Multisensory simultaneity judgment and proximity to the body

Jean-Paul Noel

Laboratory of Cognitive Neuroscience, Brain Mind Institute, School of Life Sciences, Ecole Polytéchnique Fédérale de Lausanne, Lausanne, Switzerland
Center for Neuroprosthetics, School of Life Sciences, Ecole Polytéchnique Fédérale de Lausanne, Lausanne, Switzerland
Vanderbilt Brain Institute, Vanderbilt University Medical Center, Vanderbilt University, Nashville, TN, USA ✉

Marta Łukowska

Laboratory of Cognitive Neuroscience, Brain Mind Institute, School of Life Sciences, Ecole Polytéchnique Fédérale de Lausanne, Lausanne, Switzerland
Center for Neuroprosthetics, School of Life Sciences, Ecole Polytéchnique Fédérale de Lausanne, Lausanne, Switzerland
Consciousness Lab, Institute of Psychology, Jagiellonian University, Krakow, Poland ✉

Mark Wallace

Vanderbilt Brain Institute, Vanderbilt University Medical Center, Vanderbilt University, Nashville, TN, USA
Department of Hearing and Speech, Vanderbilt University Medical Center, Vanderbilt University, Nashville, TN, USA
Department of Psychology, Vanderbilt University, Nashville, TN, USA ✉

Andrea Serino

Laboratory of Cognitive Neuroscience, Brain Mind Institute, School of Life Sciences, Ecole Polytéchnique Fédérale de Lausanne, Lausanne, Switzerland
Center for Neuroprosthetics, School of Life Sciences, Ecole Polytéchnique Fédérale de Lausanne, Lausanne, Switzerland ✉

The integration of information across different sensory modalities is known to be dependent upon the statistical characteristics of the stimuli to be combined. For example, the spatial and temporal proximity of stimuli are important determinants with stimuli that are close in space and time being more likely to be bound. These multisensory interactions occur not only for singular points in space/time, but over "windows" of space and time that likely relate to the ecological statistics of real-world stimuli. Relatedly, human psychophysical work has demonstrated that individuals are highly prone to judge multisensory stimuli as co-occurring over a wide range of time—a so-called simultaneity window (SW). Similarly, there exists a spatial representation of peripersonal space (PPS) surrounding the body in which stimuli related to the body and to external events occurring near the body are highly likely to be jointly processed. In the current study, we sought to examine the interaction between these temporal and spatial dimensions of multisensory representation by measuring the SW for

audiovisual stimuli through proximal–distal space (i.e., PPS and extrapersonal space). Results demonstrate that the audiovisual SWs within PPS are larger than outside PPS. In addition, we suggest that this effect is likely due to an automatic and additional computation of these multisensory events in a body-centered reference frame. We discuss the current findings in terms of the spatiotemporal constraints of multisensory interactions and the implication of distinct reference frames on this process.

Introduction

At the neurophysiological level, much work has gone into characterizing the response properties and profiles of multisensory neurons when presented with stimuli from multiple sensory modalities. These studies have revealed striking nonlinearities in the multisensory responses of these neurons (Benevento, Fallon, Davis, & Rezak, 1977; Bruce, Desimone, & Gross, 1981; Hikosaka, Iwai, Saito, & Tanaka, 1988; Meredith, Nemitz, & Stein, 1987; Meredith & Stein, 1986a, 1986b; Rauschecker & Korte, 1993; Wallace, Meredith, & Stein, 1992) and have elucidated a set of integrative "principles" by which these neurons operate (B. Stein & Stanford, 2008; B. E. Stein & Meredith, 1993). These principles revolve around the statistical features of the stimuli that are to be combined, including their spatial and temporal relationships to one another and their relative effectiveness. In their simplest terms, these principles state that stimuli that are spatially and temporally proximate and weakly effective (when presented on their own) give rise to the largest proportionate enhancements of response when combined. As stimuli are separated in space and/or time and as they become increasingly effective, these proportionate multisensory gains decline (Calvert, Spence, & Stein, 2004; Murray & Wallace, 2012).

Although first established at the level of the single neuron, these principles have also been shown to apply to indices of activity in larger neuronal ensembles, such as scalp (Cappe, Thelen, Romei, Thut, & Murray, 2012) and intracranial EEG (Quinn et al., 2014), fMRI (Miller & D'Esposito, 2005), and PET scanning (Macaluso, George, Dolan, Spence, & Driver, 2004), as well as to indices of animal and human behavior. Examples of multisensory-mediated benefits in the behavioral and perceptual realms include enhanced target detection (Frassinetti, Bolognini, & Ladavas, 2002; Lovelace, Stein, & Wallace, 2003), facilitated target localization (Nelson et al., 1998; Wilkinson, Meredith, & Stein, 1996), and speeded reaction times (Diederich & Colonius, 2004; Frens, Van Opstal, & Van der Willigen, 1995; Nozawa, Reuter-Lorenz, & Hughes, 1994) in response to spatiotemporally proxi-

mate and weak stimulus pairings (although see Stanford & Stein, 2007, and Spence, 2013, for relevant discussions on the applicability of the so-called principles of multisensory integration to neuroimaging and behavioral indices in humans).

Despite the large (and growing) number of studies detailing how stimulus factors shape multisensory interactions, the vast majority of previous reports have looked at each of these factors in isolation or have employed a set of features (e.g., temporal processing) in order to study the other (e.g., spatial) but have not examined their interaction. For example, by and large, studies structured to examine the impact of spatial location and correspondence between stimuli have only manipulated stimuli in the spatial dimension(s) while presenting stimuli at the same time. Temporal factors, such as the temporal discrepancy between consecutively presented stimuli, have routinely been employed to measure spatial aspects of multisensory interactions (e.g., tactile temporal order judgment tasks in which limbs are crossed/uncrossed in order to study the realignment of somatotopic and external reference frames; see Heed & Azañón, 2014, for review). Similarly, temporal factors have equally been regularly employed not only to measure, but also to manipulate spatial multisensory correspondences (e.g., synchronous administration of visuo-tactile stimulation on a rubber hand in order to elicit an illusory feeling of ownership). On the other hand, studies structured to examine the temporal dynamics underlying multisensory interactions have typically done so at a single spatial location. Although informative, it must be pointed out that such manipulations typically differ from real environmental circumstances, in which dynamic stimuli are changing frequently (and non-independently) in effectiveness and in their spatial and temporal relationships. Some recent studies have begun to examine interactions between space and time on multisensory processing at the level of the individual neuron, giving rise to the notion of spatiotemporal receptive fields (Royal, Carriere, & Wallace, 2009; Royal, Krueger, Fister, & Wallace 2010). However, fewer have systematically examined these issues in the domain of human performance (although see Slutsky & Recanzone, 2001; Stevenson, Fister, Barnet, Nidiffer, & Wallace, 2012a; Zampini, Guest, Shore, & Spence, 2005, for exceptions). In addition, those few studies that have attempted to examine multisensory spatiotemporal interactions have almost exclusively done so in azimuth and elevation and have revealed equivocal results. For example, examination of the effects of spatial alignment on temporal perception have shown cases of benefit, other cases of a deterioration in performance, and still others with no obvious interaction (Chen & Vroomen, 2013; Slutsky & Recanzone,

2001; Spence, Shore, & Klein, 2001; Vroomen & Keetels, 2010; Zampini, Shore, & Spence, 2003).

Here, thus, we propose to study interactions between spatial and temporal factors in guiding multisensory-mediated behaviors but to do so not by focusing on the spatial alignment between the sensory cues themselves, but rather by altering the spatial arrangement of audiovisual stimuli in space with respect to the perceiver's body (e.g., distance between stimuli and subjects).

The vast majority of studies detailing multisensory interactions in depth have focused on audio- or visuo-tactile interactions delineating the boundaries of peripersonal space (PPS). Namely, neuropsychological (Farne & Ladavas, 2002) as well as psychophysical (Canzoneri, Magosso, & Serino, 2012; Canzoneri, Marzolla, Amoresano, Verni, & Serino, 2013a; Canzoneri, Ubaldi, et al., 2013b; Galli, Noel, Canzoneri, Blanke, & Serino, 2015; Noel et al., 2014; Noel, Pfeiffer, Blanke, & Serino, 2015a; Serino et al., 2015) studies have demonstrated that visual and auditory stimuli most strongly interact with tactile stimulation when these are presented close to, as opposed to far from, the body. As highlighted by Van der Stoep, Nijboer, Van der Stigchel, and Spence (2015a), however, these studies do not unambiguously speak to the existence of a spatial rule of multisensory integration in depth as tactile stimulation necessarily happens on the body. Thus, by presenting visual or auditory stimuli close to the body, researchers are both varying the distance between stimuli and observer as well as the spatial discrepancy between the locations of the stimuli (audio-tactile or visuo-tactile).

In contrast, Van der Stoep, Van der Stigchel, and Nijboer (2015b) as well as Ngo and Spence (2010) report audiovisual multisensory effects in depth. Particularly, they demonstrate a spatial cueing effect (on detection and visual search, respectively) in the proximal–distal dimension. That is, the cueing was exclusively documented when cue and target were presented from the same distance. Additionally, Van der Stoep, Van der Stigchel, Nijboer, and Van der Smagt (2015c) have reported that detection of audiovisual stimuli was most facilitated by a decrease in intensity for stimuli that were presented in extrapersonal space. The same decrease in intensity for stimuli that were presented close to the body did not result in a similar facilitation. Thus, this study nicely documents the interaction between several of the traditional multisensory "principles" (the spatial rule and inverse effectiveness, in this case) as well as points toward the fact that audiovisual interactions may be stronger when presented further rather than closer for the observer.

In the present study, we turn our attention to the temporal aspects of multisensory processing and its relationship with observer–stimuli distance. Both physiological and behavioral studies have established that the temporal and spatial constraints of multisensory interactions are bounded by specific "windows" within which these interactions are highly likely. In the temporal domain, this window can be captured in the construct of the simultaneity window (SW), which is derived from perceptual decisions on the relative timing of stimuli using tasks such as simultaneity judgments (Noel, Wallace, Orchard-Mills, Alais, & Van der Burg, 2015b; Stone et al., 2001; Wallace & Stevenson, 2014). Although such tasks do not represent de facto measures of multisensory integration, they can serve as useful proxies for the temporal interval over which auditory and visual stimuli can influence one another's processing as performance on these tasks has been shown to strongly relate on a subject-per-subject basis with integration of multisensory stimuli (Stevenson, Zemtsov, & Wallace, 2012b). In the spatial domain, the concept of a window is best encapsulated by the notion of receptive fields. In terms of distance to the subject, the extension of multisensory receptive fields has been described by the notion of PPS, i.e., the region adjacent to and surrounding the body where visual or auditory stimuli related to external objects more strongly interact with somatosensory processing on the body (Graziano & Cooke, 2006; Rizzolatti, Fadiga, Fogassi, & Gallese, 1997). The notion of PPS was initially proposed by neurophysiological studies describing the properties of a particular ensemble of neurons with bimodal (tactile, auditory, visual) receptive fields anchored to specific body parts and extending over a limited sector of space surrounding the same body part (Duhamel et al., 1998; Fogassi et al., 1996; Graziano & Cooke, 2006). In humans, this concept is captured by the well-established finding that tactile processing is more strongly affected by visual or auditory inputs presented near to than far from the body, i.e., within as compared to outside the PPS (Brozzoli, Gentile, & Ehrsson, 2012; Brozzoli, Gentile, Petkova, & Ehrsson, 2011; Gentile et al., 2013; Serino, Canzoneri, & Avenanti, 2011). Despite the fact that many studies have been structured to examine the SW and PPS, no specific work has been carried out to examine both in the same study and to explore the interrelationships between them.

In order to address this question, in the current series of experiments, we presented subjects with audiovisual events with different temporal relationships (i.e., stimulus onset asynchronies ranging from 0 ms to ±250 ms) at different spatial distances from the observer; either close (30 cm, arguably within PPS) or far (100 cm, arguably outside PPS). Importantly, auditory and visual stimuli were always colocalized. We used a simultaneity judgment (SJ) task to measure the width of subjects' SW. We found that SWs are larger when stimuli are presented closer rather than farther

from participants. We then investigated this result in terms of the reference frame transformation necessary to code signals from different sensory modalities, and thus originally processed in different native frames of reference, into a common, body-centered, reference frame.

Experiment 1

Methods

Participants

Twenty-one (four females, mean age 23.4 ± 2.4, range 19–31) right-handed students from the Ecole Polytechnique Federale de Lausanne took part in this experiment. All participants reported normal auditory acuity and had normal or corrected-to-normal eyesight. All participants gave their informed consent to take part in this study, which was approved by the local ethics committee and the Brain Mind Institute Ethics Committee for Human Behavioral Research at EPFL and were reimbursed for their participation.

Stimuli and apparatus

Visual and auditory stimuli were controlled via a purpose-made Arduino[TM] microcontroller (http://arduino.cc, refresh rate 10 KHz) and driven by in-house experimental software (ExpyVR, http://lnco.epfl.ch/expyvr, direct serial port communication with microcontroller). Visual stimuli were presented by means of a red LED (7000 mcd, wavelength 640 nm, 34° radiancy angle), and auditory stimuli were generated by the activation of a piezo speaker (75 dB at 0.3 m, 3.0 kHz). Two audiovisual devices were built by assembling the auditory and visual stimuli into a 5 cm × 3 cm × 1 cm opaque rectangular box. Both visual and auditory stimuli had a duration of 10 ms and were presented within a range of stimulus onset asynchronies (SOAs) that included 0 ms, ±50 ms, ±100 ms, ±150 ms, ±200ms, or ±250 ms. Arbitrarily, positive SOAs indicate conditions in which visual stimuli preceded auditory stimuli. Participants responses were made via button press on a game pad (Xbox 360 controller, 125 Hz sampling rate, Microsoft, Redmond, WA), which was held in the right hand.

Procedure

The experiment took place in a dimly lit and sound-controlled room. Participants stood oriented straight ahead, and one audiovisual device was placed 100 cm away from the center of his or her eyes, positioned in the lower (i.e., inferior) field, 30 cm away from the

midline in the anterior–posterior axis. This device was hence within PPS as estimates of the boundary between PPS and extrapersonal space has ranged between 30 and 80 cm (Duhamel et al., 1998; Galli et al., 2015; Graziano et al., 1999; Noel et al., 2014; Noel et al., 2015a; Schlack, Sterbing-D'Angelo, Hartung, Hoffmann, & Bremmer, 2005; Serino et al., 2015). The second audiovisual device was also placed 100 cm away from the center of the participant's eyes but was directed into upper (i.e., superior) space and was also 30 cm away from the anterior–posterior axis of the body (see Figure 1a). Thus, in the latter case, the audiovisual device was effectively 100 cm away from the closest body part. In this manner, both audiovisual devices were placed at the same distance from eyes and ear, yet one of them was situated within the participant's PPS (that is, less than or equal to 30–80 cm from the closest body part), and the other was not (that is, positioned more than 30–80 cm away from the closest body part). Subjects were instructed to tilt their head (either toward the bottom or toward the top) in order to achieve a direct focal gaze toward the audiovisual devices. In this first experiment, however, no measure was taken in order to ensure that participants were not shifting their eyes in orbit. Participants were asked to judge whether an audiovisual event happened synchronously or asynchronously and to indicate their response via button press. They completed a total of 12 experimental blocks, each consisting of 66 repetitions (6 repetitions × 11 SOAs). For half the blocks, participants gazed toward the upper device, and for the other half, they gazed toward the lower one. Block order was randomized across participants, and the order of the different trials within a block was fully randomized. In total, 36 repetitions per SOA and per location were acquired. To minimize anticipatory responses the interstimulus interval was shuffled between 2.5, 3.0, and 3.5 s. Each block lasted approximately 4 min, and participants were allowed pauses between blocks.

Analyses

Each individual subject's raw data was utilized to calculate report of simultaneity (proportion) as a function of SOA and location (inside PPS vs. outside PPS). Subsequently, two psychometric sigmoid functions (MATLAB *glmfit* as in Stevenson et al., 2014) were fitted to the individual subject's average data—one on the left side (audio-first stimuli) and one on the right side (visual-first stimuli). The SOA at which these curves intersected was taken as the subject's PSS (that is, the SOA at which the subject is most likely to judge auditory and visual stimuli as synchronous). Similarly, the addition of the absolute value of the points at which these curves intersected 0.5 (that is, the SOAs at which participants were equally likely to categorize an
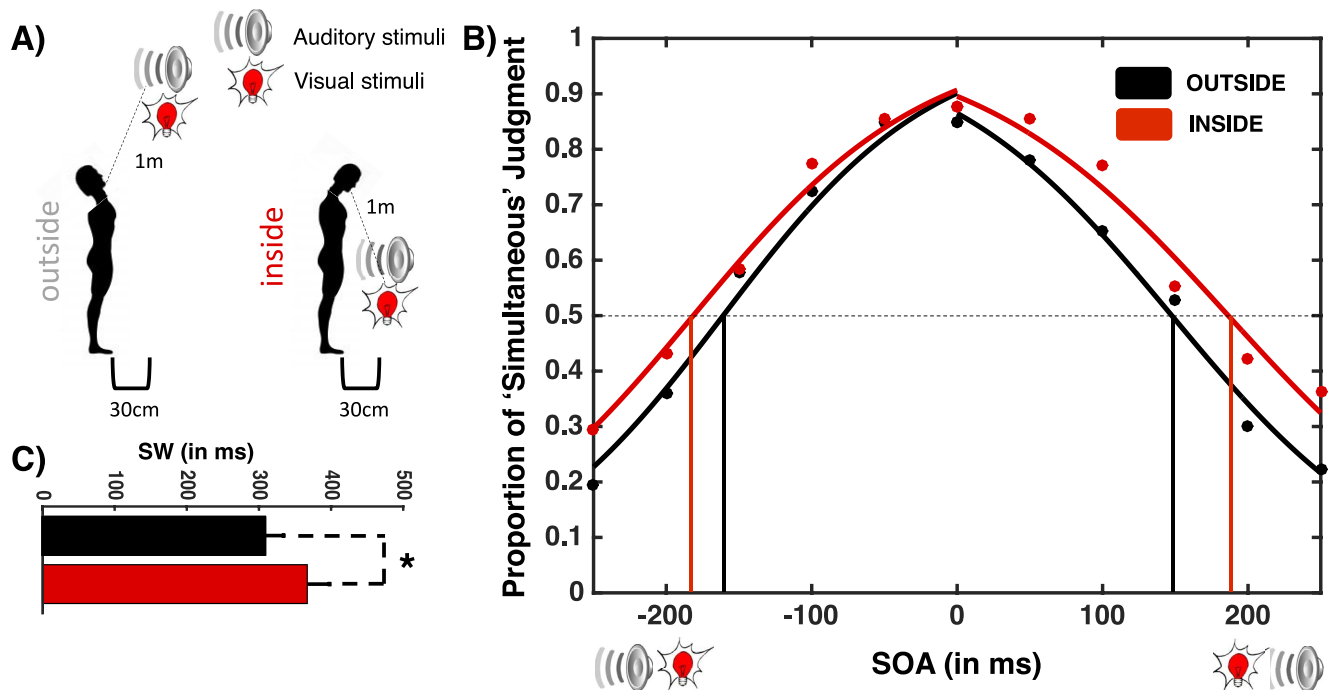
Figure 1. Audiovisual SJ inside and outside PPS. (A) Setup. In both conditions, participants tilted their head as to gaze focally at the audiovisual object. In one condition, the audiovisual event was placed outside PPS as the device was placed toward the top, and in the other, it was placed within PPS as the device was place toward the bottom, close to the participant's feet. (B) Proportion of "simultaneous" judgment between visual and auditory stimuli as a function of the different SOAs. (C) Width of SW. The SW was larger when the audiovisual events happened inside PPS than outside PPS. Error bars represent $\pm$ 1 *SEM*. *$p < 0.05$.

audiovisual presentation as synchronous and asynchronous when either audio [ASW] or visual [VSW] stimuli led) was taken as the subject's SW. In addition to this "standard" psychometric analysis, data was also fitted to an SJ-tailored independent channels model (García-Pérez & Alcalá-Quintana 2012a, 2012b; see also Alcalá-Quintana & García-Pérez, 2013, for an implementation and validation of the model) in order to estimate further parameters, putatively governing the underlying neurocognitive process involved in simultaneity judgments (see García-Pérez & Alcalá-Quintana, 2012a, 2012b, for details). None of the parameters of the independent channels model—two rate parameters indexing the neural precision of latency for each signal (auditory and visual), a processing delay parameter, and a criterion value for synchrony— significantly explained our pattern of results. Thus, although standard psychometric function fitting does not allow for ruling out the possibility that the observed effects are due to spurious variables, such as a decisional criteria, the current report focuses on the values (PSS and SW) extracted from the sigmoidal fittings as these are more commonplace (see Supplementary Material online for detailed analysis on the parameters of the independent channel model). Prior to the formal analysis described above, goodness of fit for each condition is evaluated by means of a chi-square and then compared among each other.

Results and discussion

A paired-samples *t* test (on differences of chi-square) demonstrated that the sigmoid fitting was equally well suited to both the condition in which participants gazed toward the top (outside PPS) and toward the bottom (inside PPS), $t(20) < 1$, $p = 0.89$. A paired-samples *t* test demonstrated that the width of the audiovisual SW was significantly, $t(20) = 2.38$, $p = 0.046$, larger when the stimuli were inside ($M = 368.1$ ms, $SE = 28.9$ ms) as opposed to outside ($M = 308.5$ ms, $SE = 24.2$ ms) PPS. Findings revealed no greater effects for the left or right sides of the SW as the simultaneity boundary was not selectively altered on either the auditory-first, $t(20) < 1$, $p = 0.38$, or visual-first, $t(20) = 1.24$, $p = 0.18$, side of the psychometric curve. There was no difference in the point of subjective simultaneity (PSS), $t(20) = 1.10$, $p = 0.27$.

We hypothesize that the enlargement of the SW for stimuli occurring near the body depends on a mandatory and additional recoding of external stimuli, which are initially processed exclusively in native receptor-dependent reference frames (i.e., head-cen-

tered for auditory events and eye-centered for visual events) into a more general and body-centered reference frame. External auditory and visual stimuli that are presented within PPS are processed by dedicated neural systems and interact with somatosensory representations in order to detect or anticipate contact with the body (Brozzoli, Gentile, & Ehrsson, 2012; Graziano & Cooke, 2006; Kandula, Hofman, & Dijkerman, 2015; Makin, Holmes, Brozzoli, & Farnè, 2012; Rizzolatti et al., 1997). Thus, external stimuli close to the body may be automatically integrated into a body-centered reference frame to more efficiently enable visuo-tactile and audio-tactile interactions. Indeed, interaction with and manipulation of external auditory or visual signals requires transformation from the native reference frame to the reference frame of either a particular body part or the body as a whole (Andersen & Buneo, 2002; Andersen, 1995; Andersen, Snyder, Bradley, & Xing, 1997; Colby, Duhamel, & Goldberg, 1993; Rizzolatti et al., 1997; Sereno & Huang, 2014). The computational requirement for reference frame transformations is indeed necessary to maintain receptive fields anchored on a particular body part as has been shown for the receptive fields of neurons presumed to be encoding PPS (Avillac, Denève, Olivier, Pouget, & Duhamel, 2005). The prediction here, thus, is based on the assumption that larger simultaneity temporal windows within PPS is representative of a greater window within which integration would occur and the observation that when stimuli are within PPS this integration happens (among other places) in neurons with receptive fields anchored on different body parts as well as the whole body. Within such a framework, the addition of further processing steps (or discrepancies in the form of reconciling representations in both native reference frames and body-centered reference frames) for a given stimulus modality (e.g., auditory) would result in an enlarged SW that is selectively driven by the stimulus needing additional processing. In other words, when audiovisual stimuli are presented close to the participant, they are processed not only in a native reference frame (eye-centered for visual and head-centered for auditory), but also in a reference frame taking into account body posture. Thus, when eye-centered, head-centered, and body-centered reference frames are misaligned, the reconciliation between these representations is most readily manifested in additional computation for stimuli coded in the modality whose native reference frame is misaligned from the body-centered reference frame, i.e., visual stimuli in the case of misaligned eye direction and auditory stimuli in the case of misaligned head direction. We suggest that in order to correctly bind information emanating from a common source, the neural processing scheme may allow for a greater temporal lead by the part of the sensory modality for which the native reference frame

has been misaligned with respect to the body (and thus may require additional processing time). That is, we conjecture that the computation performed in order to make SJs may adapt as to correct for physiological lags due to reference frame transformation/reconciliation (see Leone & McCourt, 2013, 2015, for a similar argument). In the second experiment, we test this prediction by dissociating either the auditory or the visual reference frame with respect to the body-centered reference frame. The hypothesis for this manipulation is that larger SWs will be seen inside PPS than outside (as in Experiment 1) but, more importantly, that these enlargements will be driven selectively by a larger ASW in the case of head/body misalignment (head tilted with respect to the whole body/trunk) and a larger VSW in the case of eyes/body misalignment (eyes in orbit tilted with respect to the whole body/trunk and head). This argument is bolstered by evidence that some visuo-tactile neurons have visual receptive fields anchored on the body, suggesting that the encoding of visual stimuli does not need to be transformed from eye-centered to head-centered to body-centered coordinates but may be directly encoded from eye-centered coordinates into body-centered coordinates (Avillac et al., 2005).

Experiment 2

In order to test the prediction that larger SWs for close stimuli are rooted in the further automatic body-centered processing and the reconciliation of this one with native reference frames, we invited two additional groups of participants to take part in another experiment. As in Experiment 1, audiovisual events were presented either inside or outside PPS, and participants were asked to make SJs concerning them. In contrast to Experiment 1, however, participants were also asked to misalign their auditory and body-centered reference frames by tilting their head with respect to their body or to misalign their visual and body-centered reference frames by gazing downward and hence changing the orientation of their eyes within their orbits.

Methods

Participants

Thirty-nine subjects took part in this experiment: 18 (four females, mean age 22.9 ± 2.8, range 19–29) in the Head/Body Misaligned group and another 21 (six females, mean age 23.1 ± 2.0, range 19–26) in the Eyes/Body Misaligned group. This variable (i.e., Misalignment Group), thus, was a between-subjects one. All participants, students from the Ecole Polytechnique Federale de Lausanne, reported normal auditory acuity

and had normal or corrected-to-normal visual acuity. All participants gave their informed consent to take part in this study, which was approved by the local ethics committee and the Brain Mind Institute Ethics Committee for Human Behavioral Research at EPFL, and were reimbursed for their time.

Stimulus and apparatus, procedure, and analyses

The stimuli, apparatus, and analyses were very similar to those used in Experiment 1. The procedure was similar with the exception that one of the audiovisual devices was placed 100 cm away, toward the bottom, from the center of the participant's eyes and equally 30 cm away from his or her midline in the anterior–posterior axis (inside PPS), and the other one was placed 100 cm away from the center of the participant's eyes; however, this time at a vertical height of each participant's eye level (outside PPS, see Figure 2a). That is, the bottom device was placed at the same location as in Experiment 1, but the second device was placed at the height of the participant's eyes instead of toward the top (in order to avoid head tilting). Within this stimulus configuration, two different reference frame misalignments were conducted. For one group, head-centered and body-centered reference frames were misaligned in the inside PPS condition as participants were asked to tilt their head downward (without moving their eyes) in order to gaze foveally at the audiovisual device. For the other group, eye-centered and body-centered reference frames were misaligned in the inside PPS condition as participants were asked to keep their head directed straightforward while moving their eyes to foveate the audiovisual device. An experimenter in the testing room continuously monitored head and eye orientation and reminded participants to comply with indications when head or eye orientation was not kept as required.

Results and discussion

Goodness-of-fit values from a chi-square test were submitted to a mixed-model ANOVA with location (Inside PPS vs. Outside PPS) as the within-subject variable, and Misalignment Group (Head vs. Eyes) as the between-subjects variable. This analysis revealed that the sigmoid fitting was equally well suited to both locations, $F(1, 37) < 1$, $p = 0.85$, and both Misalignment Groups, $F(1, 37) = 1.77$, $p = 0.34$, and that there was no interaction between these variables, $F(1, 37) = 2.01$, $p = 0.23$.

Mixed-model ANOVA with Misalignment Group (Head/Body vs. Eyes/Body) and Stimuli Location (Inside vs. Outside PPS) were run on the function's parameters. As depicted in Figure 2, results replicated those of Experiment 1 by showing that the SW inside PPS ($M = 362.4$, $SE = 34.1$) was significantly larger than outside PPS ($M = 307.2$ ms, $SE = 28.4$ ms), $F(1, 37) = 7.74$, $p = 0.039$. This effect was also independently true for each of the Misalignment Groups (Head/Body and Eyes/Body misalignment, both $ps < 0.05$). Neither the effect of Misalignment Group, $F(1, 37) < 1$, $p = 0.63$, nor an interaction between Location $\times$ Misalignment Group, $F(1, 37) < 1$, $p = 0.79$, were observed. Analyses on PSS showed no significant effect of Stimuli Location, $F(1, 37) = 1.19$, $p = 0.29$; Misalignment Group, $F(1, 37) = 2.27$, $p = 0.17$; or Stimuli Location $\times$ Misalignment Group, $F(1, 37) = 2.44$, $p = 0.20$.

Based on the experimental hypothesis that reference frame misalignment will add additional processing demands to the SJs, the most interesting results concern ASW and VSW scores. Therefore, an additional mixed-model ANOVA, incorporating the Leading Stimuli variable (visual or auditory) was performed. The three-way interaction, Leading Stimuli $\times$ Stimuli Location $\times$ Misalignment Group, was significant, $F(1, 37) = 7.35$, $p = 0.019$. Consequently, two separate within-subject ANOVAs were performed for the two Misalignment Groups in order to shed light on the nature of this interaction. As shown in Figure 3, the ANOVA carried out solely on the Head/Body Misalignment Group revealed a significant Leading Stimuli $\times$ Stimuli Location interaction, $F(1, 16) = 4.64$, $p = 0.038$, which was driven by the fact that the boundary of subjective simultaneity for the auditory-leading stimuli was larger within PPS than outside it, $t(17) = 3.47$, $p = 0.008$. In contrast, this was not the case on the visual-leading side of the psychometric curve, $t(17) = 0.20$, $p = 0.88$. In a complementary fashion, analysis for the Eye/Body Misalignment Group revealed a significant Leading Stimuli $\times$ Stimuli Location interaction, $F(1, 19) = 3.91$, $p = 0.041$, and this was driven by the fact that the boundary of subjective simultaneity for the visual-leading stimuli was larger within PPS than outside it, $t(20) = 2.27$, $p = 0.033$. This was not the case on the auditory-leading side of the psychometric curve, $t(20) = 1.03$, $p = 0.071$.

The results from Experiment 2 reinforce the results from Experiment 1 by confirming that larger SWs are exhibited inside PPS when compared with outside. Furthermore, the results of Experiment 2 extend these results by showing that when a transition between frames of reference is needed, larger SWs are most prominent for the leading stimulus in the modality whose coordinate frame was misaligned relative to the body representation (i.e., auditory leading in the head/body misalignment and visual leading in the eyes/body misalignment). That is, if the head is misaligned with respect to the trunk, auditory stimuli will have to occur earlier than visual stimuli in order to be perceived as synchronous (relative to the nonmisaligned condition).
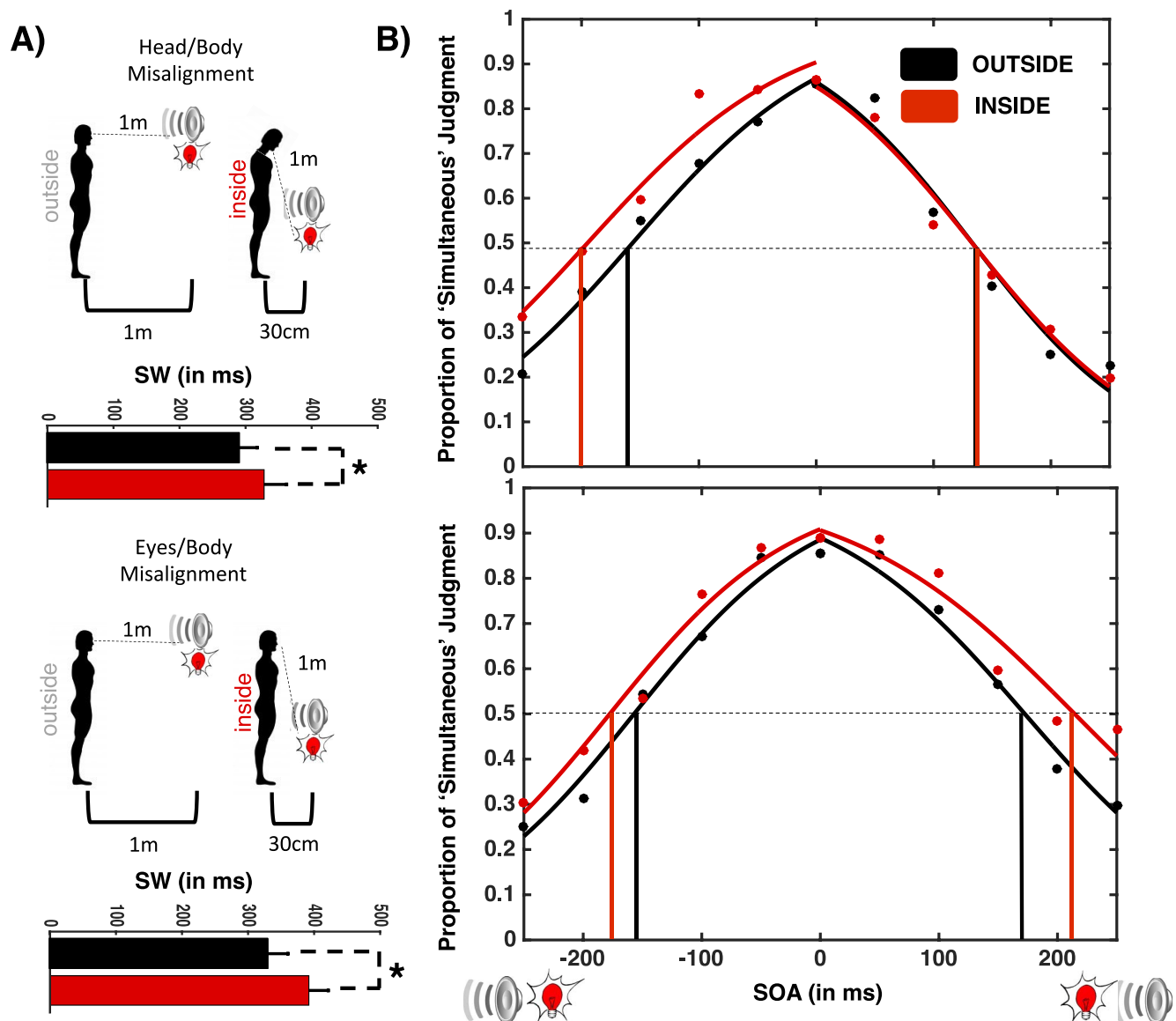
Figure 2. Audiovisual SJ inside and outside PPS with misaligned audio/body-centered reference frames (top) and visual/body-centered reference frames (bottom). (A) Setup. Participants either tilted their head in order to gaze directly at an audiovisual event inside PPS while misaligning their auditory and body-centered reference frames (top) or tilted their gaze (eyes within their orbits) in order to view the audiovisual object inside PPS while misaligning their visual and body-centered reference frames. (B) Width of SW. The SW was larger when the audiovisual events happened inside PPS than outside PPS. Error bars represent ± 1 *SEM*. *$p < 0.05$.

Conversely, if it is the eyes that are misaligned with the trunk, visual stimuli have to occur earlier in time in order to be perceived synchronously with the auditory stimuli.

In order to confirm that the effect described here is due to an automatic processing of stimuli in the whole-body reference frame and are not due to a generic cost of additional computation due to head-centered and eye-centered reference frame misalignment, we ran the same experiment keeping the audiovisual event always outside PPS while manipulating eyes and head direction. Given that the stimuli are now outside of PPS, we expected to see no difference in SW size, regardless of reference frame misalignment.

Experiment 3

In the final experiment, we had another set of participants judge the simultaneity of audiovisual events while manipulating the misalignment of their head-centered or eye-centered reference frames with their body-centered reference frame. In contrast to
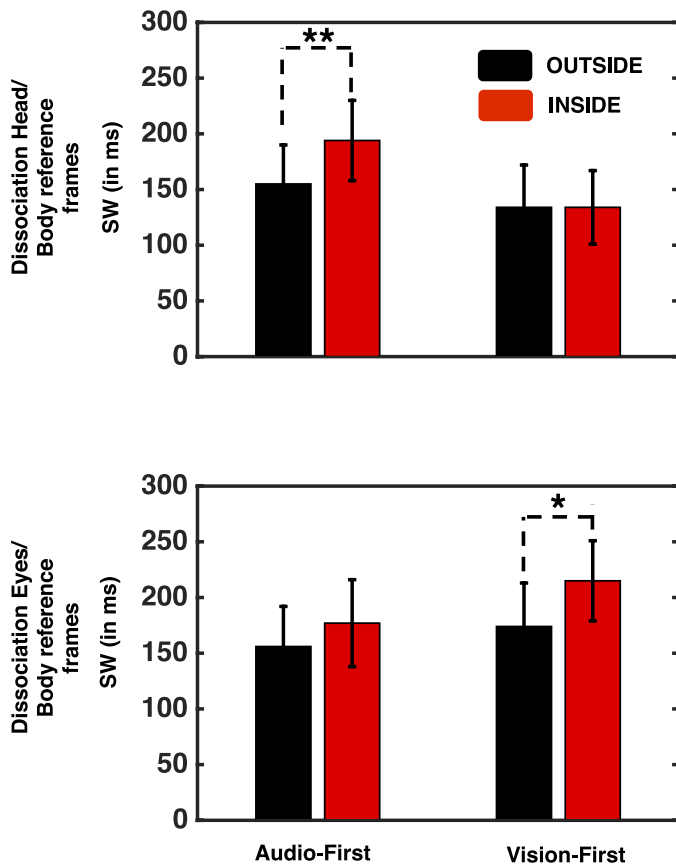
Figure 3. Simultaneity boundary for audio- and visual-leading events as a function of the location of audiovisual events and reference frame misalignment. Top: auditory and body-centered misalignment. Bottom: visual and body-centered misalignment. Results show a selective shift toward longer SOAs for the audiovisual event happening inside PPS and for which additional processing steps have been added by reference frame misalignment. Error bars indicate ± 1 *SEM*. *$p < 0.05$. **$p < 0.01$.

Experiment 2, the audiovisual events were always presented outside of PPS.

Methods

Participants

Thirty-five subjects took part in this experiment: 18 (three females, mean age 22.9 ± 3.1, range 18–28) in the Head/Body Misaligned Group and another 17 (three females, mean age 23.2 ± 2.7, range 19–27) in the Eyes/Body Misaligned Group. All participants, students at the Ecole Polytechnique Federale de Lausanne, reported normal auditory acuity and had normal or corrected-to-normal visual acuity. All subjects gave their informed consent to partake in this study, which was approved by the local ethics committee and the Brain Mind Institute Ethics Committee for Human Behavioral Research at EPFL, and were reimbursed for their participation.

Stimulus and apparatus, procedure, and analyses

All stimuli, procedures, and analyses were as in Experiment 2 except that the audiovisual device was now placed outside PPS by positioning the device 100 cm away from eyes and ears but in the upward direction (see Figure 4a.)

Results and discussion

Goodness-of-fit values from a chi-square test were submitted to a mixed-model ANOVA, with Stimuli Location (Outside PPS vs. Outside PPS-TOP) as the within-subject variable and Misalignment Group (Head/Body vs. Eyes/Body) as the between-subjects factor. This analysis revealed that the sigmoidal fitting equally suited both Stimuli Location, $F(1, 33) = 1.15$, $p = 0.29$, and both Dissociation Groups, $F(1, 33) < 1$, $p = 0.45$, and there was no interaction between these variables, $F(1, 33) < 1$, $p = 0.53$.

The function parameters were then analyzed by means of the same mixed-model ANOVA. As shown in Figure 4, there was no difference in the width of the SW for any of the conditions, including Stimuli Location, $F(1, 33) = 2.18$, $p = 0.37$, and Misalignment Group, $F(1, 33) < 1$, $p = 0.66$. Furthermore, there was no interaction between these, $F(1, 33) = 1.37$, $p = 0.31$. No effect was observed for the ASW dependent variable, Stimuli Location, $F(1, 33) < 1$, $p = 0.58$; Misalignment Group, $F(1, 33) < 1$, $p = 0.31$; or Interaction, $F(1, 33) < 1$, $p = 0.26$, nor the VSW dependent variable, Stimuli Location, $F(1, 33) = 1.87$, $p = 0.18$; Misalignment Group, $F(1, 33) < 1$, $p = 0.22$; or Interaction, $F(1, 33) = 2.12$, $p = 0.12$.

In regards to the PSS, a 2 (Stimuli Location) × 2 (Misalignment Group) mixed ANOVA performed on the PSS showed no effect, Stimuli Location, $F(1, 33) = 1.79$, $p = 0.19$; Misalignment Group, $F(1, 33) < 1$, $p = 0.63$; Interaction, $F(1, 33) < 1$, $p = 0.51$.

The results of Experiment 3 indicate that simply altering either the head-centered or eye-centered reference frame with respect the body-centered reference frame does not necessarily result in changes in the width of the multisensory SW. Rather, these changes are dependent upon the stimuli being presented within PPS. The results from Experiment 2 can thus be interpreted as a direct functional coupling between the presence of multisensory events within PPS and their temporal processing in a body reference frame.
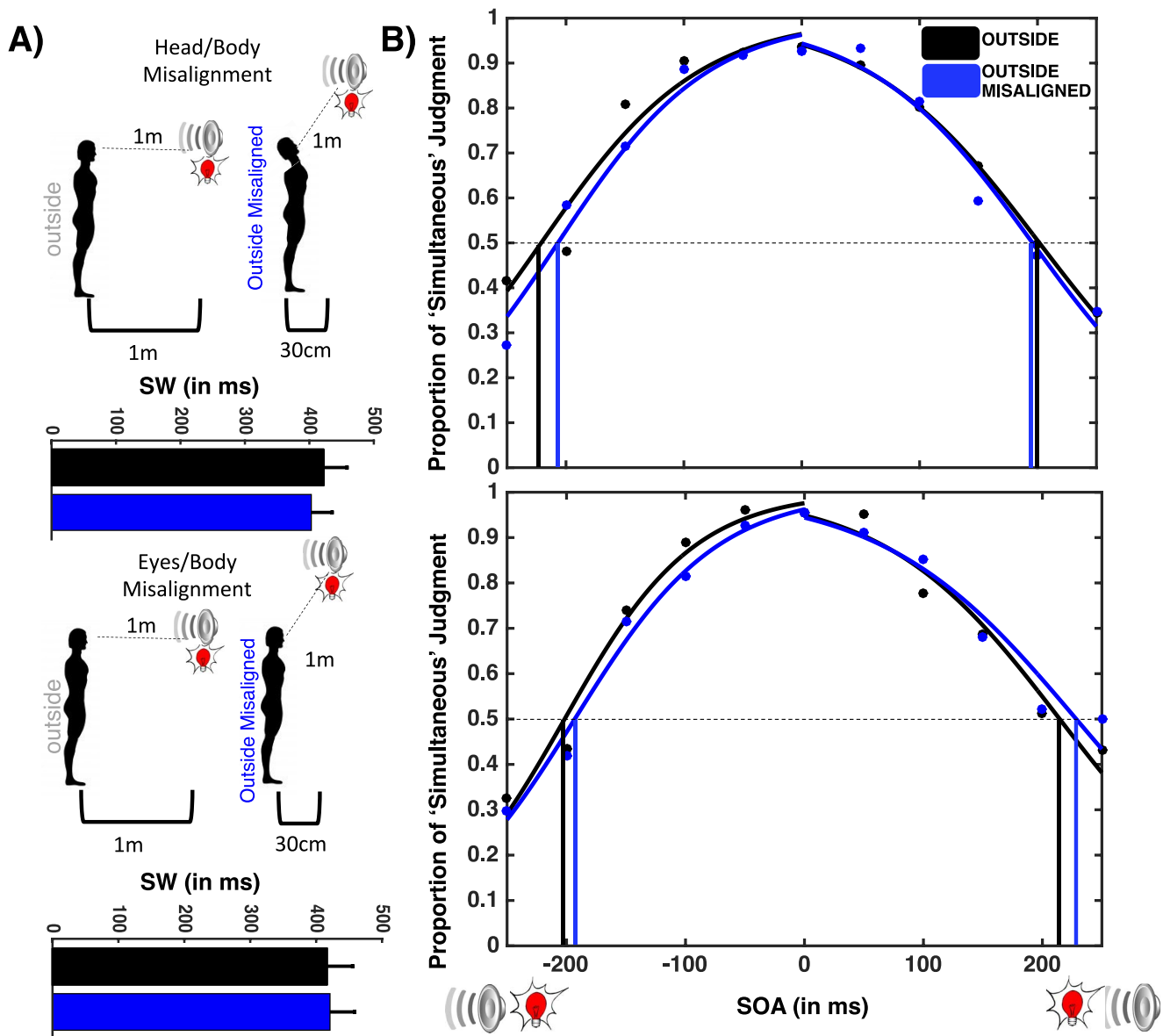
Figure 4. Audiovisual SJ outside PPS with misaligned audio/body-centered reference frames (top) and visual/body-centered reference frames (bottom). (A) Setup. Participants either tilted their head in order to gaze directly at an audiovisual event outside PPS while misaligning their auditory and body-centered reference frames (top) or tilted their gaze (eyes within their orbits) in order to view the audiovisual object outside PPS while misaligning their visual and body-centered reference frames. (B) Width of SW. The SW was unchanged by simply altering alignment between reference frames when audiovisual stimuli were presented outside PPS.

General discussion

Our findings show that the manner in which participants categorize temporal relationships between spatially congruent audiovisual events (i.e., the audiovisual SW) is not fixed throughout space. Rather, the processing architecture supporting multisensory temporal processing appears to allow for greater temporal disparity between constituent unisensory events inside as opposed to outside PPS (Experiment 1). We propose

that this effect might be due to a direct and automatic remapping (or reconciliation) of multisensory events/objects from their native reference frame (i.e., head-centered for auditory stimuli or eye-centered for visual stimuli) to (or with) a body-centered reference frame when audiovisual stimuli are within (Experiment 2), but not outside (Experiment 3) of PPS.

To the best of our knowledge, the present examination represents the most direct evidence for malleability of multisensory temporal processing across distance in the depth dimension (see Lewald & Guski,

2003, as well as Stevenson et al., 2012a, for other planes). Anecdotally, that is, not central to their question of interest, De Paepe, Crombez, Spence, and Legrain, 2014, as well as Parsons, Novich, and Eagleman, 2013, appear to have encountered the same effect presented here. De Paepe et al. acknowledge it without discussion, and Parsons et al. report a nonsignificant higher number of "simultaneity" judgments when participants were near as opposed to far from audiovisual stimulation. In the latter case, as SW were not extracted, it is hard to judge whether the effect is there or not.

Prior work has examined related questions with regards to spatially dependent changes in multisensory temporal function. For example, Sugita and Suzuki (2003) examined whether the PSS varied as a function of distance between the observer and the perceived source of the stimuli and found changes in the PSS dependent upon stimulus distance. They interpreted this effect as a consequence of a mechanism to compensate for the differences in the time of propagation of light and sound energy. Such an account cannot explain the present results as the distance between the participant's head and the source of audiovisual stimulation was kept constant at 1 m in all experiments and conditions. Thus we suggest that the present results rely on a different mechanism. We propose that SWs close to the body are larger as a consequence of an automatic remapping of stimuli into a body-centered coordinate system, which is necessary to potentially integrate visual and auditory cues related to external stimuli with tactile bodily stimulation. Note that in the current study we did not systematically manipulate hand or arm posture relative to the trunk, and thus the present results could also be a result of a remapping from native receptor-specific reference frame to the reference frame of the arm or the hand (which were aligned with that of the trunk). Regardless, the argument remains that the enlargement of multisensory SWs within PPS may derive from the process of remapping objects/events into a body or body part frame of reference.

Such a result makes sense in the context that the main function of PPS representation is to quickly and effectively process information related to external objects potentially interacting with the body (Brozzoli, Gentile, & Ehrsson, 2012; Graziano & Cooke, 2006; Kandula et al., 2015; Makin et al., 2012; Rizzolatti et al., 1997). Neurophysiological and neuroimaging evidence suggest that these functions are implemented by a frontoparietal network of multisensory neurons integrating visual and auditory cues from the space adjacent to specific body parts with tactile stimulation on the same body parts (Avillac, Ben Hamed, & Duhamel, 2007; Gentile, Petkova, & Ehrsson, 2011; Serino et al., 2011) and that project to the motor

system in order to support rapid responses (Avenanti, Annella, & Serino, 2012; Cooke & Graziano, 2004; Cooke, Taylor, Moore, & Graziano, 2003; Serino, Annella, & Avenanti, 2009). We suggest that audiovisual events that occur within the PPS are automatically recoded into a reference frame allowing for immediate processing and integration with tactile information. It is well known that the transformation needed to code stimuli from different modalities into a common reference frame takes additional processing time (Azañón, Camacho, & Soto-Faraco, 2010; Azañón, Longo, Soto-Faraco, & Haggard, 2010; Azañón & Soto-Faraco, 2008; Heed & Röder, 2010). For example, prior work shows that temporal order judgments are altered when limbs are crossed with respect to the midline. Shore, Spry, and Spence (2002) as well as Cadieux, Barnett-Cowan, and Shore (2010) suggest that tactile stimulation is initially coded according to a body-centered reference frame, and then it is rapidly and automatically remapped into an external spatial reference frame. When the hands are crossed, it is the conflict between both these representations (somatotopic/body-centered and spatiotopic/world-centered) that leads to an increased error rate in determining temporal order. Our work is in agreement with this account and extends these observations to show that it is not only touch that is automatically remapped between external and body-centered spatial coordinates, but also visual and auditory stimuli occurring within PPS (most likely through the intermediary of audio-tactile and visuo-tactile PPS neurons as, to the best of our knowledge, multisensory audiovisual neurons with receptive fields adjacent to and grounded on the body are not known; Schlack et al., 2005). This account would also explain why limb-posture effects on temporal order judgments are not only observed when limbs are crossed with respect to the midline, but also when they are merely close together (Shore, Gray, Spry, & Spence, 2005) or even just perceived to be close together (Gallese & Spence, 2005).

A recent finding has indicated that audiovisual integration is enhanced for stimuli that are presented further from as opposed to closer to the body (Van der Stoep et al., 2015c). This observation could be interpreted as being in direct opposition to the pattern of results revealed here. However, it is important to highlight that Van der Stoep et al. (2015c) utilized a redundant target task (indexing magnitude of multisensory integration) and did not vary SOAs. Here, on the other hand, via a proxy measure, such as a SJ, we attempted to index the temporal extent over which multisensory interactions are likely and not the magnitude of these interactions. Thus, it is conceivable that multisensory facilitation is greater farther away than closer but that multisensory effects are most

resistant to temporal disparities closer versus farther. On the contrary, we judge Van der Stoep et al.'s (2015c) findings to make a good deal of ecological sense as audiovisual stimuli typically decrease in visual angle and intensity as they occur farther from the observer. Thus, the findings reported by Van der Stoep et al. (2015c) highlight the evolutionary advantage that can be conferred by multisensory integration and illustrate the importance of the depth dimension in these interactions. The current study further details the importance of stimulus depth in gating multisensory function and specifically shows a previously unknown relationship between depth (specifically within vs. outside PPS) and multisensory temporal function. Indeed, the pattern of findings reported here, in conjunction with those by Van der Stoep et al. (2015c), would suggest that as multisensory SWs enlarge, multisensory benefits may be observed at wider temporal discrepancies between the paired unisensory stimuli; however, those benefits may be smaller in amplitude. This speculation however remains to be tested.

Evolutionarily, smaller SWs (as opposed to larger as reported here), in particular within PPS, would seem advantageous. However, it may well be that having a broader temporal filter (and thus a more liberal binding criterion) would allow for a greater weighting of stimuli within PPS. Such within-PPS binding could be yoked to processes, such as speeded responses to bound audio-visual stimuli, which would thus confer a strong adaptive advantage. Second, and in a related manner, the nature of the physical world may dictate the construction of these representations For example, stimuli that are approaching from a distance are also growing in size, and this increase in size may necessitate a rescaling of temporal function given the spatiotemporal contingencies of real moving stimuli. Finally, it may well be that the enlarged SW within PPS is derivative of the greater/different computations needed in order to do the transformations described in the current study and thus has little true adaptive value unto itself apart from compensating for neural/physiological lag between different neural processing schemes.

The notion that the SW for multisensory stimuli is larger within PPS than outside it also has important implications for the emerging body of evidence suggesting atypical multisensory temporal function in clinical conditions. For example, it has been recently shown that individuals with autism spectrum disorder (ASD) have abnormally large audiovisual SWs (most notably for complex audiovisual stimuli, such as speech; Stevenson et al., 2014). These researchers further posit that temporal processing alterations may play an important role in the higher-order deficits that are well established in domains such as social commu-

nication. In the context of the current results, one possibility explaining the expanded SW observed in ASD may be a difference in PPS size. Indeed, recent studies have indicated that ASD individuals possess an altered representation of personal space (Ferri, Ardizzi, Ambrosecchia, & Gallese, 2013; Gessaroli, Santelli, di Pellegrino, & Frassinetti, 2013; Kennedy & Adolphs, 2014). Further work is needed in order to test these links between spatial and temporal function in clinical populations.

Last, we must point out that throughout the report we have assumed that 30 cm is "within PPS" and 100 cm is "outside PPS" based on prior audio-tactile and visuo-tactile work (Noel et al., 2014; Noel et al., 2015a; Serino et al., 2015). However, in the current study, we neither measured nor attempted to manipulate the size of PPS. Clearly, in further work, it would be interesting to manipulate PPS size. Nonetheless, the results reported here undoubtedly demonstrate that when audiovisual stimuli are presented closer to the body, participants are more likely to categorize these presentations as co-occurring in time.

To conclude, previous evidence shows interdependencies between temporal and spatial perception. Our results add that spatiotemporal effects in multisensory processing are not only influenced by the spatial relationships between stimuli that are integrated, but also depend on the spatial relationship between those external stimuli and the body of the perceiver. This is an important factor to consider when studying spatiotemporal interactions between disparate stimuli. Although such linkages are intuitive given the seamless integration between time and space in the world in which we live, experimental approaches have largely focused on one particular dimension or domain at a time (with good reason in many respects given the high dimensionality of the stimulus space once time and space are taken into account). However, we argue here that work moving forward must increasingly explore the interactions and interdependencies between these dimensions as it is only under such circumstances that we will begin to provide a unified view into sensory, perceptual, and cognitive processes.

Acknowledgments

Corresponding authors: Jean-Paul Noel; Andrea Serino.
Email: jean-paul.noel@vanderbilt.edu; Andrea.Serino@epfl.ch.
Address: Vanderbilt University Medical Center, Vanderbilt University, Nashville, TN, USA; School of Life Sciences, Ecole Polytéchnique Fédérale de Lausanne, Lausanne, Switzerland.

References

Alcalá-Quintana, R., & García-Pérez, M. A. (2013). Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. *Behavior Research Methods, 45*, 972–998, doi:10.3758/s13428-013-0325-2.

Andersen, R. A. (1995). Encoding of intention and spatial location in the posterior parietal cortex. *Cerebral Cortex, 5*(5), 457–469.

Andersen, R. A., & Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. *Annual Review of Neuroscience, 25*, 189–220.

Andersen, R. A., Snyder, L. H., Bradley, D. C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience, 20*, 303–330.

Avenanti, A., Annella, L., & Serino, A. (2012). Suppression of premotor cortex disrupts motor coding of peripersonal space. *Neuroimage, 63*, 281–288.

Avillac, M., Ben Hamed, S., & Duhamel, J.-R. (2007). Multisensory integration in the ventral intrapariel area of the macaque monkey. *Journal of Neuroscience, 27*(8), 1922–1932.

Avillac, M., Denève, S., Olivier, E., Pouget, A., & Duhamel, J. R. (2005). Reference frames for representing visual and tactile locations in parietal cortex. *Nature Neuroscience, 8*, 941–949.

Azañón, E., Camacho, K., & Soto-Faraco, S. (2010). Tactile remapping beyond space. *European Journal of Neuroscience, 31*, 1858–1867.

Azañón, E., Longo, M. R., Soto-Faraco, S., & Haggard, P. (2010). The posterior parietal cortex remaps touch into external space. *Current Biology, 20*, 1304–1309.

Azañón, E., & Soto-Faraco, S. (2008). Changing reference frames during the encoding of tactile events. *Current Biology, 18*, 1044–1049.

Benevento, L. A., Fallon, J. H., Davis, B., & Rezak, M. (1977). Auditory-visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. *Experimental Neurology, 57*, 849–872.

Brozzoli, C., Gentile, G., & Ehrsson, H. H. (2012). That's near my hand! Parietal and premotor coding of hand-centered space contributes to localization and self-attribution of the hand. *Journal of Neuroscience, 32*(42), 14573–14582.

Brozzoli, C., Gentile, G., Petkova, V. I., & Ehrsson, H. H. (2011). fMRI-adaptation reveals a cortical mechanism for the coding of space near the hand. *Journal of Neuroscience, 31*(24), 9023–9031.

Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology, 46*, 369–384. [PubMed: 6267219]

Cadieux, M. L., Barnett-Cowan, M., & Shore, D. I. (2010). Crossing the hands is more confusing for females than males. *Experimental Brain Research, 204*, 431–446, doi:10.1007/s00221-010-2268-5.

Calvert, G. A., Spence, C., & Stein, B. E. (2004). *The handbook of multisensory processes*. Cambridge, MA: The MIT Press.

Canzoneri, E., Magosso, E., & Serino, A. (2012). Dynamic sounds capture the boundaries of peripersonal space representation in humans. *PLoS ONE, 7*(9), e44306, doi:10.1371/journal.pone.0044306.

Canzoneri, E., Marzolla, M., Amoresano, A., Verni, G., & Serino, A. (2013a). Amputation and prosthesis implantation shape body and peripersonal space representations. *Scientific Reports, 3*, 2844, doi:10.1038/srep02844.

Canzoneri, E., Ubaldi, S., Rastelli, V., Finisguerra, A., Bassolino, M., & Serino, A. (2013b). Tool-use reshapes the boundaries of body and peripersonal space representations. *Experimental Brain Research, 228*(1), 25–42, doi:10.1007/s00221-013-3532-2.

Cappe, C., Thelen, A., Romei, V., Thut, G., & Murray, M. M. (2012). Looming signals reveal synergistic principles of multisensory interactions. *Journal of Neuroscience, 32*, 1171–1182.

Chen, L., & Vroomen, J. (2013). Intersensory binding

across space and time: A tutorial review. *Attention, Perception & Psychophysics, 75*, 790–811.

Colby, C. L., Duhamel, J. R., & Goldberg, M. E. (1993). Ventral intraparietal area of the macaque: Anatomic location and visual response properties. *Journal of Neurophysiology, 69*, 902–914, http://jn.physiology.org/content/69/3/902.full-text.pdf + html.

Colby, C. L., & Goldberg, M. E. (1998). Space and attention in parietal cortex. *Annual Review of Neuroscience, 22*, 319–349.

Cooke, D. F., & Graziano, M. A. S. (2004). Sensori-motor integration in the precentral gyrus: Poly-sensory neurons and defensive movements. *Journal of Neurophysiology, 91*, 1648–1660.

Cooke, D. F., Taylor, C. S. R., Moore, T., & Graziano, M. S. A. (2003). Complex movements evoked by microstimulation of Area VIP. *Proceedings of the National Academy of Sciences, USA, 100*, 6163–6168.

De Paepe, A. L., Crombez, G., Spence, C., & Legrain, V. (2014). Mapping nocioceptive stimuli in a peripersonal frame of reference: Evidence from a temporal order judgment task. *Neuropsychologia, 56*, 219–228.

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perceptions & Psychophysics, 66*, 1388–1404.

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1998). Ventral intraparietal area of the macaque: Congruent visual and somatic response properties. *Journal of Neurophysiology, 79*, 126–136, http://jn.physiology.org/content/79/1/126.full-text.pdf + html.

Farne, A., & Ladavas, E. (2002). Auditory peripersonal space in humans. *Journal of Cognitive Neuroscience, 14*, 1030–1043.

Ferri, F., Ardizzi, M., Ambrosecchia, M., & Gallese, V. (2013). Closing the gap between *the inside* and *the outside*: Interoceptive sensitivity and social distances. *PLoS ONE, 8*(10), e75758, doi:10.1371/journal.pone.0075758.

Fogassi, L., Gallese, V., Fadiga, L., Luppino, G., Matelli, M., & Rizzolatti, G. (1996). Coding of peripersonal space in inferior premotor cortex (area F4). *Journal of Neurophysiology, 76*, 141–157, http://www.ncbi.nlm.nih.gov/pubmed/8836215.

Frassinetti, F., Bolognini, N., & Ladavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research, 147*, 332–343.

Frens, M. A., Van Opstal, A. J., & Van der Willigen, R. F. (1995). Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Perception & Psychophysics, 57*, 802–816.

Gallese, A., & Spence, C. (2005). Visual capture of apparent limb position influences tactile temporal order judgments. *Neuroscience Letters, 379*, 63–68, doi:10.1016/j.neulet.2004.12.052.

Galli, G., Noel, J. P., Canzoneri, E., Blanke, O., & Serino, A. (2015). The wheelchair as a full-body tool extending the peripersonal space. *Frontiers in Psychology, 6*, 639, doi:10.3389/fpsyg.2015.00639.

García-Pérez, M. A., & Alcalá-Quintana, R. (2012a). On the discrepant results in synchrony judgment and temporal-order judgment tasks: A quantitative model. *Psychonomic Bulletin & Review, 19*, 820–846, doi:10.3758/s13423-012-0278-y.

García-Pérez, M. A., & Alcalá-Quintana, R. (2012b). Response errors explain the failure of independent-channels models of perception of temporal order. *Frontiers in Psychology–Perception Science, 3*, 94, doi:10.3389/fpsyg.2012.00094.

Gentile, G., Guterstam, A., Brozzoli, C., & Ehrsson, H. H. (2013). Disintegration of multisensory signals from the real hand reduces default limb self-attribution: An fMRI study. *The Journal of Neuroscience, 33*(33), 13350–13366.

Gentile, G., Petkova, V. I., & Ehrsson, H. H. (2011). Integration of visual and tactile signals from the hand in the human brain: An FMRI study. *Journal of Neurophysiology, 105*, 910–922.

Gessaroli, E., Santelli, E., di Pellegrino, G., & Frassinetti, F. (2013). Personal space regulation in childhood autism spectrum disorders. *PLoS ONE, 8*, e74959.

Graziano, M. S., Reiss, L. A. J., & Gross, C. G. (1999, Feb 4). A neuronal representation of the location of nearby sounds. *Nature, 397*, 428–430.

Graziano, M. S., & Cooke, D. F. (2006). Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia, 44*(13), 2621–2635.

Heed, T., & Azañón, E. (2014). Using time to investigate space: A review of tactile temporal order judgments as a window onto spatial processing in touch. *Frontiers in Psychology, 5*(76), 1–16.

Heed, T., & Röder, B. (2010). Common anatomical and external coding for hands and feet in tactile attention: Evidence from event-related potentials. *Journal of Cognitive Neuroscience, 22*(1), 184–202.

Hikosaka, K., Iwai, E., Saito, H., & Tanaka, K. (1988). Polysensory properties of neurons in the anterior

bank of the caudal superior temporal sulcus of the macaque monkey. *Journal of Neurophysiology, 60*, 1615–1637. [PubMed: 2462027].

Kandula, M., Hofman, D., & Dijkerman, H. C. (2015). Visuo-tactile interactions are dependent on the predictive value of the visual stimulus. *Neuropsychologia*, *70*, 358–366, http://dx.doi.org/10.1016/j.neuropsychologia.2014.12.008.

Kennedy, D. P., & Adolphs, R. (2014). Violations of personal space by individuals with autism spectrum disorder. *PLoS ONE, 9*(8), e103369, doi:10.1371/journal.pone.0103369.

Leone, L., & McCourt, M. E. (2013). The roles of physical and physiological simultaneity in audiovisual multisensory facilitation. *I-Perception, 4*, 213–228.

Leone, L., & McCourt, M. E. (2015). Dissociation of perception and action in audiovisual multisensory integration. *European Journal of Neuroscience*, *42*, 2915–2922.

Lewald, J., & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cognitive Brain Research, 16*(3), 468–478. [S0926641003000740]

Lovelace, C. T., Stein, B. E., & Wallace, M. T. (2003). An irrelevant light enhances auditory detection in humans: A psychophysical analysis of multisensory integration in stimulus detection. *Cognitive Brain Research, 17*, 447–453.

Macaluso, E., George, N., Dolan, R., Spence, C., & Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: A PET study. *Neuroimage, 21*, 725–732.

Makin, T. R., Holmes, N. P., Brozzoli, C., & Farnè, A. (2012). Keeping the world at hand: Rapid visuomotor processing for hand-object interactions. *Experimental Brain Research, 219*, 421–428.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience, 7*, 3215–3229.

Meredith, M. A., & Stein, B. E. (1986a). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research, 365*, 350–354.

Meredith, M. A., & Stein, B. E. (1986b). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology, 56*, 640–662.

Miller, L. M., & D'Esposito, M. (2005). Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. *Journal of Neuroscience, 25*(25), 5884–5893.

Murray, M. M., & Wallace, M. T. (2012). *The neural bases of multisensory processes*. Boca Raton, FL: CRC Press.

Nelson, W. T., Hettinger, L. J., Cunningham, J. A., Brickman, B. J., Haas, M. W., & McKinley, R. L. (1998). Effects of localized auditory information on visual target detection performance using a helmet-mounted display. *Human Factors, 40*(3), 452–460.

Ngo, M. K., & Spence, C. (2010). Auditory, tactile, and multisensory cues can facilitate search for dynamic visual stimuli. *Attention, Perception, & Psychophysics, 72*, 1654–1665.

Noel, J. P., Grivaz, P., Marmaroli, P., Lissek, H., Blanke, O., & Serino, A. (2014). Full body action remapping of peripersonal space: The case of walking. *Neuropsychologia, 70*, 375–384, doi:10.1016/j.neuropsychologia.2014.08.030.

Noel, J. P., Pfeiffer, C., Blanke, O., & Serino, A. (2015a). Peripersonal space as the space of the bodily self. *Cognition, 144*, 49–57, doi:10.1016/j.cognition.2015.07.012.

Noel, J. P., Wallace, M. T., Orchard-Mills, E., Alais, D., & Van der Burg, E. (2015b). True and perceived synchrony are preferentially associated with particular sensory pairings. *Scientific Reports, 5*, 17467, doi:10.1038/srep17467.

Nozawa, G., Reuter-Lorenz, P. A., & Hughes, H. C. (1994). Parallel and serial processes in the human oculomotor system: Bimodal integration and express saccades. *Biological Cybernetics, 72*, 19–34.

Parsons, B., Novich, S. D., & Eagleman, D. M. (2013). Motor-sensory recalibration modulates perceived simultaneity of cross-modal events. *Frontiers in Psychology, 4*, 46.

Quinn, B. T., Carlson, C., Doyle, W., Cash, S. S., Devinsky, O., Spence, C., . . . Thesen, T. (2014). Intracranial cortical responses during visual-tactile integration in humans. *Journal of Neuroscience, 34*(1), 171–181.

Rauschecker, J. P., & Korte, M. (1993). Auditory compensation for early blindness in cat cerebral cortex. *Journal of Neuroscience, 13*, 4538–4548. [PubMed: 8410202]

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1997). The space around us. *Science, 277*, 190–191.

Royal, D. W., Carriere, B. N., & Wallace, M. T. (2009). Spatiotemporal architecture of cortical receptive fields and its impact on multisensory interactions. *Experimental Brain Research, 198*(2–3), 127–136, doi:10.1007/s00221-009-1772-y.

Royal, D. W., Krueger, J., Fister, M. C., & Wallace, M. T. (2010). Adult plasticity of spatiotemporal receptive fields of multisensory superior colliculus neurons following early visual deprivation. *Restorative Neurology & Neuroscience, 28*, 259–270.

Schlack, A., Sterbing-D'Angelo, S. J., Hartung, K., Hoffmann, K. P., & Bremmer, F. (2005). Multisensory space representations in the macaque ventral intraparietal area. *Journal of Neuroscience, 25*, 4616–4625.

Sereno, M. I., & Huang, R. S. (2014). Multisensory maps in parietal cortex. *Current Opinion in Neurobiology, 24*, 39–46.

Serino, A., Annella, L., & Avenanti, A. (2009). Motor properties of peripersonal space in humans. *PLoS ONE, 4*, e6582.

Serino, A., Canzoneri, E., & Avenanti, A. (2011). Fronto-parietal areas necessary for a multisensory representation of peripersonal space in humans: An rTMS study. *Journal of Cognitive Neuroscience, 23*(10), 2956–2967.

Serino, A., Noel, J. P., Galli, G., Canzoneri, E., Marmaroli, P., Lissek, H., & Blanke, O. (2015). Body part-centered and full body-centered peripersonal space representations. *Science Reports, 5*, 18603, doi:10.1038/srep18603.

Shore, D. I., Gray, K., Spry, E., & Spence, C. (2005). Spatial modulation of tactile temporal-order judgments. *Perception, 34*, 1251–1262, doi:10.1068/p3313.

Shore, D. I., Spry, E., & Spence, C. (2002). Confusing the mind by crossing the hands. *Cognitive Brain Research, 14*, 153–163, doi:10.1016/S0926-6410(02)00070-8.

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport, 12*, 7–10.

Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences, 1296*, 31–49.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General, 130*, 799–832.

Stanford, T. R., & Stein, B. E. (2007). Superadditivity in multisensory integration: Putting the computation in context. *Neuroreport, 18*, 787–792.

Stein, B., & Stanford, T. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nature Reviews Neuroscience, 9*, 255–266, doi:10.1038/nrn2331.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT.

Stevenson, R. A., Fister, J. K., Barnet, Z. P., Nidiffer, A. R., & Wallace, M. T. (2012a). Interactions between the spatial and temporal stimulus factors that influence multisensory integration in human performance. *Experimental Brain Research, 219*(1), 121–137.

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*, 691–697.

Stevenson, R. A., Zemtsov, R. K., & Wallace, M. T. (2012b). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance, 38*(6), 1517–1529.

Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., . . . Porter, N. R. (2001). When is now? Perception of simultaneity. *Proceedings Biological Sciences / The Royal Society, 268*(1462), 31–38, doi:10.1098/rspb.2000.1326.

Sugita, Y., & Suzuki, Y. (2003, Feb 27). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature, 421*(6926), 911, doi:10.1038/421911a 421911a.

Van der Stoep, N., Nijboer, T. C. W., Van der Stigchel, S., & Spence, C. (2015a). Multisensory interactions in the depth plane in front and rear space: A review. *Neuropsychologia, 70*, 335–349, doi:10.1016/j.neuropsychologia.2014.12.007.

Van der Stoep, N., Van der Stigchel, S., & Nijboer, T. C. W. (2015b). Exogenous spatial attention decreases audiovisual integration. *Attention, Perception, & Psychophysics, 77*(2), 464–482.

Van der Stoep, N., Van der Stigchel, S., Nijboer, T. C. W., & Van der Smagt, M. J. (2015c). Audiovisual integration in near and far space: Effects of changes in distance and stimulus effectiveness. *Experimental Brain Research, 1*–14, doi:10.1007/s00221-015-4248-2.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchronie: A tutorial review. *Attention, Perception, and Psychophysics, 72*, 871–884.

Wallace, M. T., Meredith, M. A., & Stein, B. E. (1992). Integration of multiple sensory modalities in cat cortex. *Experimental Brain Research, 91*, 484–488.

Wallace, M. T., & Stevenson, R. (2014). The construct of the multisensory temporal binding window and

its dysregulation in developmental disabilities. *Neuropsychologia, 64*, 105–123.

Wilkinson, L. K., Meredith, M. A., & Stein, B. E. (1996). The role of anterior ectosylvian cortex in cross-modality orientation and approach behavior. *Experimental Brain Research, 112*(1), 1–10.

Zampini, M., Guest, S., Shore, D., & Spence, C. (2005). Audio-visual simultaneity judgments. *Perception & Psychophysics, 67*(3), 531–544.

Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental Brain Research, 152*, 198–210.