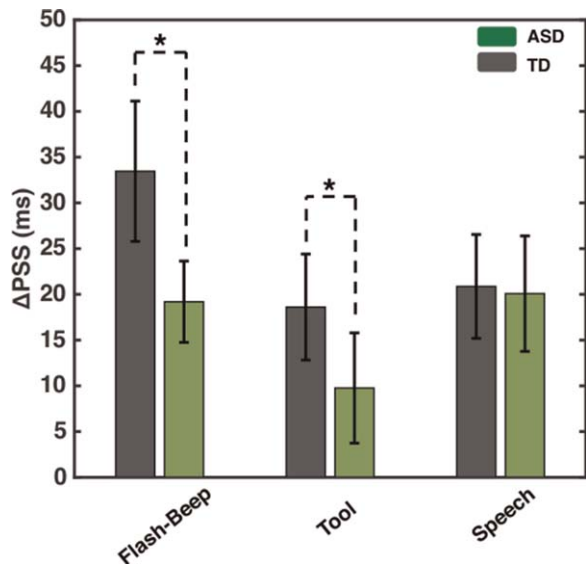


Figure 3. Degree of change in PSS (in ms; Δ PSS = PSS audio-leading – PSS visual-leading) as a function of Task Complexity (x-axis) and Group (ASD in green and TD in black). * indicates $P < 0.05$ (independent samples *t*-test), and error bars represent ± 1 S.E.M.

significance for Group ($F(1, 50) = 2.674, P = 0.108$), and importantly, a significant Stimulus Complexity X Group interaction ($F(2, 75) = 3.451, P = 0.0369$). Again consistent with prior work, this difference was driven by changes in the TBW in response to audiovisual speech stimuli ($t(50) = 2.28, P = 0.026$; TD, $M = 228.54, SD = 68.45$; ASD, $M = 265.69, SD = 53.45$). The size of the TBW did not differ between groups for the tool ($P = 0.37$) or flash-beep stimuli ($P = 0.25$). See Supporting Information online for more detail regarding TBW results.

With regard to the degree that prior trial history drove changes in the TBW across experimental group and stimuli complexity, a 2 (Group: TD vs. ASD) X 3 (Task Complexity: flash-beep, tool, speech) mixed-model ANOVA demonstrated no main effects (Task Complexity, $F(2, 75) = 1.051, P = 0.63, 1 - \beta = 0.64$; Group, $F(1, 50) = 0.08, P = 0.928, 1 - \beta = 0.51$), nor an interaction between Task Complexity and Group ($F(2, 75) = 0.891, P = 0.34, 1 - \beta = 0.154$).

Finally, Van der Burg et al. [2013] report a significant positive relationship between the size of an individual's TBW and the size of the rapid recalibration effect. In the current work we report larger TBWs in ASD compared to TD for speech stimuli, as well as greater recalibration for speech stimuli compared with non-speech stimuli, yet we found no significant relationship in either experimental groups or stimulus complexity (all $R^2 < 0.36$, all $P > 0.08$, correlations performed within each group). Lastly, although our TD and ASD groups were matched for age, they spanned a large age range

(7–17 years old), thus allowing for the possibility of developmental effects masked by the collapsed nature of the analyses. Correlational analysis between the raw PPS and TBWs values, as well as between the magnitude of rapid recalibration and age, revealed no significant effects within this dataset (all $R < 0.34$, all $P > 0.09$).

Discussion

The core finding of the current study was the differences in audiovisual rapid temporal recalibration for children and adolescents with ASD. Whereas those with ASD were found to recalibrate on a trial-per-trial basis similarly to their TD counterparts for speech stimuli, they differed from their TD peers for both complex non-speech (tools) and for simple (flash-beep) stimuli. Intriguingly, this pattern of differences is exactly opposite to that found with regard to more global measures of audiovisual temporal acuity [Bebko et al., 2006; Foss-Feig et al., 2010; Kwakye et al., 2011; Stevenson et al., 2014]. In the Stevenson et al. study, individuals with ASD were found to have large audiovisual TBWs, but only for speech-level stimuli. Indeed, in the current study, we replicate this effect when we examine for TBW size. An initial speculation concerning these two observations (i.e., enlarged TBW and large recalibration effects for speech stimuli in ASD) is that they may represent the same underlying temporal processes (and neural networks). Thus, one could envision a mechanistically linked set of processes such that short-term recalibration scaffolds the creation and maintenance of windows of integration or binding. Such a possibility is supported by the observations of Van der Burg and colleagues [2013], in which individuals with larger TBWs were found to exhibit larger rapid temporal recalibration effects. However, in the current dataset we found no evidence for such a relationship between TBW size and magnitude of rapid recalibration. This lack of a relationship in the current study may suggest a lack of mechanistic commonality for building (and maintaining) the TBW and the processes that underpin the short-term shifts in audiovisual temporal acuity that are driven by the dynamic nature of multisensory temporal relations. The reasons for this are not clear, but as all previous studies on rapid recalibration were performed in adults, one possibility is that the relationship between TBW size and rapid recalibration emerges with age. Further developmental work both in TD and ASD would shed new light on the developmental time course for maturation of the TBW and for rapid recalibration and the relationship between the two. This would be a promising line of enquiry for two reasons: first, from a Bayesian view of autistic perception [Pellicano & Burr, 2012] because many aspects of

Bayesian sensory integration are known to change dramatically through childhood and adolescence [Gori, Sandini, & Burr, 2012], and second, prior work from our group has indicated a protracted developmental time course for the maturation of the TBW in a TD population which only reaches the adult state in late adolescence (18–23 years old; Hillock-Dunn et al., 2011; Hillock-Dunn and Wallace, 2012).

Our observation regarding to extent to which individuals with ASD temporally recalibrate to audiovisual stimuli may provide important insights into how sensory information interacts with and influences existing or prior sensory representations. Despite the growing influence of Bayesian theories of ASD and their emphasis on sensory reliability/noise and updating of prior/posterior representations, to date, little work has been done to directly characterize perceptual plasticity/stability in the context of ASD. The few studies that have been done on experience-dependent adaptation in ASD have shown impairments for both low-level [Turi et al., 2015] and complex, socially relevant, visual stimuli [Pellicano, Jeffery, Burr, & Rhodes, 2007]. However, these studies have been limited to long adaptation procedures and have focused exclusively on a single sensory modality (i.e., vision). Given the increasing recognition that sensory changes in autism extend into the multisensory arena, our study was designed to expand our understanding of experience-dependent plasticity in ASD in response to multisensory stimuli and to probe these changes within a rapid adaptation framework.

It has recently been suggested that individuals with ASD may exhibit attenuated Bayesian priors, resulting in a tendency to perceive the world primarily based on incoming sensory information rather than in the context of past sensory experiences [Pellicano & Burr, 2012]. An attenuated prior should result in a heavier weighting of incoming sensory cues, and thus in more veridical judgments about these cues. In contrast, these individuals would struggle to flexibly adapt (i.e., recalibrate) because of the absence of a sensory referent with which to compare the incoming sensory information. In other words, an attenuated internal representation would mean the discrepancy between the expected and actual sensory signal would be diminished, resulting in weaker recalibration. Concordant with this idea, although not reaching significance, the general trend in the current data for all stimulus types is for less overall recalibration in the ASD group as compared to the TD group.

Complementing the idea that individuals with ASD exhibit attenuated priors, it has been suggested that high level or top-down precision may be attenuated in ASD relative to bottom-up or sensory precision (Lawson, Rees, & Friston, 2014). From this perspective,

prior beliefs should generate top-down predictions about the expectations of sensory input. In the context of recalibration, a discrepancy between top-down predictions, (the internal sensory representation), and sensory inputs will be minimized by optimizing posterior beliefs. Thus, the differences in recalibration may arise from differences between the strength of internal representations and sensory noise, which changes as a function of stimulus complexity (i.e. variability in the sensory information).

With regard the TD individuals; this group also showed greater recalibration effects for the low-level stimuli compared to complex non-speech (tools) and speech stimuli. One interpretation of this could be that the low-level stimuli had less sensory noise when compared with the speech stimuli. For the TD group, the internal representation or expected distribution of sensory noise for low-level stimuli is less than the actual “noisy” distribution of the presented stimuli, thus generating an error signal which drives recalibration in this group. For the ASD group, a noisier internal representation for the lower-level cues, when paired with the same sensory signals, does not produce an error signal, thus resulting in an inability of ASD subjects to recalibrate. In contrast, the greater degree of noise associated with speech signals, coupled with a top-down expectation of a noisier distribution, could be sufficient to drive the recalibration process in both the TD and ASD groups. These interpretations are consistent with the enlarged TBW for speech stimuli in general (when compared with simpler stimuli), as well as with the enlarged TBW for speech stimuli in ASD. Under such a framework, the TBW width can be seen as a proxy measure for the amount of noise contained in the internal representations of the speech signals.

Two further points should be mentioned. First, although statistical analyses did not point toward gender differences in any of the variables of interest, our TD and ASD groups were not matched for gender (reflecting the gender bias in ASD). Thus, we cannot entirely rule out the possibility that the present findings are partially influenced by such a factor. Second, the TD and ASD groups differed in vocabulary proficiency, reflecting a well-established phenotypic characteristic of autism. Although correlational analyses did not reveal a relationship between vocabulary proficiency and any of the measures of interest, it is possible that differences in this variable could interact with the psychophysical measures of interest here—particularly in the case of the speech stimuli. Thus, in future work it will be important to further scrutinize co-variables that may either drive or reflect differences in multisensory temporal acuity and rapid recalibration in TD and ASD. In addition, future work should also focus on better outlining the developmental trajectory of these

rapid recalibration effects, as well as on showing how these short-term changes (dependent on rapid recalibration) interact with and potentially scaffold longer-term cumulative changes in sensory and multisensory representations. In the clinical arena, and most notably in autism, such studies could be of enormous value, not only in providing empirical data important in the evaluation of global models of autism such as weakened predictive coding, but also for the design and implementation of (multi)sensory training protocols and regimens that could be important in autism therapy.

In conclusion, our results confirm prior work in showing that changes in audiovisual temporal acuity (i.e., the TBW) are most prominent for speech stimuli (when compared with simple and complex non-speech stimuli) in those with ASD. In striking contrast to this result, the work also shows that speech is the only stimulus class that exhibits a normal pattern of rapid recalibration. Two interesting questions derive from these results. First, in the absence of rapid recalibration effects for speech, why do individuals with ASD exhibit specific temporal processing deficits for multisensory speech stimuli? Second, and perhaps more vexing, why do individuals with ASD show deficits in multisensory rapid recalibration for simple and complex non-speech stimuli, yet have normal temporal acuity for such stimuli as measured via the TBW? Hence, the work points to a disconnect between the manner in which prior exposure affects the processing of current stimuli. One possibility is that the representations of speech stimuli are so noisy in ASD that rapid recalibration effects fail to accrue with time. Future work should be directed toward understanding the links between rapid audiovisual temporal recalibration and audiovisual temporal acuity, as the current work suggests a dissociation between these that may reflect mechanistic differences between these processes.

Acknowledgments

The project was supported by the Simons Foundation Autism Research Initiative, the Wallace Foundation, NIH T32MH064913, NIH U54HD083211, NIH 1F32 DC011993, a Banting Fellowship from the Canadian National Science and Engineering Research Council, and the Autism Research Training Program funded by the Canadian Health Institutes of Health Research. All authors were involved in the ideation and design of the experiment, MDN and RS performed the experiments, JPN and MDN analyzed the data and thus had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis, JPN and MW wrote the manuscript, which was edited by RS, DA, and MDN. The authors

declare no financial interests or other potential conflicts of interest.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders-V-TR*. Washington, DC: American Psychological Association.
- Bebko, J.M., Weiss, J.A., Demark, J.L., & Gomez, P. (2006). Discrimination of temporal synchrony in intermodal events by children with autism and children with developmental disabilities without autism. *Journal of Child Psychology and Psychiatry*, 47(1), 88–98.
- Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436. pmid:9176952. doi: 10.1163/156856897x00357
- Brandwein, A.B., Foxe, J.J., Butler, J.S., Russo, N.N., Altschuler, T.S., Gomes, H., & Molholm, S. (2013). The development of multisensory integration in high functioning autism: high-density electrical mapping and psychophysical measures reveal impairments in the processing of audiovisual inputs. *Cereb Cortex*, 23, 1329–1341.
- Foss-Feig, J.H., Kwakye, L.D., Cascio, C.J., Burnette, C.P., Kadivar, H., Stone, W.L., et al. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, 203, 381–389. doi: 10.1007/s00221-010-2240-4
- Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, 7, 773–778.
- Gori, M., Sandini, G., & Burr, D. (2012). Development of visuo-auditory integration in space and time. *Frontiers in Integrative Neuroscience*, 6, 77.
- Harvey, C., Van der Burg, E., & Alais, D. (2014). Rapid temporal recalibration occurs crossmodally without stimulus specificity but is absent unimodally. *Brain Research*, 1585, 120–130.
- Hillock-Dunn, A.R., & Wallace, M.T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental Science*, 15:688–696.
- Hillock, A.R., Powers, A.R., & Wallace, M.T. (2011) Binding of sights and sounds: Age-related changes in multisensory processes. *Neuropsychologia* 49:461–467.
- Kwakye, L.D., Foss-Feig, J.H., Cascio, C.J., Stone, W.L., & Wallace, M.T. (2011). Altered auditory and multisensory temporal processing in autism spectrum disorders. *Frontiers in Integrative Neuroscience*, 4, 129
- Lawson, R.P., Rees, G., & Friston, K.J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience*. 8. <http://dx.doi.org/10.3389/fnhum.2014.00302>.
- Marco, E.J., Hinkley, L.B., Hill, S.S., & Nagarajan, S.S. (2011) Sensory processing in autism: a review of neurophysiologic findings. *Pediatric Research*, 69, 48R–54R.
- Noel, J.P., Lukowska, M., Wallace, M.T., & Serino, A. (2016). Multisensory simultaneity judgment and distance from the body. *Journal of Vision*, 16(3), 21. doi: 10.1167/16.3.21
- Noel, J.P., Wallace, M.T., Orchard-Mills, E., Alais, D., & Van der Burg, E. (2015). True and perceived synchrony are

- preferentially associated with particular sensory pairings. *Scientific Reports*, 5, 17467.
- Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442. pmid:9176953. doi: 10.1163/156856897x00366
- Pellicano, E., & Burr, D. (2012). When the world becomes too real: a Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, 16 (10), 504–510. <http://dx.doi.org/10.1016/j.tics.2012.08.009>.
- Pellicano, E., Jeffery, L., Burr, D., & Rhodes, G. (2007). Abnormal adaptive face-coding mechanisms in children with autism spectrum disorder. *Current Biology*, 17, 1508–1512. pmid:17764946. doi: 10.1016/j.cub.2007.07.065
- Pöppel, E. (1985). *Grenzen des Bewusstseins*. Stuttgart: Deutsche Verlags-Anstalt, translated as *Mindworks: Time and Conscious Experience*. New York: Harcourt Brace Jovanovich, 1988.
- Quinto, L., Thompson, W.F., Russo, F.A., & Trehub, S.E. (2010). A comparison of the McGurk effect for spoken and sung syllables. *Attention, perception & psychophysics*, 72(6), 1450–1454.
- Rosenberg, A., Patterson, J.S., & Angelaki, D.E. (2015). A computational perspective on autism. *Proc Natl Acad Sci U S A*, 112(30), 9158–9165.
- Smith, E.G., & Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry*, 48, 813–821.
- Stevenson, R.A., Segers, M., Ferber, S., Barense, M.D., Camarata, S., & Wallace, M.T. (2014). The impact of multisensory integration deficits on speech perception in children with autism spectrum disorder. *Front. Psychol.* 5, 379.
- Stevenson, R.A., Siemann, J.K., Schneider, B.C., Eberly, H.E., Woynaroski, T.G., Camarata, S.M., & Wallace, M.T. (2014). Multisensory temporal integration in autism spectrum disorders. *Journal of Neuroscience*, 34(3), 691–697.
- Stevenson, R.A., & Wallace, M.T. (2013). Multisensory temporal integration: Task and stimulus dependencies. *Experimental Brain Research*, 227(2), 249–261.
- Stevenson, R.A., Zemtsov, R.K., & Wallace, M.T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance*, 38(6), 1517.
- Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature*, 421(6926), 911, doi:10.1038/421911a
- Sumby, W.H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, 26, 212–215. doi: 10.1121/1.1907309
- Turi, M., Burr, D., Iglizzi, R., Aagten-Murphy, D., Filippo, M., & Pellicano, E. (2015). Children with autism spectrum disorder show reduced adaptation to number. *PNAS*, 12(25), 7868–7872.
- Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid recalibration to asynchronous audiovisual stimuli. *Journal of neuroscience*, 33, 633–637. doi:10.1523/JNEUROSCI.1182-13.2013
- Van der Burg, E., & Goodbourn, P.T. (2015). Rapid, generalized adaptation to asynchronous audiovisual speech. *Proceedings of the Royal Society*, 282, 20143083. <http://dx.doi.org/10.1098/rspb.2014.3083>
- Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics*, 72(4), 871–884. doi:10.3758/APP.72.4.871
- Vroomen, J., Keetels, M., De Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research*, 22, 32–35.
- Wallace, M.T., & Stevenson, R.A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia*, 64, 105–123.
- Zaidel, A., Goin-Kochel, R.P., Angelaki, D.E. (2015). Self-motion perception in autism is compromised by visual noise but integrated optimally across multiple senses. *Proceedings of the National Academy of Sciences of the United States of America*, 112(20), 6461–6466.