Individual Differences in the Multisensory Temporal Binding Window Predict Susceptibility to Audiovisual Illusions

Ryan A. Stevenson, **Raquel K. Zemtsov**, and **Mark T. Wallace**

Ryan A. Stevenson, Department of Hearing and Speech Sciences, Vanderbilt University Medical Center; Raquel K. Zemtsov, Undergraduate Neuroscience Program, Vanderbilt University; Mark T. Wallace, Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Vanderbilt University Kennedy Center, Department of Psychology, Vanderbilt Brain Institute, Vanderbilt University

Abstract

Human multisensory systems are known to bind inputs from the different sensory modalities into a unified percept, a process that leads to measurable behavioral benefits. This integrative process can be observed through multisensory illusions, including the McGurk effect and the sound-induced flash illusion, both of which demonstrate the ability of one sensory modality to modulate perception in a second modality. Such multisensory integration is highly dependent upon the temporal relationship of the different sensory inputs, with perceptual binding occurring within a limited range of asynchronies known as the temporal binding window (TBW). Previous studies have shown that this window is highly variable across individuals, but it is unclear how these variations in the TBW relate to an individual's ability to integrate multisensory cues. Here we provide evidence linking individual differences in multisensory temporal processes to differences in the individual's audiovisual integration of illusory stimuli. Our data provide strong evidence that the temporal processing of multiple sensory signals and the merging of multiple signals into a single, unified perception, are highly related. Specifically, the width of right side of an individuals' TBW, where the auditory stimulus follows the visual, is significantly correlated with the strength of illusory percepts, as indexed via both an increase in the strength of binding synchronous sensory signals and in an improvement in correctly dissociating asynchronous signals. These findings are discussed in terms of their possible neurobiological basis, relevance to the development of sensory integration, and possible importance for clinical conditions in which there is growing evidence that multisensory integration is compromised.

Keywords

multisensory integration; cross-modal; McGurk; sound-induced flash illusion; perception; temporal processing

Human sensory systems have the ability to transduce and process information from distinct types of energy in the environment. In addition to the workings of the individual sensory systems, multisensory networks have the ability to bind these signals into a coherent, unified perception of external events (Gaver, 1993). This integration of information across the senses improves our ability to interact with the environment, allowing for enhanced detection (Lovelace, Stein, & Wallace, 2003; Stein & Wallace, 1996), more accurate

localization (Nelson et al., 1998; Wilkinson, Meredith, & Stein, 1996), and faster reactions (Diederich & Colonius, 2004; Hershenson, 1962). In addition to these highly adaptive benefits, multisensory integration can also be seen in a host of perceptual illusions, where the presentation of a signal in one sensory modality can modulate perception in a second sensory modality. One of the more compelling of these illusions is the sound-induced flash illusion, in which an observer is presented with a single visual flash paired with multiple auditory cues (i.e., beeps) in rapid succession. The participant is instructed to ignore the beeps and to report only the number of visual flashes. Despite the presence of only a single visual flash, observers often report seeing multiple flashes when paired with two or more auditory cues (Shams, Kamitani, & Shimojo, 2000). Whereas this illusion takes advantage of very low-level stimulus pairings (i.e., flashes and beeps) other illusions can be demonstrated using stimuli rich in semantic content. In the speech realm, this is best illustrated by the McGurk effect, in which the combination of a visual /ga/ and an auditory /ba/ often results in the perception of a novel syllable (i.e., /da/ (McGurk & MacDonald, 1976).

These interactions between the senses, including the illusory examples described above, are highly dependent upon the characteristics of the stimuli that are combined. These findings were originally reported in animal models with neuronal recordings, showing that the more spatially congruent (Meredith & Stein, 1986a) and temporally synchronous (Meredith, Nemitz, & Stein, 1987) the stimuli are to one another, the larger the change (i.e., gain) seen under multisensory circumstances. In addition, the less effective the individual stimuli are in generating a response, the greater the multisensory effect when they are combined (Meredith & Stein, 1986b). These effects have been extended to human psychophysical and neuroimaging studies with a great deal of interest directed of late to the temporal features of these multisensory interactions. In large measure these studies have paralleled the animal model work referenced above, showing that the largest multisensory interactions occur when stimuli are combined in close temporal proximity. Evidence for this has been gathered in behavioral studies (Conrey & Pisoni, 2004; Conrey & Pisoni, 2006; Dixon & Spitz, 1980; Foss-Feig et al., 2010; Hillock, Powers, & Wallace, 2011; Keetels & Vroomen, 2005; Powers, Hillock, & Wallace, 2009; van Atteveldt, Formisano, Blomert, & Goebel, 2007; van Wassenhove, Grant, & Poeppel, 2007; Wallace et al., 2004; Zampini, Guest, Shore, & Spence, 2005), in event-related potentials (Schall, Quigley, Onat, & Konig, 2009; Senkowski, Talsma, Grigutsch, Herrmann, & Woldorff, 2007; Talsma, Senkowski, & Woldorff, 2009), and in functional MRI studies (James & Stevenson, 2012; Macaluso, George, Dolan, Spence, & Driver, 2004; Miller & D'Esposito, 2005; Stevenson, Altieri, Kim, Pisoni, & James, 2010; Stevenson, VanDerKlok, Pisoni, & James, 2011). In addition to these findings, it has been shown that multisensory stimuli presented in close temporal proximity are often integrated into a single, unified percept (Andersen, Tiippana, & Sams, 2004; McGurk & MacDonald, 1976; Shams et al., 2000; Stein & Meredith, 1993). This perceptual binding over a given temporal interval is best captured in the construct of a multisensory temporal binding window (TBW; Colonius & Diederich, 2004; Foss-Feig et al., 2010; Hairston, Burdette, Flowers, Wood, & Wallace, 2005; Powers et al., 2009).

The TBW is a probabilistic concept, reflecting the likelihood that two stimuli from different modalities will or will not result in alterations in behavior and/or perception that likely reflect the perceptual binding of these stimuli across a range of stimulus asynchronies. The TBW typically exhibits a number of characteristic features (for review, see Vroomen & Keetels, 2010). First, as the temporal interval between the stimuli (i.e., the stimulus onset asynchrony, SOA) increases, the likelihood of multisensory interactions decreases (Conrey & Pisoni, 2006; Hirsh & Fraisse, 1964; Keetels & Vroomen, 2005; Miller & D'Esposito, 2005; Powers et al., 2009; Spence, Baddeley, Zampini, James, & Shore, 2003; Stevenson et al., 2010; van Atteveldt, et al., 2007; van Wassenhove, et al., 2007; Vatakis & Spence,

2006). This reflects the statistics of the natural environment, where sensory inputs which are not closely related in time are less likely to have originated from a single event. Second, the TBW is often asymmetrical, with the right side (reflecting conditions in which visual stimuli precedes auditory stimuli) being wider (Conrey & Pisoni, 2006; Dixon & Spitz, 1980; Hillock et al., 2011; Stevenson et al., 2010; van Atteveldt, et al., 2007; van Wassenhove, et al., 2007; Vroomen & Keetels, 2010). This asymmetry has been attributed to the difference in propagation times of visual and auditory stimulus energies, where auditory lags increase as stimulus distance grows (Pöppel, Schill, & von Steinbüchel, 1990). In addition, there are substantial differences in the timing of the acoustic and visual transduction processes, as well as in neural conduction times for these modalities (Corey & Hudspeth, 1979; King & Palmer, 1985; Lamb & Pugh, 1992; Lennie, 1981). As such, the SOA at which the probability of multisensory interactions peaks is often also asymmetrical, generally being greatest when the auditory signal slightly lags the visual signal (Dixon & Spitz, 1980; Meredith et al., 1987; Roach, Heron, Whitaker, & McGraw, 2011; Seitz, Holloway, & Watanabe, 2006; Zampini et al., 2005; Zampini, Shore, & Spence, 2003, 2005). Third, the width of the TBW varies with stimulus type, with simple stimuli with sharp rise times generally giving rise to narrower TBWs (Hirsh & Sherrick, 1961; Keetels & Vroomen, 2005; Zampini et al., 2005; Zampini et al., 2003) relative to those observed with more complex stimuli such as speech (Conrey & Pisoni, 2006; Miller & D'Esposito, 2005; Stevenson et al., 2010; van Atteveldt, et al., 2007; Van der Burg, Cass, Olivers, Theeuwes, & Alais, 2010; van Wassenhove, et al., 2007; Vatakis & Spence, 2006). Finally, as a prelude to the current study, in many of the aforementioned studies it is clear that there are striking individual differences in each of these attributes of the TBW, the most salient of which are its width, asymmetry, and the SOA at which the probability of integration peaks (Miller & D'Esposito, 2005).

While it is clear that more synchronous stimulus presentations are typically associated with increased multisensory integration, it is unclear how individual differences in the TBW are related to an individual's integrative abilities. We suspect that the TBW and the capacity for and magnitude of multisensory integration may be related based upon three premises. First, the developmental chronology of the TBW (Hillock et al., 2011) and the susceptibility to multisensory illusions such as the McGurk Effect (McGurk & MacDonald, 1976) follow a similar time course. Second, a number of clinical populations exhibit comorbid impairments in both the TBW and in the perception of multisensory illusions. These include autism (Foss-Feig et al., 2010; Kwakye, Foss-Feig, Cascio, Stone, & Wallace, 2011; Taylor, Isaac, & Milne, 2010), dyslexia (Bastien-Toniazzo, Stroumza, & Cavé, 2009; Hairston et al., 2005), and schizophrenia (Foucher, Lacambre, Pham, Giersch, & Elliott, 2007; Pearl et al., 2009). Finally, the neural underpinnings of the sound-induced flash illusion (Bolognini, Rossetti, Casati, Mancini, & Vallar, 2011; Watkins, Shams, Josephs, & Rees, 2007; Watkins, Shams, Tanaka, Haynes, & Rees, 2006) and McGurk effect (Beauchamp, Nath, & Pasalar, 2010; Nath & Beauchamp, 2011b; Sekiyama, Kanno, Miura, & Sugita, 2003) share common substrates that show modulations in functional activity with variations in real and perceived asynchrony (Macaluso et al., 2004; Miller & D'Esposito, 2005; Stevenson et al., 2010; Stevenson et al., 2011) and with multisensory temporal processing (Dhamala, Assisi, Jirsa, Steinberg, & Kelso, 2007; Noesselt et al., 2007). Our results will be discussed specifically in light of each of these previous findings with the overarching hypothesis that individuals with a narrower TBW will encounter fewer occurrences of stimulus inputs that are perceived to be temporally synchronous and that perceived synchronous events will be more tightly bound. Such a view results in the predictions that those with greater precision in the temporal constraints of their multisensory networks (i.e., narrower temporal binding windows) will show (a) an increased ability to dissociate, or failure to bind, sensory signals that are *asynchronous* (i.e., a correct rejection);.and (b) greater enhancements associated with integration of *synchronous* sensory inputs.

In the experiments described here, we seek to provide the first comprehensive evidence linking individual differences in multisensory temporal processes (i.e., the TBW) to differences in the individuals' audiovisual integration as indexed through two robust multisensory illusions. Specifically, we determined individuals' TBWs through the use of a simultaneity judgment task employing low-level visual and auditory stimuli (i.e., flashes and beeps) and, in the same individuals, measured the strength of the McGurk effect and the sound-induced flash illusion. Our hypothesis that individuals with narrower TBWs would exhibit greater magnitude of multisensory integration with synchronous inputs predicts that these individuals should show a stronger McGurk effect. Conversely, the hypothesis that individuals with narrower windows should be more temporally precise and subsequently fail to integrate *asynchronous* stimulus inputs leads to the prediction that these individuals will be less susceptible to the sound-induced flash illusion, which requires the integration of auditory and visual inputs that are not temporally aligned. Our results suggest that this pattern is indeed seen in the relationship between the TBW and the ability to form fused, unified percepts from multisensory inputs, a finding that has important mechanistic implications for the neural networks that subserve multisensory binding.

Method

Participants

Participants included 31, right-handed, native-English speaking Vanderbilt students (13 male, mean age = 21 years, SD = 4 years, age range = 18–35 years) and who were compensated with class credit. Participants reported normal hearing and normal or corrected-to-normal vision. All recruitment and experimental procedures were approved by the Vanderbilt University Institutional Review Board.

Stimuli

All stimuli throughout the study were presented using MATLAB (MATHWORKS Inc., Natick, MA) software with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). Visual stimuli were presented on a NEC MultiSync FE992 monitor at 100 Hz at a distance of approximately 60 cm from the participants. Auditory stimuli were presented binaurally via Phillips noise-cancelling SBC HN-110 headphones. The duration of all visual and auditory stimuli, as well as the SOAs, was confirmed using a Hameg 507 oscilloscope with a photovoltaic cell and microphone.

Two categories of audiovisual stimuli were presented: simple paired flash-beep stimuli for the 2-AFC simultaneity judgment task and sound-induced flash illusion task and single syllable utterances for the McGurk task. The visual component of the sound-induced flash illusion task (Figure 1A) consisted of a white ring circumscribing the visual fixation cross on a black background. Visual stimulus duration was 10 ms. Auditory stimuli consisted of a 3500 Hz pure tone with a duration of 7 ms. When multiple flashes were presented in the sound-induced flash illusion task, they were separated by 43 ms intervals (Figure 1A).

Single syllable utterances for the McGurk task were selected from a stimulus set that has been previously used successfully in studies of multisensory integration (Quinto, Thompson, Russo, & Trehub, 2010). Stimuli consisted of two audiovisual clips of a female speaker uttering single instances of the syllables /ga/ and /ba/ (Figures 1B and 1C, respectively). Visual stimuli were cropped to square, down-sampled to a resolution of 400×400 pixels spanning 18.25 cm per side, and converted from color to grayscale. Presentations were shortened to 2 s, with each presentation containing the entire articulation of the syllable, including prearticulatory gestures.

Procedure

Each participant completed three separate experiments within a larger battery of experiments that were spread over 4 days. Experiment orders were randomized across participants.

The first task was a two-alternative forced-choice (2AFC) simultaneity judgment, used to derive the TBW. In this task, participants were presented with a series of flash-beep stimuli at parametrically varied SOAs. SOAs were 0, 10, 20, 50, 80, and 100–300 ms in 50 ms increments, with both auditory-preceding-visual and visual-preceding-auditory stimulus presentations. Thus, a total of 21 SOAs were used, with 20 trials per condition. The participants were asked to report whether each presentation was temporally synchronous or asynchronous via a button press.

The second task was the sound-induced flash illusion paradigm. Participants were presented with a number of flashes and beeps and were asked to indicate how many flashes they perceived. Participants were explicitly instructed to respond according to their visual perception only. Trials included a single flash presented with 0–4 beeps as well as control presentations of 2–4 flashes with either a single beep or no beep. Thus, a total of 11 conditions were presented with 25 trials per condition. Participants responded via button press. In all conditions with auditory stimuli, the first auditory beep was presented synchronously with the visual flash.

The third and final task was a test of the McGurk effect. Twenty-six of the participants completed this task. Participants were presented with audio-only (with the fixation cross remaining on the screen), visual-only, and audiovisual versions of the /ba/ and /ga/ stimuli described above. Additionally, they were also presented with a McGurk stimulus in which the visual /ga/ was presented with the auditory /ba/. Thus, a total of seven stimulus conditions were presented, with 20 trials in each condition. Participants were asked to report what syllable they perceived by indicating what the first letter of the syllable was via button press, including the options to respond with /ba/, /ga/, /da/, and /tha/.

For all tasks, participants were seated inside an unlit sound attenuating WhisperRoom™ (Model SE 2000; Whisper Room Inc, Morristown, TN), were asked to fixate a central cross, and were monitored by close circuit infrared cameras throughout the experiment to ensure fixation. Each task began with an instruction screen, after which the participant was asked if he or she understood the instructions. Each trial began with a fixation screen for 500 ms plus a random jitter ranging from 1 to 1000 ms, after which the stimulus(i) was presented and then followed by an additional 250 ms fixation screen. A response screen was then presented during which the participant responded via button press ("Was the presentation synchronous," "How many flashes did you see," or "What did she say," for the perceived simultaneity, sound-induced flash illusion, and McGurk tasks, respectively). Following the response, the fixation screen beginning the subsequent trial was presented. The order of trial types was randomized with all tasks, and was randomly generated for each participant for each experiment.

Analysis

Responses from the simultaneity judgment task in which participants reported if they perceived the audiovisual presentation as either simultaneous or not were used to calculate TBWs for each participant. The first step in this process was to calculate a rate of perceived simultaneity with each SOA, which was simply the percentage of trials in a given condition in which the individual reported that the presentation was simultaneous. Two psychometric sigmoid functions were then fit to the rates of perceived simultaneity across SOAs, one to the audio-first presentations and a second to the visual-first presentations which will be referred to as the left (AV) and right (VA) windows, respectively (Figure 2A). This data was

used to create both best-fit sigmoids included the simultaneous, 0 ms SOA condition. These best-fit functions were calculated using the *glmfit* function in MATLAB. Each participant's individual left and right TBWs were then estimated as the SOA at which the best-fit sigmoid's y-value equaled a 50% rate of perceived simultaneity (Stevenson, Zemtsov, & Wallace, under review). Group TBWs were then calculated by taking the arithmetic mean of the respective left and right TBWs from each participant.

Responses to the sound-induced flash illusion were calculated using trials in which a single flash was presented with varying numbers of beeps (0–4). For each of these five trial types, each participant's mean response was calculated, and a group average response was calculated as a mean of individuals' means (Figure 3A). An additional index was calculated to describe each individual's susceptibility to the illusion. This index consisted of the individual's average change from the nonillusory condition (one flash paired with one beep) to each of the three illusory conditions (one flash paired with 2–4 beeps) relative to the number of illusion-inducing beeps. The usefulness of this index is twofold. First, it accounts for any individuals who may have consistently reported perceiving multiple flashes when there was only a single flash presented. Second, this index gives a single numerical index for each individual across conditions, ensuring that the pattern of correlations reported here are driven by a consistent relationship within each individual (e.g., ensuring that the correlation between the 2-beep and 3-beep illusory condition and the TBW are not each driven by different individual outliers in those specific conditions). The susceptibility index was calculated using the following equation,

$$Sl = \frac{\left[\frac{R_2 - R_1}{1} + \frac{R_3 - R_1}{2} + \frac{R_4 - R_1}{3} \right]}{3}$$

where R_n represents the individuals mean response to the condition with one flash and *n* beeps. This index gives a single metric for the individual's susceptibility to the illusion where an index of 1 would reflect a scenario in which each additional auditory cue increased the number of perceived flashes by 1, and an index of 0 would reflect a scenario in which the addition of auditory cues did not change the visual perception. It should be noted that it was possible that participants could deduce that every time multiple beeps were heard, the correct response was to report a single visual flash; however, these results suggest that this did not occur.

Responses to the visual-only conditions were calculated using trials in which 1–4 flashes were presented without beeps. For each of these four trial types, each participant's mean responses were calculated, and group average responses were calculated as a mean of individuals' means (Figure 3B).

Responses with the speech stimuli were used to calculate each participant's strength of the McGurk effect. First, a baseline was calculated for each participant equal to the percentage of unisensory trials in which they perceived a unisensory /ba/ or /ga/ as either /da/ or /tha/ (Figure 4B). Next, the percentage of trials in which each participant reported the fused (i.e., McGurk) percept, an auditory /ba/ paired with a visual /ga/ as either /da/ or /tha/ was calculated (Figure 4A). For simplicity, we will henceforth refer to illusory perceptions only as /da/. Each participant's McGurk score was calculated as the percentage of perceived McGurk illusions relative to their unisensory baseline using the equation:

$$p(AV\,McGurk) \times [1 - p(Unisensory/da/)],$$

where $p(AV McGurk)$ is the individual's rate of McGurk percepts with audiovisual McGurk stimuli, and $p(Unisensory / da)$ is the rate at which the individual reported perceiving /da/ with unisensory /ba/ and /ga/ stimuli. This choice of baseline was made due to the ambiguity of the visual /ga/, where it was often misperceived as /da/ (Figure 4A), a misperception that has been frequently reported in the literature (Erber, 1972; Massaro, 1998; Massaro, 2004; Summerfield, 1987). When all unisensory components produce highly accurate responses, no such unisensory baseline is needed and one can simply compare the illusory percepts (for an example, see Grant, Walden, & Seitz, 1998). However, our choice of baseline ensures that a participant's perceptions of the McGurk effect are not due to incorrect perceptions of the unisensory components of the McGurk stimuli. Additionally, percent correct responses were calculated to the audiovisual congruent /ba/ and /ga/ presentations to ensure recognition of the utterances (Figure 4B).

Results

Determination of the Temporal Binding Window

The temporal binding window (TBW) was calculated for each participant on the basis of their simultaneity judgment responses, with the full window being created by merging the sigmoid functions used to create the left (AV) and right (VA) sides of the response distributions (see methods for additional detail on the creation of these windows). On average, the width of the right TBW (215 ms ± 19 ms SEM) was significantly larger than the left TBW (156 ms ± 10 ms SEM; $t = 4.56$, $p < .0001$; Figure 2A). The relationship between left and right TBWs for individuals was also tested, with the two showing a significant correlation (see Figure 2B for detailed statistics).

Correlation Between the TBW and Sound-Induced Flash Illusion Performance

In these same individuals, the pattern of results on the sound-induced flash illusion task revealed results in keeping with the original description of the illusion (Shams et al., 2000). Thus, and as illustrated in Figure 3A and Table 1, the pairing of a single flash with multiple beeps drove perceptual reports toward multiple flashes. Trials on which a single flash was presented concomitant with either no beep or a single auditory beep were associated with the least number of perceived flashes, and did not differ from each other. Trials with multiple auditory beeps were associated with a greater number of perceived flashes than in the 0- and 1-beep conditions. Additionally, each addition of a beep significantly increased the number of perceived flashes. For detailed statistics of each pairwise comparison, see Table 1.

Responses with the control, visual-only conditions that were interleaved with the sound-induced flash illusion illusory trials were also analyzed (Figure 3B). As for the illusory trials, results were in keeping with the original report (Shams et al., 2000), with parametric increases in perceived flashes as the actual number of flashes increased. For detailed statistics of each pairwise comparison, see Table 2.

Each individual's left and right TBWs were then correlated with their sound-induced flash illusion responses as measured with one, two, three, and four flashes (Figure 3C). No significant correlation was seen in the nonillusory, single-flash, single-beep condition. However, in each of the illusory conditions, the width of the TBW was significantly correlated with sound-induced flash illusion strength at an -value of 0.05. When corrected for multiple comparisons (10 comparisons including susceptibility index correlations, Greenhouse-Geisser corrected = .005), only the right TBW was significantly correlated with sound-induced flash illusion strength. In all cases, the strength of the illusion *decreased* with narrower TBWs. See Figure 3C for detailed statistics for each correlation.

In addition to these singular measures, a more global susceptibility index for the illusion was calculated (see Method) and was also found to vary greatly between individuals, with a range from 0 (no illusory reports) to 0.85 (high degree of bias associated with number of beeps). The mean susceptibility index was 0.38 with a standard deviation of 0.24. Supporting the individual measures, this index showed a correlation with both the left and right TBW (Figure 3D), with the correlation only being significant for the right TBW after correction for multiple comparisons. For detailed statistics, see Figure 3D.

Correlation Between the TBW and McGurk Performance

In addition to sound-induced flash illusion performance, behavioral responses associated with the presentation of auditory, visual, and audiovisual speech syllables were also analyzed for interactions. The paradigm was structured to examine the McGurk effect, in which the presentation of auditory /ga/ and visual /ba/ syllables often give rise to a novel and synthetic percept (i.e., /da/). Of the individuals that reported such fusions ($n = 21$), the illusory /da/ percept was reported on 58.4% of trials when the auditory /ga/ and visual /ba/ were paired (Figure 4A). In contrast to these incongruent trials, accuracy rates for congruent /ba/ and /ga/ multisensory trials were near 100%. In addition, accuracy rates were very high for unisensory (i.e., visual alone, auditory alone) presentations, with the exception of the visual /ga/, which was more ambiguous (Figure 4B). When correcting for these inaccurate perceptions of the unisensory components of the McGurk stimuli (see Method), the average rate of perceptual fusion was 52.3% ($SE = 6.7\%$). Additionally, with McGurk stimulus presentations, participants were more likely to perceive /da/ than either /ba/ or /ga/ ($p < 8.99\mathrm{e}^{-5}$, $t = 4.88$; $p < .03$, $t = 2.44$, respectively).

Similar to the sound-induced flash illusion analysis, the rate of each individual's perception of the McGurk effect was then correlated with his or her left and right TBWs (Figure 4C). There was no correlation between left TBWs and a participant's rate of reporting the McGurk effect, even at an uncorrected $= .05$, but there was a significant negative correlation between width of the participant's right TBW and his or her rate of McGurk perception. Intriguingly, this relationship is the opposite of that seen for the sound-induced flash illusion, with the rate of the McGurk effect *decreasing* with wider right TBWs. See Figure 4C for detailed statistics.

Correlations Between the Sound-Induced Flash Illusion and McGurk Tasks

Finally, the rate of each individual's perception of the McGurk effect was correlated with their perception of the sound-induced flash illusion (see Figure 5). Correlations relating the McGurk effect with the 1–4 beep conditions as well as the susceptibility index were calculated. Significant negative correlations were found with the 2–4 beep conditions as well as with the susceptibility index (five comparisons, corrected $= .01$). See Figure 5 for detailed statistics of each correlation.

Discussion

In these experiments, we have described the relationship between individuals' temporal constraints in integrating multisensory cues (i.e., the temporal binding window) and their strength of integrating or binding these cues as indexed through illusory percepts. Three main conclusions can be drawn from these data. First, there is a clear and significant correlation between the width of the temporal binding window (TBW), as measured through a simultaneity judgment task, and the strength of audiovisual integration with synchronous sensory stimuli, as measured via the McGurk effect. Second, the width of the TBW was correlated with the temporal precision of individual's multisensory integration, in that individuals with narrower TBWs were able to better dissociate asynchronous audiovisual

inputs, measured via the sound-induced flash illusion. Third, and perhaps most intriguing, these relationships are specific to the right side of the TBW (i.e., reflecting when an auditory stimulus lags a visual stimulus).

The TBW as a construct relates to the ability of our perceptual systems to bind inputs from multiple sensory modalities into a singular perceptual gestalt. This ability is constrained by the temporal relationship of the unisensory components of an audiovisual stimulus pair; within a limited temporal range of offsets, binding or integration is highly likely to occur. However, as the offset increases, the probability of binding decreases. Both current and prior work has shown marked variability in the size of the TBW across individuals, suggesting that the degree of tolerance for temporal asynchrony in multisensory inputs is highly variable (Conrey & Pisoni, 2006; Dixon & Spitz, 1980; Miller & D'Esposito, 2005; Powers et al., 2009; Stevenson et al., 2010). The observation of these striking differences coupled with previous findings led to the underlying hypotheses that motivated the current study— that a narrower TBW will be associated with an increased strength of integration for synchronous sensory (i.e., audiovisual) inputs due an increase in the uniqueness of such perceived synchronous events. The corollary to this increase in integration of synchronous stimuli is that these individuals should also be better able to parse, or correctly fail to bind, stimulus inputs that are not synchronous. Together these improvements in temporal precision would manifest as greater gains in accuracy, reaction time (RT), and other measures that can be indexed via multisensory tasks. These three previous findings included (a) the concurrent developmental trajectories of the TBW and susceptibility of multisensory illusions, (b) clinical populations with comorbid impairments in multisensory integration and temporal processing discussed below, and (c) overlapping neural architecture supporting the perception of multisensory illusions, perception of multisensory asynchrony, and multisensory temporal processing. Support for such a relationship can be found in the current data relating TBW width (specifically the width of the right TBW) with the rate of perceiving the McGurk effect. Individuals with narrower TBWs, and thus, those individuals for which sensory signals that are perceived as synchronous in the natural environment are more unique, showed increases in their rates of integration as indexed by the McGurk Effect.

Support for the idea that narrower TBWs are associated with a better ability to dissociate inputs that are not synchronous is found in the current data relating the TBW and the sound-induced flash illusion. This illusion contains an inherent asynchrony. Consequently, to perceive the illusion an individual must perceptually bind an audiovisual stimulus pair that is temporally asynchronous. Individuals with narrower windows were found to be less likely to perceive the illusion, suggesting that they are more likely to dissociate temporally asynchronous inputs. More specifically, the sound-induced flash illusion is comprised of the simultaneous presentation of a flash and a beep followed by one to three successive beeps. As such, there is an inherent asynchrony within the stimulus structure used to generate the illusion. Given the requirement that an individual must bind temporally nonsimultaneous events in order to perceive the illusion, an individual that has a wider TBW (and thus, is more likely to bind nonsimultaneous inputs) is more likely to perceive the illusion. In the case of the three- and four-beep conditions, the time interval between the first and last auditory onsets was 100 and 150 ms, respectively, outside of many individuals' TBW measured in this study and in previous studies using simple stimuli (Hirsh & Sherrick, 1961; Keetels & Vroomen, 2005; Zampini et al., 2005; Zampini et al., 2003). This was further supported by the finding that as the number of beeps increased from two to four, and thus, the inherent asynchrony in the sound-induced flash illusion increased, the correlation between the TBW and strength of the illusion became stronger.

The finding that the strength of the McGurk effect and sound-induced flash illusion were both significantly correlated with the right but not the left TBW is intriguing. In terms of multisensory events that occur in the natural environment, auditory and visual signals frequently have significant differences in their time of impact on neural circuits, a function of the different propagation times for their energies and the differences in the transduction and conduction times of signals in the different sensory systems. Thus, as a visual-auditory event happens further from the observer, the arrival time of auditory information (regardless of brain structure) is increasingly delayed. Given that real world events are specified in this way (i.e., where auditory signals are progressively delayed relative to visual signals), it should come as little surprise that the right TBW (in which auditory signals lag visual signals) is wider than the left, or that the right TBW is related to the strength of the integration of auditory and visual signals. Further evidence for the right TBW being more ethologically relevant is seen in the ability of perceptual learning to narrow the right but not left TBW (Powers et al., 2009).

In addition to these observations in adults, developmental studies also highlight the differences between the left and right halves of the temporal binding window. Early in development, the TBW is markedly wider than the mature TBW. As maturation progresses with exposure to the inherent temporal statistics of the environment, an asymmetry develops through a narrowing of the TBW, with narrowing of the left side preceding the right (Hillock et al., 2011). Furthermore, the development of the TBW appears to parallel the development of the perception of the McGurk effect. While children as young as four to five months have been reported to perceive the McGurk effect (Burnham & Dodd, 2004; Rosenblum, Schmuckler, & Johnson, 1997), younger children are less likely to perceive the effect than adolescents and adults (Massaro, 1984; Massaro, Thompson, Barron, & Laren, 1986; McGurk & MacDonald, 1976; Tremblay et al., 2007), and adolescents are less likely to perceive the effect than adults (Hillock et al., 2011). These parallels in the developmental trajectories of the TBW and the perception of multisensory illusions provide further support that the two may be mechanistically related.

The body of literature focusing on the neural architecture underlying these processes also suggests that there may be a relationship between the TBW and the strength of multisensory illusions. One of the primary brain regions implicated in the fusion of auditory and visual information is the posterior superior temporal sulcus (pSTS; Beauchamp, 2005; Beauchamp, Argall, Bodurka, Duyn, & Martin, 2004; Beauchamp, Lee, Argall, & Martin, 2004; James, Stevenson, & Kim, in press; James, Stevenson, & Kim, 2009; James, VanDerKlok, Stevenson, & James, 2011; Macaluso et al., 2004; Miller & D'Esposito, 2005; Nath & Beauchamp, 2011a; Stevenson, Geoghegan, & James, 2007; Stevenson & James, 2009; Stevenson, Kim, & James, 2009; Stevenson et al., 2011; Wallace & Murray, 2011; Werner & Noppeney, 2009, 2010). This region shows greater activation when individuals are presented with multisensory stimuli relative to unisensory stimuli, and the activity in the pSTS is specifically modulated when the auditory and visual presentations are bound into a unified percept. Furthermore, fMRI studies of the McGurk effect have also implicated the pSTS as a region in which neural activity is correlated with the perception (or lack thereof) of the McGurk effect (Beauchamp et al., 2010; Nath & Beauchamp, 2011b; Sekiyama et al., 2003). Likewise, a series of imaging studies of the sound-induced flash illusion have implicated the STS (in addition to occipital regions including V1) as a region in which neural activity is correlated with the perception (or lack thereof) of the audiovisual illusion (Bolognini et al., 2011; Watkins et al., 2007; Watkins et al., 2006).

Studies of the impact that temporal simultaneity has on multisensory processing have also implicated the pSTS. Thus, variations in the temporal disparity between cross-modal stimuli have been shown to modulate the amount of activity measured in pSTS in a number of

studies (Macaluso et al., 2004; Miller & D'Esposito, 2005; Stevenson et al., 2010; Stevenson et al., 2011). The findings that variation in both the temporal structure of multisensory stimuli as well as reports of multisensory illusions modulating neural activity in similar brain regions further suggest that a link between the width of the TBW and the strength of binding processes that are indexed by multisensory illusions.

Finally, a connection between the width of an individual's TBW and that person's ability to perceive multisensory illusions has also been seen in clinical populations in which impairments in multisensory processing have been seen, including autism spectrum disorders (ASD; Foss-Feig et al., 2010; Kwakye et al., 2011), dyslexia (Bastien-Toniazzo et al., 2009), and schizophrenia (Bleich-Cohen, Hendler, Kotler, & Strous, 2009; Stone et al., 2011; Williams, Light, Braff, & Ramachandran, 2010). In each of these conditions, sensory impairments include dysfunctions in perceptual fusion. Perhaps the strongest evidence has been seen in autism, in which individuals are often characterized by an impaired ability to combine multisensory information into a unified percept and, more specifically, show a weakened ability to perceive the McGurk effect (Mongillo et al., 2008; Taylor et al., 2010; Williams, Massaro, Peel, Bosseler, & Suddendorf, 2004). Recent evidence also shows that autistic individuals exhibit an atypically wide TBW (Foss-Feig et al., 2010; Kwakye et al., 2011). As in the evidence discussed above for typical populations, the relationship between the TBW and the ability to create unified multisensory perceptions in these clinical populations converges with neural evidence. Studies investigating the neural correlates of autism show irregularities in regions known to be involved with multisensory integration such as the pSTS (Boddaert et al., 2003; Boddaert et al., 2004; Boddaert & Zilbovicius, 2002; Gervais et al., 2004; Levitt et al., 2003; Pelphrey & Carter, 2008a, 2008b). Again, this finding of atypically wide TBWs being associated with decreases in the perception of multisensory illusions provides additional evidence for a mechanistic relationship between TBWs and strength of an individual's multisensory integration, and also suggests it may have clinical implications. This evidence has led multiple investigators to propose that the impairments in multisensory binding stem from deficits in temporal processing or that both deficits have a common etiology (Brunelle, Boddaert, & Zilbovicius, 2009; Kwakye et al., 2011; Stevenson et al., 2011; Zilbovicius et al., 2000). Given this relationship between temporal processing and multisensory integration, and the known behavioral and perceptual benefits of multisensory integration, perceptual training focused on improving multisensory temporal processing as has been carried out in TD individuals (Powers et al., 2009) may prove to be a useful means to improve sensory function, and the higher order processes dependent on multisensory binding, in these populations.

Outside of the relationship between multisensory illusions and the TBW, the measurement of the TBW produced several additional noteworthy findings. Our data replicated the commonly shown finding that the right side of the TBW is wider than the left side (Figure 2A). Our data also show that while the absolute values of these two sides are different, these widths remain proportionate across individuals as can be seen in the correlation between individuals' left and right TBWs (Figure 2B). This finding suggests that while there are notable differences in synchrony perception and temporal binding dependent upon which stimulus modality is presented first, the width of the total TBW reflects a single underlying cognitive process.

The data presented here provide strong supportive evidence that the temporal processing of multiple sensory signals and the merging of these signals into a single, unified perception, are highly related. Individuals with a narrower TBW will statistically encounter fewer occurrences of stimulus inputs that are perceived to be temporally synchronous and subsequently, perceived synchronous sensory signals will be more unique, leading to greater perceptual binding of synchronous inputs, and increased ability to dissociate asynchronous

inputs. This finding has strong relevance for the neural networks subserving multisensory processing, in that they suggest a common mechanistic link. In addition, the clinical implications of this work, coupled with previous demonstrations highlighting the plasticity of the TBW, suggest that through manipulations of the TBW one may be able to strengthen the ability of an individual to integrate information from multiple sensory modalities.

Acknowledgments

References

Andersen TS, Tiippana K, Sams M. Factors influencing audiovisual fission and fusion illusions. Cognitive Brain Research. 2004; 21:301–308. [PubMed: 15511646]

Bastien-Toniazzo M, Stroumza A, Cavé C. Audio-visual perception and integration in developmental dyslexia: An exploratory study using the McGurk effect. Current Psychology Letters. 2009; 25:1–14.

Beauchamp MS. Statistical criteria in FMRI studies of multisensory integration. Neuroinformatics. 2005; 3:93–113. [PubMed: 15988040]

Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A. Unraveling multisensory integration: Patchy organization within human STS multisensory cortex. Nature Neuroscience. 2004; 7:1190–1192.

Beauchamp MS, Lee KE, Argall BD, Martin A. Integration of auditory and visual information about objects in superior temporal sulcus. Neuron. 2004; 41:809–823. [PubMed: 15003179]

Beauchamp MS, Nath AR, Pasalar S. fMRI-Guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. Journal of Neuroscience. 2010; 30:2414–2417. [PubMed: 20164324]

Bleich-Cohen M, Hendler T, Kotler M, Strous RD. Reduced language lateralization in first-episode schizophrenia: An fMRI index of functional asymmetry. Psychiatry Research. 2009; 171:82–93. [PubMed: 19185468]

Boddaert N, Belin P, Chabane N, Poline JB, Barthelemy C, Mouren-Simeoni MC, Zilbovicius M. Perception of complex sounds: Abnormal pattern of cortical activation in autism. American Journal of Psychiatry. 2003; 160:2057–2060. [PubMed: 14594758]

Boddaert N, Chabane N, Gervais H, Good CD, Bourgeois M, Plumet MH, Zilbovicius M. Superior temporal sulcus anatomical abnormalities in childhood autism: A voxel-based morphometry MRI study. Neuroimage. 2004; 23:364–369. [PubMed: 15325384]

Boddaert N, Zilbovicius M. Functional neuroimaging and childhood autism. Pediatric Radiology. 2002; 32:1–7. [PubMed: 11819054]

Bolognini N, Rossetti A, Casati C, Mancini F, Vallar G. Neuromodulation of multisensory perception: A tDCS study of the sound-induced flash illusion. Neuropsychologia. 2011; 49:231–237. [PubMed: 21094177]

Brainard DH. The psychophysics toolbox. Spatial Vision. 1997; 10:433–436. [PubMed: 9176952]

Brunelle F, Boddaert N, Zilbovicius M. Autisme et imagerie ce´re´brale [Autism and brain imaging]. Bulletin De l'Académie Nationale De Me´decine. 2009; 193:287–297. discussion 297–288.

Burnham D, Dodd B. Auditory-visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. Developmental Psychobiology. 2004; 45:204–220. [PubMed: 15549685]

Colonius H, Diederich A. Multisensory interaction in saccadic reaction time: A time-window-of-integration model. Journal of Cognitive Neuroscience. 2004; 16:1000–1009. [PubMed: 15298787]

Conrey, BL.; Pisoni, DB. Detection of auditory-visual asynchrony in speech and nonspeech signals. In: Pisoni, DB., editor. Research on spoken language processing. Vol. Vol. 26. Bloomington, IN: Indiana University; 2004. p. 71-94.

Conrey BL, Pisoni DB. Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. Journal of the Acoustical Society of America. 2006; 119:4065–4073. [PubMed: 16838548]

Corey DP, Hudspeth AJ. Response latency of vertebrate hair cells. Biophysical Journal. 1979; 26:499–506. [PubMed: 318064]

Dhamala M, Assisi CG, Jirsa VK, Steinberg FL, Kelso JA. Multisensory integration for timing engages different brain networks. Neuroimage. 2007; 34:764–773. [PubMed: 17098445]

Diederich A, Colonius H. Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. Perception and Psychophysics. 2004; 66:1388–1404. [PubMed: 15813202]

Dixon NF, Spitz L. The detection of auditory visual desynchrony. Perception. 1980; 9:719–721. [PubMed: 7220244]

Erber NP. Auditory, visual, and auditory-visual recognition of consonants by children with normal and impaired hearing. Journal of Speech and Hearing Research. 1972; 15:413–422. [PubMed: 5047880]

Foss-Feig JH, Kwakye LD, Cascio CJ, Burnette CP, Kadivar H, Stone WL, Wallace MT. An extended multisensory temporal binding window in autism spectrum disorders. Experimental Brain Research. 2010; 203:381–389.

Foucher JR, Lacambre M, Pham BT, Giersch A, Elliott MA. Low time resolution in schizophrenia Lengthened windows of simultaneity for visual, auditory and bimodal stimuli. Schizophrenia Research. 2007; 97:118–127. [PubMed: 17884350]

Gaver WW. What in the world do we hear?: An ecological approach to auditory event perception. Ecological Psychology. 1993; 5:1–29.

Gervais H, Belin P, Boddaert N, Leboyer M, Coez A, Sfaello I, Zilbovicius M. Abnormal cortical voice processing in autism. Nature Neuroscience. 2004; 7:801–802.

Grant KW, Walden BE, Seitz PF. Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. Journal of the Acoustical Society of America. 1998; 103:2677–2690. [PubMed: 9604361]

Hairston WD, Burdette JH, Flowers DL, Wood FB, Wallace MT. Altered temporal profile of visual-auditory multisensory interactions in dyslexia. Experimental Brain Research. 2005; 166:474–480.

Hershenson M. Reaction time as a measure of intersensory facilitation. Journal of Experimental Psychology. 1962; 63:289–293. [PubMed: 13906889]

Hillock AR, Powers AR, Wallace MT. Binding of sights and sounds: Age-related changes in multisensory temporal processing. Neuropsychologia. 2011; 49:461–467. [PubMed: 21134385]

Hirsh IJ, Fraisse P. Simultanéitéet succession de stimuli hétérogénes [Simultaneous character and succession of heterogenous stimuli]. L'Année Psychologique. 1964; 64:1–19.

Hirsh IJ, Sherrick CE Jr. Perceived order in different sense modalities. Journal of Experimental Psychology. 1961; 62:423–432. [PubMed: 13907740]

James, TW.; Stevenson, RA. The use of fMRI to assess multisensory integration. In: Wallace, MH.; Murray, MM., editors. Frontiers in the neural basis of multisensory processes. London, UK: Taylor & Francis; 2012. p. 131-146.

James, TW.; Stevenson, RA.; Kim, S. Inverse effectiveness in multisensory processing. In: Stein, BE., editor. The new handbook of multisensory processes. Cambridge, MA: MIT Press; (in press)

James, TW.; Stevenson, RA.; Kim, S. Assessing multisensory integration with additive factors and functional MRI; Dublin, Ireland. Paper presented at the The International Society for Psychophysics; Oct. 2009

James TW, VanDerKlok RM, Stevenson RA, James KH. Multisensory perception of action in posterior temporal and parietal cortices. Neuropsychologia. 2011; 49:108–114. [PubMed: 21036183]

Keetels M, Vroomen J. The role of spatial disparity and hemifields in audio-visual temporal order judgments. Experimental Brain Research. 2005; 167:635–640.

King AJ, Palmer AR. Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. Experimental Brain Research. 1985; 60:492–500.

Kwakye LD, Foss-Feig JH, Cascio CJ, Stone WL, Wallace MT. Altered auditory and multisensory temporal processing in autism spectrum disorders. Frontiers in Integrative Neuroscience. 2011; 4:129. [PubMed: 21258617]

Lamb TD, Pugh EN Jr. A quantitative account of the activation steps involved in phototransduction in amphibian photoreceptors. Journal of Physiology. 1992; 449:719–758. [PubMed: 1326052]

Lennie P. The physiological basis of variations in visual latency. Vision Research. 1981; 21:815–824. [PubMed: 7314459]

Levitt JG, Blanton RE, Smalley S, Thompson PM, Guthrie D, McCracken JT, Toga AW. Cortical sulcal maps in autism. Cerebral Cortex. 2003; 13:728–735. [PubMed: 12816888]

Lovelace CT, Stein BE, Wallace MT. An irrelevant light enhances auditory detection in humans: A psychophysical analysis of multisensory integration in stimulus detection. Cognitive Brain Research. 2003; 17:447–453. [PubMed: 12880914]

Macaluso E, George N, Dolan R, Spence C, Driver J. Spatial and temporal factors during processing of audiovisual speech: A PET study. Neuroimage. 2004; 21:725–732. [PubMed: 14980575]

Massaro DW. Children's perception of visual and auditory speech. Child Development. 1984; 55:1777–1788. [PubMed: 6510054]

Massaro, DW. Perceiving talking faces: From speech perception to a behavioral principle. Cambridge, MA: MIT Press; 1998.

Massaro, DW. From multisensory integration to talking heads and language learning. In: Calvert, G.; Spence, C.; Stein, BE., editors. The handbook of multisensory processes. Cambridge, MA: MIT Press; 2004. p. 153-176.

Massaro DW, Thompson LA, Barron B, Laren E. Developmental changes in visual and auditory contributions to speech perception. Journal of Experimental Child Psychology. 1986; 41:93–113. [PubMed: 3950540]

McGurk H, MacDonald J. Hearing lips and seeing voices. Nature. 1976; 264:746–748. [PubMed: 1012311]

Meredith MA, Nemitz JW, Stein BE. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. Journal of Neuroscience. 1987; 7:3215–3229. [PubMed: 3668625]

Meredith MA, Stein BE. Spatial factors determine the activity of multisensory neurons in cat superior colliculus. Brain Research. 1986a; 365:350–354. [PubMed: 3947999]

Meredith MA, Stein BE. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. Journal of Neurophysiology. 1986b; 56:640–662. [PubMed: 3537225]

Miller LM, D'Esposito M. Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. Journal of Neuroscience. 2005; 25:5884–5893. [PubMed: 15976077]

Mongillo EA, Irwin JR, Whalen DH, Klaiman C, Carter AS, Schultz RT. Audiovisual processing in children with and without autism spectrum disorders. Journal of Autism and Developmental Disorders. 2008; 38:1349–1358. [PubMed: 18307027]

Nath AR, Beauchamp MS. Dynamic changes in superior temporal sulcus connectivity during perception of noisy audiovisual speech. Journal of Neuroscience. 2011a; 31:1704–1714. [PubMed: 21289179]

Nath AR, Beauchamp MS. A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. Neuroimage. 2011b; 59:781–787. [PubMed: 21787869]

Nelson WT, Hettinger LJ, Cunningham JA, Brickman BJ, Haas MW, McKinley RL. Effects of localized auditory information on visual target detection performance using a helmetmounted display. Human Factors. 1998; 40:452–460. [PubMed: 9849103]

Noesselt T, Rieger JW, Schoenfeld MA, Kanowski M, Hinrichs H, Heinze HJ, Driver J. Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. Journal of Neuroscience. 2007; 27:11431–11441. [PubMed: 17942738]

Pearl D, Yodashkin-Porat D, Katz N, Valevski A, Aizenberg D, Sigler M, Kikinzon L. Differences in audiovisual integration, as measured by McGurk phenomenon, among adult and adolescent

patients with schizophrenia and age-matched healthy control groups. Comprehensive Psychiatry. 2009; 50:186–192. [PubMed: 19216897]

Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision. 1997; 10:437–442. [PubMed: 9176953]

Pelphrey KA, Carter EJ. Brain mechanisms for social perception: Lessons from autism and typical development. Annals of the New York Academy of Sciences. 2008a; 1145:283–299. [PubMed: 19076404]

Pelphrey KA, Carter EJ. Charting the typical and atypical development of the social brain. Development and Psychopathology. 2008b; 20:1081–1102. [PubMed: 18838032]

Pöppel E, Schill K, von Steinbüchel N. Sensory integration within temporally neutral systems states: A hypothesis. Naturwissenschaften. 1990; 77:89–91. [PubMed: 2314478]

Powers AR 3rd, Hillock AR, Wallace MT. Perceptual training narrows the temporal window of multisensory binding. Journal of Neuroscience. 2009; 29:12265–12274. [PubMed: 19793985]

Quinto L, Thompson WF, Russo FA, Trehub SE. A comparison of the McGurk effect for spoken and sung syllables. Attention, Perception and Psychophysics. 2010; 72:1450–1454.

Roach NW, Heron J, Whitaker D, McGraw PV. Asynchrony adaption reveals neural population code for audio-visual timing. Proceedings of the Royal Society. 2011; 278:9.

Rosenblum LD, Schmuckler MA, Johnson JA. The McGurk effect in infants. Perception and Psychophysics. 1997; 59:347–357. [PubMed: 9136265]

Schall S, Quigley C, Onat S, Konig P. Visual stimulus locking of EEG is modulated by temporal congruency of auditory stimuli. Experimental Brain Research. 2009; 198:137–151.

Seitz AR, Nanez JE, Holloway SR, Watanabe T. Perceptual learning of motion leads to faster flicker perception. PLoS ONE. 2006; 1:e28. [PubMed: 17183655]

Sekiyama K, Kanno I, Miura S, Sugita Y. Auditory-visual speech perception examined by fMRI and PET. Neuroscience Research. 2003; 47:277–287. [PubMed: 14568109]

Senkowski D, Talsma D, Grigutsch M, Herrmann CS, Woldorff MG. Good times for multisensory integration: Effects of the precision of temporal synchrony as revealed by gamma-band oscillations. Neuropsychologia. 2007; 45:561–571. [PubMed: 16542688]

Shams L, Kamitani Y, Shimojo S. Illusions. What you see is what you hear. Nature. 2000; 408:788. [PubMed: 11130706]

Spence C, Baddeley R, Zampini M, James R, Shore DI. Multisensory temporal order judgments: When two locations are better than one. Perception and Psychophysics. 2003; 65:318–328. [PubMed: 12713247]

Stein, B.; Meredith, MA. The merging of the senses. Boston, MA: MIT Press; 1993.

Stein BE, Wallace MT. Comparisons of cross-modality integration in midbrain and cortex. Progress in Brain Research. 1996; 112:289–299. [PubMed: 8979836]

Stevenson RA, Altieri NA, Kim S, Pisoni DB, James TW. Neural processing of asynchronous audiovisual speech perception. Neuroimage. 2010; 49:3308–3318. [PubMed: 20004723]

Stevenson RA, Geoghegan ML, James TW. Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. Experimental Brain Research. 2007; 179:85–95.

Stevenson RA, James TW. Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. Neuroimage. 2009; 44:1210–1223. [PubMed: 18973818]

Stevenson RA, Kim S, James TW. An additive-factors design to disambiguate neuronal and areal convergence: Measuring multisensory interactions between audio, visual, and haptic sensory streams using fMRI. Experimental Brain Research. 2009; 198:183–194.

Stevenson RA, VanDerKlok RM, Pisoni DB, James TW. Discrete neural substrates underlie complementary audiovisual speech integration processes. Neuroimage. 2011; 55:1339–1345. [PubMed: 21195198]

Stevenson RA, Zemtsov RK, Wallace MT. (under review). Consistency in individual's multisensory temporal binding windows across task and stimulus complexity. Experimental Brain Research.

Stone DB, Urrea LJ, Aine CJ, Bustillo JR, Clark VP, Stephen JM. Unisensory processing and multisensory integration in schizophrenia: A high-density electrical mapping study. Neuropsychologia. 2011

Summerfield, Q. Some preliminaries to a comprehensive account of audio-visual speech perception. In: Dodd, B.; Campbell, BA., editors. Hearing by eye: The psychology of lip reading. London, UK: Erlbaum; 1987. p. 3-52.

Talsma D, Senkowski D, Woldorff MG. Intermodal attention affects the processing of the temporal alignment of audiovisual stimuli. Experimental Brain Research. 2009; 198:313–328.

Taylor N, Isaac C, Milne E. A comparison of the development of audiovisual integration in children with autism spectrum disorders and typically developing children. Journal of Autism and Developmental Disorders. 2010; 40:1403–1411. [PubMed: 20354776]

Tremblay C, Champoux F, Voss P, Bacon BA, Lepore F, Theoret H. Speech and non-speech audio-visual illusions: A developmental study. PLoS One. 2007; 2:e742. [PubMed: 17710142]

van Atteveldt NM, Formisano E, Blomert L, Goebel R. The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. Cerebral Cortex. 2007; 17:962–974. [PubMed: 16751298]

Van der Burg E, Cass J, Olivers CN, Theeuwes J, Alais D. Efficient visual search from synchronized auditory signals requires transient audiovisual events. PLoS One. 2010; 5:e10664. [PubMed: 20498844]

van Wassenhove V, Grant KW, Poeppel D. Temporal window of integration in auditory-visual speech perception. Neuropsychologia. 2007; 45:598–607. [PubMed: 16530232]

Vatakis A, Spence C. Audiovisual synchrony perception for music, speech, and object actions. Brain Research. 2006; 1111:134–142. [PubMed: 16876772]

Vroomen J, Keetels M. Perception of intersensory synchrony: A tutorial review. Attention, Perception and Psychophysics. 2010; 72:871–884.

Wallace, MH.; Murray, MM., editors. Frontiers in the neural basis of multisensory processes. London, UK: Taylor & Francis; 2011.

Wallace MT, Roberson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA. Unifying multisensory signals across time and space. Experimental Brain Research. 2004; 158:252–258.

Watkins S, Shams L, Josephs O, Rees G. Activity in human V1 follows multisensory perception. Neuroimage. 2007; 37:572–578. [PubMed: 17604652]

Watkins S, Shams L, Tanaka S, Haynes JD, Rees G. Sound alters activity in human V1 in association with illusory visual perception. Neuroimage. 2006; 31:1247–1256. [PubMed: 16556505]

Werner S, Noppeney U. Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. Cerebral Cortex. 2009; 20:1829–1842. [PubMed: 19923200]

Werner S, Noppeney U. Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. Journal of Neuroscience. 2010; 30:2662–2675. [PubMed: 20164350]

Wilkinson LK, Meredith MA, Stein BE. The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. Experimental Brain Research. 1996; 112:1–10.

Williams JH, Massaro DW, Peel NJ, Bosseler A, Suddendorf T. Visual-auditory integration during speech imitation in autism. Research in Developmental Disabilities. 2004; 25:559–575. [PubMed: 15541632]

Williams LE, Light GA, Braff DL, Ramachandran VS. Reduced multisensory integration in patients with schizophrenia on a target detection task. Neuropsychologia. 2010; 48:3128–3136. [PubMed: 20600181]

Zampini M, Guest S, Shore DI, Spence C. Audio-visual simultaneity judgments. Perception and Psychophysics. 2005; 67:531–544. [PubMed: 16119399]

Zampini M, Shore DI, Spence C. Audiovisual temporal order judgments. Experimental Brain Research. 2003; 152:198–210.

Zampini M, Shore DI, Spence C. Audiovisual prior entry. Neuroscience Letters. 2005; 381:217–222. [PubMed: 15896473]

Zilbovicius M, Boddaert N, Belin P, Poline JB, Remy P, Mangin JF, Samson Y. Temporal lobe dysfunction in childhood autism: A PET study. Positron emission tomography. American Journal of Psychiatry. 2000; 157:1988–1993. [PubMed: 11097965]
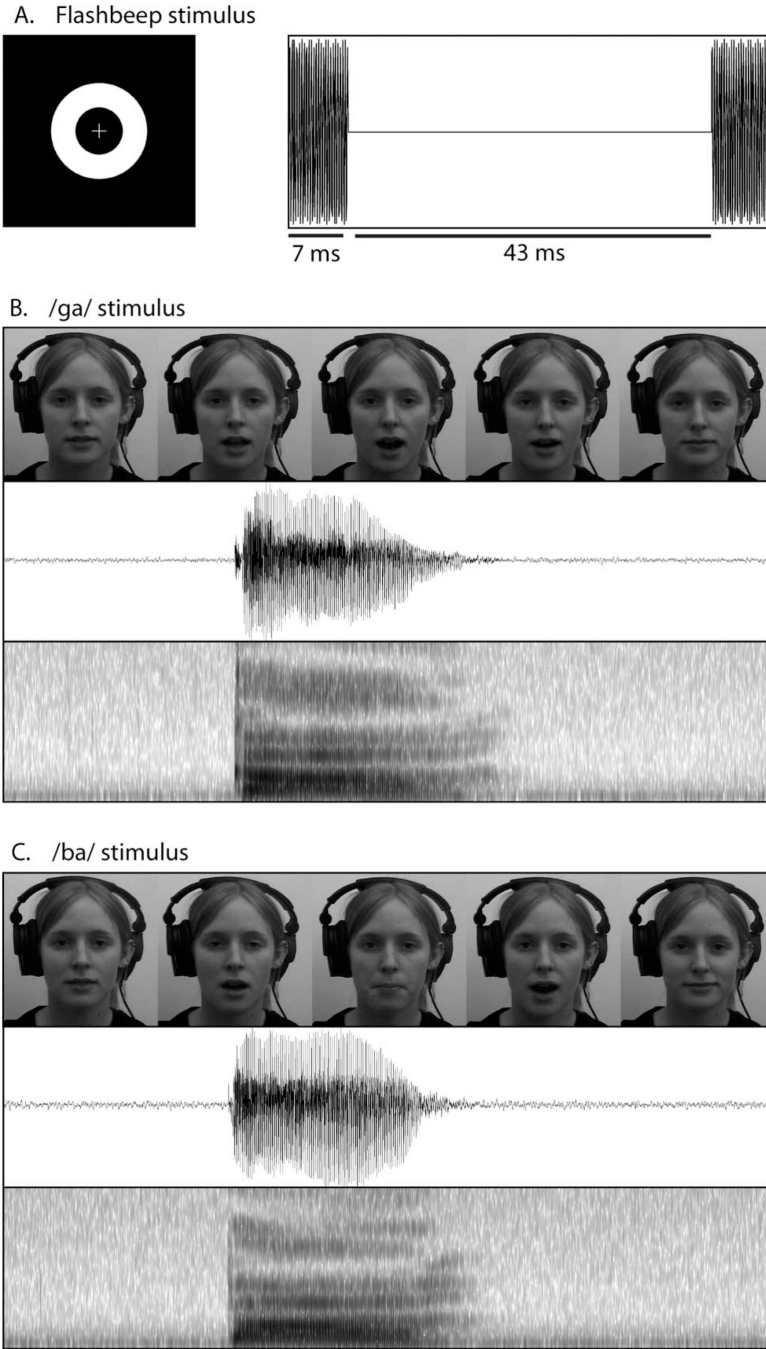
A. Flashbeep stimulus

B. /ga/ stimulus

C. /ba/ stimulus

Figure 1.
Stimuli used for the simultaneity judgment, sound-induced flash illusion and McGurk tasks.
A. The stimuli used for the simultaneity judgment task and the sound-induced flash illusion task are identical, and include the illumination of a white annulus (left) and a pure-tone beep (right). B and C. Frames from the visual stimuli as well as waveforms and spectrograms of the auditory stimuli used for the two syllables (/ga/ and /da/) in the McGurk task.
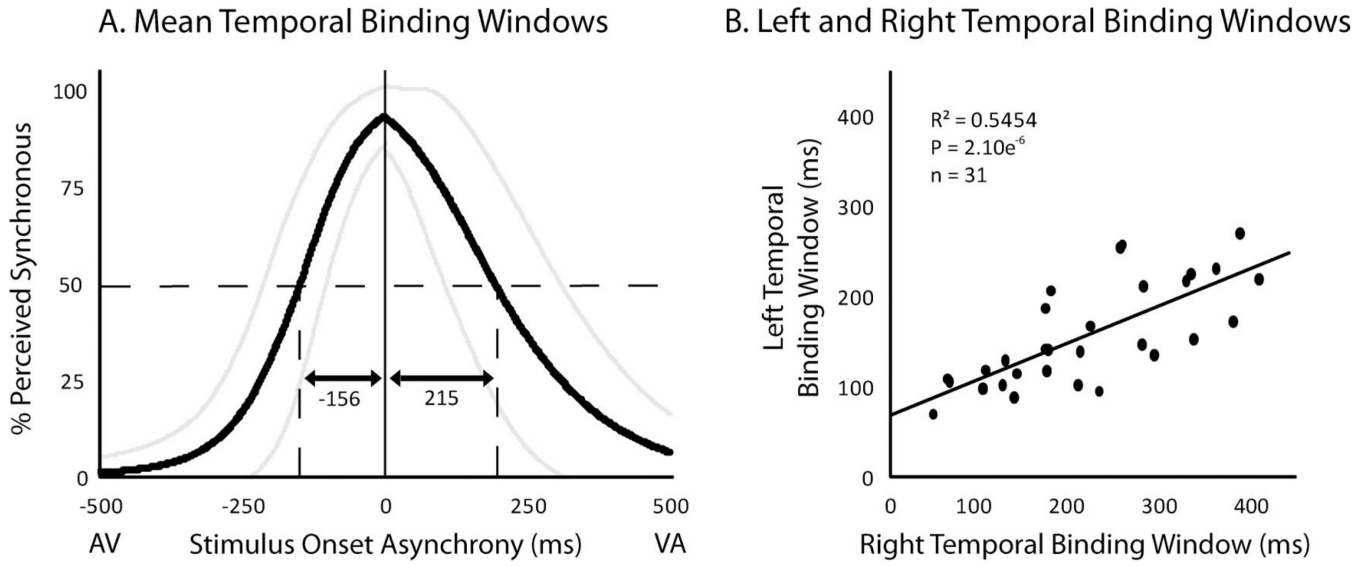
A. Mean Temporal Binding Windows

B. Left and Right Temporal Binding Windows

Figure 2.
Temporal Binding Windows defined using the simultaneity judgment task. A. Group mean curves used to derive each individual's temporal binding windows (standard deviations shown in gray). Note that the total window is comprised of individually defined and fitted left and right windows. B. Correlation in the size of each individual's left and right temporal binding windows.
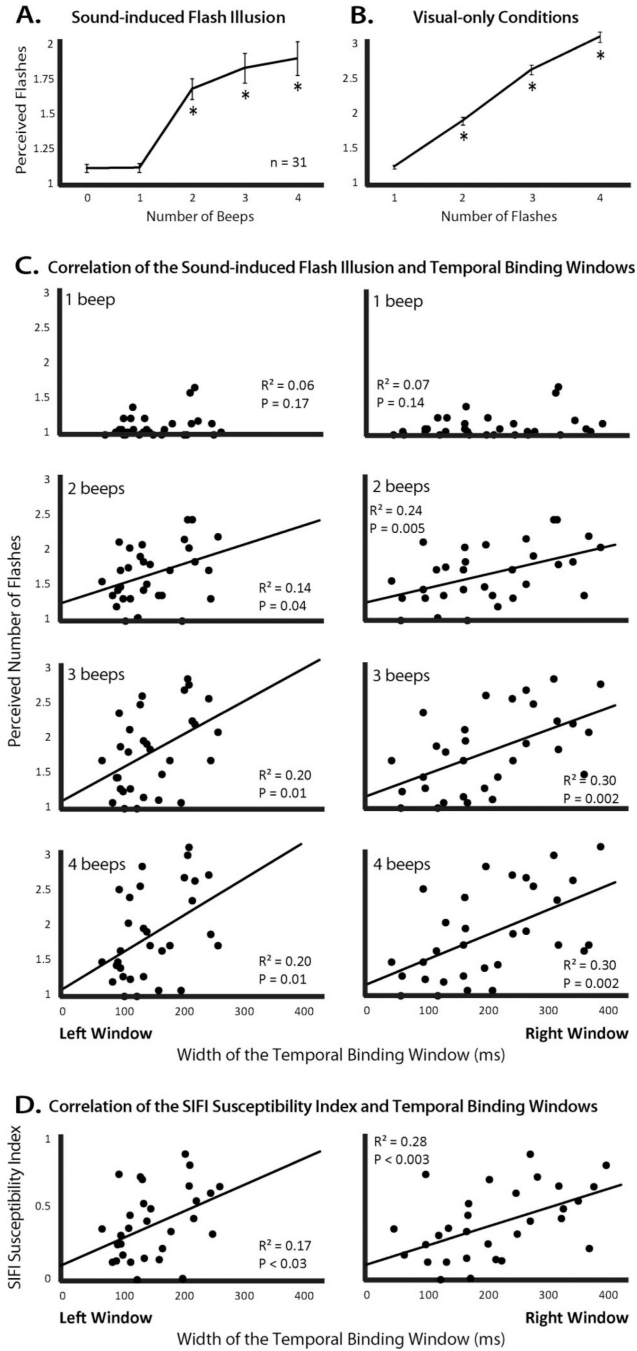
Figure 3.
Sound-induced flash illusion and its relationship to the TBW. A. Group mean responses show a rise in the perceived number of flashes as a function of the number of beeps. B. Control trials show a rise in the number of perceived flashes when the number of visual stimuli is increased. C. Correlations between the strength of each individual's perceived illusion (i.e., perceived number of flashes) and left and right TBW width for the 1, 2, 3, and 4 beep conditions. Note the presence of significant correlations for each of the illusory (i.e., 2 beep) conditions. D. Correlation between individuals' sound-induced flash illusion susceptibility index and their left and right temporal binding window.
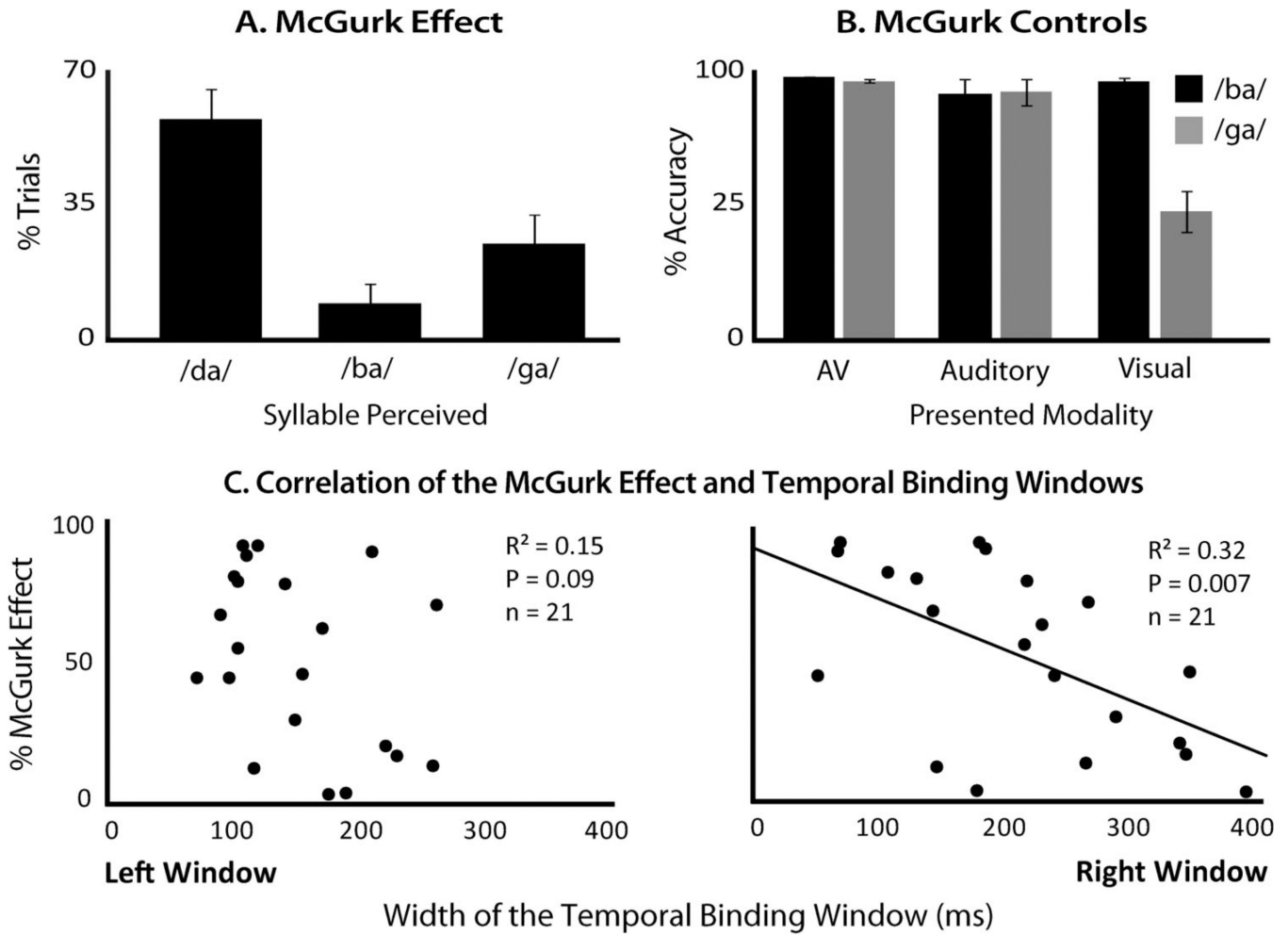
Figure 4.
McGurk Effect and its relationship to the TBW. A. Group mean responses to the
incongruent pairing of auditory /ba/ with visual /ga/ show a high incidence of reports of the
illusory fused percept /da/ B. Group mean responses to the congruent multisensory and the
individual unisensory components. C. Correlation of each individual's McGurk effect score
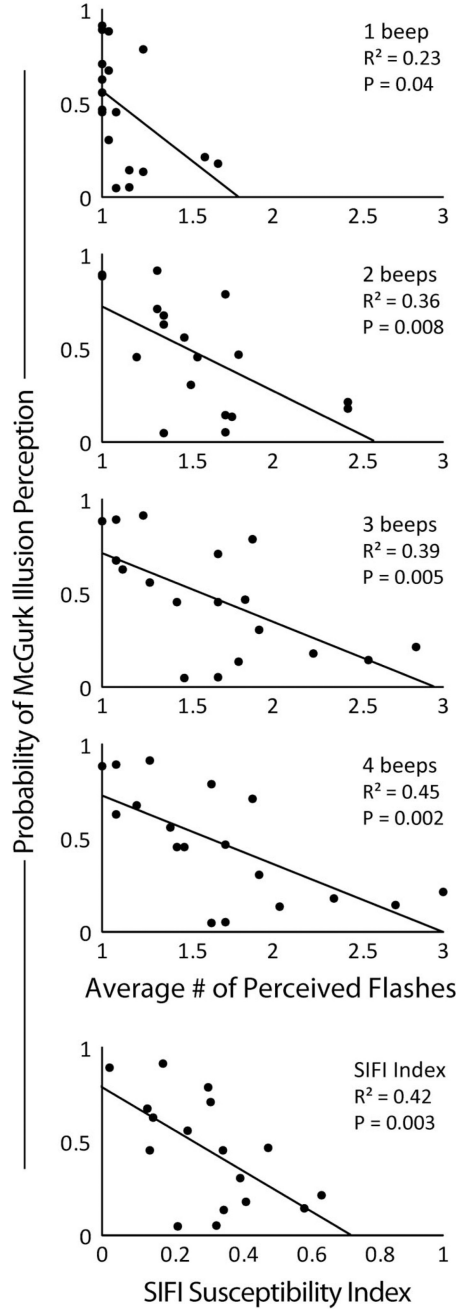(i.e., % of fusions) with their left and right temporal binding window.

Figure 5.
Relationship between the McGurk Effect sound-induced flash illusion. Individual's rate of perception of the McGurk Effect were inversely correlated with the strength of their perception of the sound-induced flash illusion in conditions with 2–4 beeps, as well as with the sound-induced flash illusion susceptibility index.

Table 1

Sound-Induced Flash Illusion Responses

# beeps	Mean (SEM) perceived flashes	from 1 beep		from 2 beeps		from 3 beeps		from 4 beeps	
		t	p	t	p	t	p	t	p
0	1.12 (0.03)	0.08	n.s.	9.95	$3.33e^{-10}$	7.77	$3.98e^{-8}$	7.41	$7.20e^{-8}$
1	1.12 (0.03)	—	—	9.41	$1.87e^{-10}$	7.49	$2.40e^{-8}$	7.33	$3.63e^{-8}$
2	1.66 (0.07)	9.41	$1.87e^{-10}$	—	—	2.71	0.02	3.03	0.005
3	1.81 (0.10)	7.49	$2.40e^{-8}$	2.71	0.02	—	—	2.24	0.04
4	1.87 (0.12)	7.33	$3.63e^{-8}$	3.03	0.005	2.24	0.04	—	—

Table 2

Responses to Visual-Only Control Trials

# flashes	Mean (SEM) perceived flashes	from 1 flash		from 2 flashes		from 3 flashes		from 4 flashes	
		t	p	t	p	t	p	t	p
1	1.12 (0.03)	—	—	12.04	$5.04e^{-13}$	20.37	$4.07e^{-19}$	21.93	$5.05e^{-20}$
2	1.74 (0.05)	12.04	$5.04e^{-13}$	—	—	10.43	$2.49e^{-11}$	15.06	$2.99e^{-15}$
3	2.43 (0.06)	20.37	$4.07e^{-19}$	10.43	$2.49e^{-11}$	—	—	11.73	$2.56e^{-12}$
4	2.87 (0.07)	21.93	$5.05e^{-20}$	15.06	$2.99e^{-15}$	11.73	$2.56e^{-12}$	—	—