# The Consistency of Crossmodal Synchrony Perception across the Visual, Auditory, and Tactile Senses

Machulla, Tonja-Katrin; Di Luca, Massimiliano; Ernst, Marc

*DOI:*
10.1037/xhp0000191

Running Head: CONSTISTENCY OF CROSSMODAL SYNCHRONY PERCEPTION

# The Consistency of Crossmodal Synchrony Perception across the Visual, Auditory, and Tactile Senses

Tonja-Katrin Machulla

Max Planck Institute for Biological Cybernetics

Massimiliano Di Luca

University of Birmingham

Marc O. Ernst

Bielefeld University

## Author Note

Tonja-Katrin Machulla, Department of Human Perception, Cognition and Action, Max Planck Institute for Biological Cybernetics; Massimiliano Di Luca, School of Psychology, University of Birmingham; Marc O. Ernst, Faculty of Biology, Bielefeld University.

This work was supported by Deutsche Forschungsgesellschaft (Sonderforschungsbereich 550-A11), EU Grants “Immer-Sence” (IST- 2006-027141), and “THE” (IST-2009-248587) and the Max Planck Society. T.-K.M. and M.D.L. contributed equally to this work. We thank Rolf Ulrich for helpful discussions during initial stages of this work, Salvador Soto-Faraco and three anonymous reviewers for their insightful comments.

Correspondence concerning this article should be addressed to Tonja-Katrin Machulla, Max Planck Institute for Biological Cybernetics, Spemannstr. 38, 72076 Tübingen, Germany. E-mail: tonja.machulla@tuebingen.mpg.de.

Abstract

Crossmodal judgments of relative timing commonly yield a non-zero point of subjective simultaneity (PSS). Here, we test whether subjective simultaneity is coherent across all pair-wise combinations of the visual, auditory, and tactile modalities. To this end, we examine PSS estimates for transitivity: if stimulus A has to be presented *x* ms before stimulus B to result in subjective simultaneity, and B *y* ms before C, then A and C should appear simultaneous when A precedes C by *z* ms, where $z=x+y$. We obtained PSS estimates via two different timing judgment tasks — temporal order judgments (TOJ) and synchrony judgments (SJ) — thus allowing us to examine the relationship between TOJ and SJ. We find that (i) SJ estimates do not violate transitivity, and that (ii) TOJ and SJ data are linearly related. Together, these findings suggest that both TOJ and SJ access the same perceptual representation of simultaneity and that this representation is globally coherent across the tested modalities. Furthermore, we find that (iii) TOJ estimates are intransitive. This is consistent with the proposal that while the perceptual representation of simultaneity is coherent, relative timing judgments that access this representation can at times be incoherent with each other due to post-perceptual response biases.

*Keywords:* time perception, crossmodal synchrony, temporal order judgments, synchrony judgments

# The Consistency of Crossmodal Synchrony Perception across the Visual, Auditory, and Tactile Senses

Being able to assess the temporal relationship of sensory signals is considered an important prerequisite to correctly bind corresponding features of a perceptual scene (Eagleman, 2010; King, 2005; Spence & Squire, 2003) and to infer causality (Stetson, Cui, Montague, & Eagleman, 2006). Perhaps counter-intuitively, previous research suggests substantial and systematic deviations of the percept from physical reality: When two discriminable stimuli are presented simultaneously, they are often judged as being successive (e.g., Dinnerstein & Zlotogura, 1968; Dixon & Spitz, 1980; Exner, 1875; Rutschmann & Link, 1964; Rutschmann, 1966; Sternberg & Knoll, 1973; Zampini, Shore, & Spence, 2003). Hence, stimuli have to be presented with a temporal offset to be judged as being simultaneous events. This given offset is commonly referred to as the point of subjective simultaneity (PSS).

This incongruence between physical stimuli and the reported perceptual experience raises an interesting problem. Events in physical time are transitive: If A occurs simultaneously with B, and B simultaneously with C, we can conclude that A is also simultaneous with C. One would assume that such transitive relation also holds true for the subjective temporal relationships between perceptual events. Otherwise, it is not clear how the brain maintains a temporally coherent representation of our environment.

In the present paper, we test whether subjective simultaneity, as measured by the PSS, for all pair-wise combinations of stimuli from three modalities—audiovisual (AV), audiotactile (AT), and visuotactile (TV)—observes a transitive relation. As an example, transitivity of PSS estimates implies that if stimulus A has to be presented 10 ms before

stimulus B for the two to be perceptually simultaneous, and B 20 ms before C, then presenting A 30 ms before C should result in their perceived simultaneity.

In general, a test for transitivity can be interpreted as an evaluation of the consistency of several binary choices. It has been used in research on decision-making and preference choices (Luce & Suppes, 1965). Transitivity of choices suggests a preference ranking that is globally coherent: If bananas are preferred to apples, apples to pears, and bananas to pears, the choices are transitive and the inferred preference ranking is: bananas, apples, pears. On the other hand, if choices are intransitive (bananas are preferred to apples, apples to pears, and *pears to bananas*), no global ranking of preferences can be established. Likewise, transitivity of timing judgments would suggest a temporal representation of the investigated stimulus types (in our case, the sensory modalities that the stimuli are presented to) that is globally coherent. Conversely, intransitivity would suggest that separate processes exist, each of which is specific to a pair-wise combination of stimulus types. Thus, (in)transitivity of a set of crossmodal relative-timing judgments should reveal aspects of the neural processes that underlie multisensory synchrony perception. More pragmatically, it can serve as a generative model for whether results from bimodal experiments are generalizable to multimodal scenarios that involve more than two modalities. If observers' judgments are intransitive, the perceived relationships of three or more concurrently presented stimuli cannot be predicted from bimodal experiments.

While it is intuitively plausible that perception should be coherent—i.e., transitive—across all sources of concurrently presented information, there are two reasons why there may be incoherence across individually elicited bimodal percepts. First, each combination pair of modalities may be processed by a dedicated neural mechanism,

potentially resulting in conflicting estimates of crossmodal synchrony. Subsequently, these conflicts may be reconciled by a higher order integrative process. Second, even if all stimuli were processed by the same neural mechanism, perceptual decisions concerning the relative timing of stimuli from two modalities may vary depending on the presence or absence of stimuli from a third modality. A similar effect has recently been demonstrated (Roseboom, Nishida, & Arnold, 2009): audiovisual timing judgments changed after an additional auditory or visual stimulus was added to the presentation.

While timing judgments are widely employed in cross-modal perception research, most studies employ only one modality-pairing per study (e.g., Allan, 1975; Cairney, 1975; Keetels & Vroomen, 2005; Lewald & Guski, 2003; Rutschmann & Link, 1964; Spence, Shore, & Klein, 2001; Teatini, Farne, Verzella, & Berruecos, 1976; Zampini, 2003; Zampini, Brown, et al., 2005). In the infrequent instances where several pairings are assessed within the same study, the derived PSS estimates are often not submitted to a test of consistency (e.g., Dinnerstein & Zlotogura, 1968; Fink, Ulbrich, Churan, & Wittmann, 2006; Hanson, Heron, & Whitaker, 2008; Harrar & Harris, 2008; Hirsh & Sherrick, 1961). Sternberg and Knoll (1973) review three previous studies that have investigated the transitivity of PSS estimates and report inconsistent findings. In two studies (Corwin & Boynton, 1968; von Békésy, 1963), the estimates did not violate the prediction of transitivity (though one of these studies does not provide the supporting data). In the third study, the estimates are intransitive (Efron, 1963). A more recent report finds that the PSS estimates for the pair-wise combinations of 4 different visual stimuli (two increments in contrast and two orientations of Gabor patches) are intransitive (Cardoso-Leite, Gorea, & Mamassian, 2007). Notably, all previous research on PSS transitivity included intra-modal relative-timing

judgments. The present study is the first to investigate the transitivity of PSS estimates across three different sensory modalities.

There are two commonly used tasks for obtaining PSS estimates—temporal order judgments (TOJ) and simultaneity judgments (SJ). Both are often treated as equivalent and, hence, used interchangeably. In both tasks, observers are presented with two stimuli separated by a varying stimulus onset asynchrony (SOA). The SJ task requires observers to indicate for each presentation whether the two stimuli are synchronous or not (Allan, 1975; Dixon & Spitz, 1980; Fujisaki, Shimojo, Kashino, & Nishida, 2004; Mitrani, Shekerdjiiski, & Yakimof, 1986; Zampini, Guest, Shore, & Spence, 2005; Zampini, Shore, & Spence, 2005). The SOA that corresponds to the maximum proportion of 'synchronous' responses is used as an estimate of the PSS. Alternatively, in the TOJ task observers report which stimulus they perceived first (Alais & Carlile, 2006; Exner, 1875; Hirsh & Sherrick, 1961; Jaskowski, 1993; Machulla, Di Luca, Froehlich, & Ernst, 2012; Rutschmann, 1966; Spence, Baddeley, Zampini, James, & Shore, 2003). The SOA that results in maximal uncertainty about the order of presentation (i.e., 50/50 choices) is taken to correspond to the perception of synchrony. Since TOJ and SJ are assumed to assess the same psychological phenomenon, the point of maximal uncertainty about order in the TOJ task should coincide with the point of maximum synchrony perception as measured by the SJ task. However, several studies have found the PSS estimates obtained with the two tasks to differ and/or to be uncorrelated (Linares & Holcombe, 2014; Love, Petrini, Cheng, & Pollick, 2013; Maier, Di Luca, & Noppeney, 2011; van Eijk, Kohlrausch, Juola, & van de Par, 2008; Vatakis, Navarra, Soto-Faraco, & Spence, 2008). Therefore, in the current study, we employed and compared both tasks to assess the transitivity of PSS estimates across three sensory modalities—vision, touch, and audition.

## Method

### Participants

Twelve participants (6 male; aged 23–28 with a mean age of 26 years) were recruited via the subject database of the Max Planck Institute for Biological Cybernetics, Tübingen. They were paid 8 Euro per hour of participation. The participants were naïve to the purpose of the experiment and reported normal or corrected-to-normal vision, good audition, and no somatosensory disorders. One subject did not finish all experimental sessions; the incomplete dataset was subsequently excluded from further analysis, leaving 11 complete datasets. The study was approved by the local Ethics committee of Tübingen University and all participants gave their informed consent prior to the study.

### Apparatus and Stimuli

Stimuli were generated using Matlab (Mathworks) and a custom-made apparatus (for a picture see Di Luca, Machulla, & Ernst, 2009), which is capable of producing co-located sound, light, and vibration stimuli with high temporal accuracy (0.1 ms). Observers sat at about 50 cm from the apparatus in a dark, sound-attenuated chamber with their extended left index finger placed on the stimulus surface of the apparatus (i.e., the LED light that was mounted on the vibratory shaker between the speakers; see below). They were instructed to maintain fixation at that location throughout the experiment. Tactile stimuli were presented via a vibration device (electro-magnetic shaker, Monacor Bass Rocker BR25), which was mounted on a damping mass, and thus produced tactile stimulation without audible noise. Visual stimuli were presented via an LED display that was mounted in the center on top of the vibration device, serving as the finger contact surface transmitting vibrations as well as the light source (7 x 5 red LEDs, 1.6 x 1.3 cm). The LED array was bigger than the finger, such that the light could easily be seen behind the

finger. Auditory stimuli were presented via two identical speakers, mounted vertically 7.5 cm apart above and below the LED array. Due to the vertical configuration of the speakers the perceived sound location was right at the center, co-located with the light and the touch stimuli. A multi-channel sound card (M-audio 1010LT) together with three identical power amplifiers was used to generate all three stimuli. All stimuli were sinusoidal modulations that were 20 ms in duration, with 5 ms linearly ramped onset and offset. Specifically, the stimuli were a 145 Hz uniform red light, a 1000 Hz audio tone and a 40 Hz tactile vibration. Stimulus intensities were 41 cd/m$^2$ for the visual and 76 dB SPL for the auditory stimuli. Tactile stimuli were of an intensity that created the sensation of a light "tap" on the finger, similar to the vibrating alarm in a cell phone.

**Design and Procedure**

The experiment was run in 6 sessions of 1.5 hours each over the course of 2–3 weeks. Participants were presented with pairings of stimuli in the three possible crossmodal combinations of auditory and visual (AV), auditory and tactile (AT), and tactile and visual (TV). The two stimuli of each pairing were presented at different stimulus onset asynchronies (SOA). In half of the sessions, participants made unspeeded judgments as to which of the two presented stimuli of each pairing had occurred first (temporal order judgment, TOJ). Here, 13 different SOAs ranging from −240 to 240 ms in steps of 40 ms were used. Negative values of SOAs indicate a physical lead of the visual stimulus for the AV and TV stimulus pairings and the tactile stimulus for the AT stimulus pairings. In the other half of the sessions, participants' task was to decide whether the two stimuli had been synchronous or asynchronous (synchrony judgment, SJ). Stimuli were presented with the same SOAs as in the TOJ task with the addition of two SOAs of −400 and 400 ms. That

is, in the SJ task there were 15 different SOAs in total. The order of the two response tasks (TOJ and SJ) was randomly chosen for each participant.

All combinations of modality-pairing and SOA were presented twenty times at random during each experimental session (with a 10-minute rest break in the middle of each session) for a total of 60 repetitions per SOA and modality-pairing for each response task. After each presentation, participants entered their responses with their right hand over the arrow keys of the keyboard. In the TOJ task subjects were asked which modality occurred first, whereas in the SJ task they had to respond whether the stimuli were simultaneous. Since all three modality-pairings were randomly presented in a session, subjects had three possible choices in the TOJ task ("left" arrow key modality 1, "down" arrow key modality 2, and "right" arrow key modality 3) but only two choices ("left" and "right" arrow) for the SJ task (synchronous/non synchronous). The order of assignment of response to key were randomly chosen for each participant and held consistent over all sessions. The next trial was initiated 1.5 seconds after the participant's response. Data from TOJ trials with inappropriate key presses (e.g., a "tactile" response to an audiovisual stimulus) were removed before data analysis (on average, this occurred in only 1 % of the trials).

Before the first session of each kind of response task, participants were acquainted with the procedure and the response button mapping via 100 practice trials. During this practice session, each trial consisted of a randomly selected modality-pairing presented with a random SOA (range 0 ± 1000 ms for TOJ and 0±400 for SJ). During the first 20 of these training trials, auditory feedback was provided after each response indicating which response key had been pressed (however, not whether the response had been correct).

After the practice session, participants were allowed to ask for further clarification or to repeat the training.

**Data Analysis**

For each participant, the data was analyzed separately for each combination of modality-pairing and response task.

**Temporal order judgments.** First, the proportion of "auditory first" responses for AV and AT stimuli and the proportion of "tactile first" responses for TV stimuli was computed as a function of SOA. For TOJ, we define the PSS as the SOA between stimuli X and Y at which participants are equally likely to respond "X first" and "Y first", i.e., the SOA that corresponds to an response probability of 0.5. We also obtained a measure of judgment sensitivity—the *just noticeable difference* (JND), which we defined as half of the difference between the SOAs that correspond to the response probabilities of 0.75 and 0.25. To find these SOAs, we interpolated the response proportions using a nonparametric method and the associated software made available by Zychaluk & Foster (2009). This method is based on local linear fitting; the bandwidth for the local polynomial estimate was chosen from a search interval ranging from 40 ms (minimum difference between the different levels of the predictor variable, SOA) to 480 ms (difference between the maximum and the minimum level of the predictor variable, SOA) by a cross-validation procedure that was provided in the software package. The average bandwidths (standard errors of the mean in parentheses) were 68 ms (8 ms), 63 ms (6 ms), and 70 ms (6 ms) for the AV, AT, and TV pairings, respectively. By using a non-parametric fit, we minimize the assumptions concerning the generative model underlying our data.

**Synchrony judgments.** Here, we computed the proportion of "synchronous" responses as a function of SOA for each of the three modality-pairings (AV, AT, and TV). In

order to estimate the PSS, the values were interpolated with the same method as described for TOJ above. Here, the search interval ranged from 40 ms to 800 ms. The average bandwidths (the standard error of the mean) for the local linear fit were 261 ms (25 ms), 314 ms (30 ms), 295 ms (20 ms) for the AV, AT, and TV pairings, respectively. For SJ, we defined the PSS as the SOA at which participants show the highest probability of reporting "synchronous" perception, i.e., the SOA that corresponded to the maximum value of the fitted function.

To obtain a sensitivity estimate for the SJ that corresponded conceptually to the JND estimate for the TOJ, we proceeded as follows. If both TOJ and SJ were interchangeable and unbiased measures of the same phenomenon, the SJ data, when cumulated and normalized, should be identical to the TOJ data along the entire stimulus dimension, i.e., at every SOA, not only at the estimated PSS. That is, the cumulation of the roughly bell-shaped SJ curve should lead to a sigmoidal curve that is similar in shape to the response curve obtained with the TOJ task. Thus, we cumulated the SJ responses up to the PSS estimate (i.e., the data on the left side of the maximum value) and normalized the resulting data such that the maximum value corresponded to 0.5. The same procedure was then applied to the responses on the other side of the maximum value and 0.5 added to these data. From the resulting cumulative distribution, the JND estimates were obtained as described for TOJ.

**Test for transitivity.** For each participant, we obtained three PSS values ($PSS_{AV}$, $PSS_{AT}$, and $PSS_{TV}$) for both TOJ and SJ. If these estimates were transitive, individual data should adhere to the equation $PSS_{AT} + PSS_{TV} - PSS_{AV} = 0$. This equation describes a planar surface in a three-dimensional PSS space, henceforth referred to as *transitivity plane*. In this space, each participant's PSS data can be conjointly represented by a point that should

lie on the transitivity plane if a given participant exhibits transitivity in his PSS data. If deviations from this prediction were the result of random errors in the estimation of single PSS estimates, data points should be equally likely to be situated to both sides of the plane. Therefore, we computed the orthogonal signed distance of each data point from the transitivity plane (henceforth, transitivity distance) and conducted a sign test on the results.

## Results

### Analysis of TOJ and SJ Data

Figure 1 shows the SJ and TOJ data together with the fitted psychometric functions for one participant. TOJ data are indicated by light grey circles, SJ data by dark grey circles. Mean PSS and JND values for the three modality-pairings and mean distance from the transitivity plane with corresponding standard errors are presented in Table 1. Since we were interested in whether non-zero PSS estimates of different modality-pairings are coherent with each other, at least one of the three PSS distributions (AV, AT, or TV) has to differ from zero. Otherwise, a transitive result would be trivial (i.e., $0\ ms\ (PSS_{AT}) + 0\ ms\ (PSS_{TV}) - 0\ ms\ (PSS_{AV}) = 0\ ms$). Therefore, we tested for each combination of modality-pairing and judgment task whether the mean PSS value differed from zero. For the TOJ data, this was the case for the mean $PSS_{AT}$ and $PSS_{TV}$ estimates (Table 2a; test statistics and p-values are summarized in Table 2). For the SJ data, mean $PSS_{AT}$ estimates differed significantly from 0 (Table 2b). This provides some indication that transitivity of PSS estimates would be nontrivial for both judgment tasks.

Nevertheless, this analysis cannot rule out that transitivity is ‘trivial’ at the level of individual participants. To test for this possibility, we conducted an additional analysis. We applied a bootstrap procedure to each individual’s TOJ and SJ data. This analysis simulates

the experiment repeatedly based on the empirically obtained data. The resulting distributions of PSS estimates allow us to test for statistically significant deviations from 0 on the level of individual participants. At each SOA, we sampled from a binomial distribution with parameters *p* and *n*, where *n* is the number of presentations of the stimulus pair with a particular SOA (e.g., 60 ms) and *p* is the proportion of 'synchronous' responses (for the SJ task) or the proportion of 'modality x-first' responses (for the TOJ task; 'x' stands for 'auditory' in the case of AV and AT pairs and 'tactile' in the case of VT stimulus pairs) by the participant. The resulting data was fitted as described above and the PSS estimate extracted. This procedure was repeated 1000 times, for each modality-pairing. The resulting PSS estimate distributions were used to obtain confidence intervals for participants' original PSS data by taking the 2.5 and 97.5 percentiles of the bootstrap distribution as the lower and upper bounds, respectively. In case of TOJ, for each participant the confidence intervals for at least one modality combination did not include 0. In case of SJ, for 10 out of 11 participants the confidence intervals for at least one modality combination did not include 0. Thus, transitivity would be non-trivial for these individual data sets.

Figure 2 shows each participant's PSS data in relation to the transitivity plane from two different vantage points. For better visualization, the transitivity distance is also available in a box-plot representation in Figure 3. To reiterate, if deviations from the transitivity plane were due to random error, data points should be located to both sides of the plane. For the TOJ task, data from all participants fall to the same side of the transitivity plane. That is, the PSS estimates obtained with TOJ are not transitive (sign test: $p<0.001$). For the SJ task, the pattern of transitivity distances clearly differs from the one found with TOJ data: Data points are located to both sides of the transitivity plane (sign-

test: p = 0.55). In other words, in contrast to the TOJ data, the SJ data is statistically indistinguishable from being transitive (that is, the sum $PSS_{AT} + PSS_{TV} – PSS_{AV}$ is indistinguishable from zero).

To strengthen this conclusion, we conducted a power analysis for the SJ data. Specifically, we computed the power for the specific intransitivity instance where all data points are located to the same side of the transitivity plane (as in the case of TOJ), using the power function provided by Dixon (1953). For this, we assumed that, given the H1-hypothesis, the SJ data were a randomly generated sample from a binomial distribution with parameters n=11 and p=0.99 (i.e., a close approximation of the case p=1.0, where all data points will always fall on the same side of the plane). The obtained power is 0.9948, which is sufficient to support our conclusion (Cohen, 1988). Furthermore, the transitivity distance estimates obtained with TOJ and SJ are significantly different from each other (two-tailed paired-sample t-test, t(10) = 3.46, p = 0.0025) and do not correlate (r = 0.54, p = 0.09; differences and similarities in PSS and JND estimates across SJ and TOJ are discussed later). Together, these results indicate that it is very unlikely that the SJ data is intransitive in the same way as the TOJ data.

However, these analyses can potentially hide individual data patterns that do not conform to the conclusions drawn at the group level. For instance, it is possible that the SJ PSS data are intransitive at the level of individual participants, even though data points are distributed evenly to both sides of the transitivity plane. Therefore, we tested individual participants' data for deviations from transitivity, separately for the TOJ and the SJ methods. For this, we used the $PSS_{AV}$, $PSS_{AT}$, and $PSS_{TV}$ estimates obtained via the above-mentioned bootstrap procedure to compute multiple estimates of the distance from the transitivity plane, yielding 1000 distance estimates for each participant. The resulting

distribution of estimates was used to construct 95% confidence intervals for each participant's distance estimate (the 2.5 and 97.5 percentiles of the distribution served as the lower and upper bounds, respectively). In general, the results substantiate the findings obtained at group level (Figure 4). For the TOJ, 9 of the 11 participants have confidence intervals that do not overlap with the transitivity plane. In other words, their PSS estimates are intransitive. In contrast, all except for one participant's confidence intervals for the SJ data overlap with the transitivity plane. That is, the SJ PSS values are indistinguishable from being transitive (for nine of these data sets transitivity can be considered non-trivial). As for the one participant whose distance estimate deviates significantly from transitivity, we did not find a convincing explanation for why this data set might have diverged from the general result. His PSS values were neither extreme, nor his pattern of intransitivity particularly distinctive.

**Comparison of TOJ and SJ Data**

TOJ and SJ are often treated as interchangeable measures of the same underlying psychological phenomenon — the perception of simultaneity of two stimuli. However, a number of studies report that the two tasks can yield different results (Love et al., 2013; Maier et al., 2011; van Eijk et al., 2008; Vatakis et al., 2008). Similarly, we also find that data collected with TOJ and SJ differ: While the PSS estimates obtained with TOJ are intransitive, the PSS estimates obtained with SJ are not. Here, we conduct additional comparisons of the data obtained with the two tasks to further characterize the relationship between TOJ and SJ.

Figure 5 displays the distributions of individual PSS (upper panel) and JND (lower panel) values obtained with both tasks plotted against each other. There are a number of observations we can make from these data; test statistics and p-values are again

summarized in Table 2. First, the $PSS_{AV}$ and $PSS_{AT}$ estimates obtained with the two judgment tasks are significantly correlated and the correlation of the $JND_{TV}$ is marginally significant (Table 2c and 2d). However, in the case of $PSS_{AT}$ statistical significance depends on the inclusion of one data point that deviates to some degree from the norm. As a result, only one of the six correlation analyses between TOJ and SJ data (i.e., audiovisual PSS) provides us with a clear indication that the two tasks measure the same underlying phenomenon. This failure to find stronger evidence for a relationship between TOJ and SJ data may have resulted from the comparatively low number of data points entering into each correlation analysis. The limited number of points could have introduced random factors with sufficient noise to potentially mask an actual relationship. To alleviate this shortcoming, we conducted two separate linear mixed effects analyses of TOJ as a function of SJ, one for the PSS and one for the JND data. In essence, this type of analysis allowed us to pool the data across the different modality-pairings, thus increasing the number of data points entering into the analysis. This procedure is justifiable because here we were interested in the relationship between TOJ and SJ regardless of the nature of the stimuli that had to be judged. In our model, we predicted the TOJ data from the fixed effect *SJ data* and the random effects *participant* and *modality-pairing*, allowing for individual intercepts for the random effects. There were no systematic deviations in the residuals and they were indistinguishable from being normally distributed (as determined by visual inspection). To screen for influential data points (i.e, data points that significantly alter the outcome of the analysis), the analysis was repeated multiple times, each time on the whole data set minus the data point under investigation. No influential data points were detected in this manner. Statistical significance of the linear relationship was determined via a

likelihood test of the full model against a model without the fixed effect. The strength of the relationship between TOJ and SJ data was assessed by computing two versions of $R^2$: marginal $R^2$, which is based on the proportion of variance explained by the fixed effect, and conditional $R^2$, which is based on the proportion of variance explained by fixed and random effects together (Nakagawa & Schielzeth, 2013). For both PSS and JND data, performance on the SJ task predicted performance on the TOJ task (PSS: $\chi^2(1) = 8.9$, $p = 0.003$, $R^2_{marginal} = 0.26$, $R^2_{conditional} = 0.59$; JND: $\chi^2(1) = 9.45$, $p = 0.002$, $R^2_{marginal} = 0.24$, $R^2_{conditional} = 0.62$). Further, in both cases, the estimated slope for the fixed effect was positive indicating that larger TOJ PSS were associated with larger SJ PSS and larger TOJ JND with larger SJ JND (slopes of 0.84 and 0.42, respectively). These findings provide substantial evidence for a positive linear relationship of moderate strength between the data collected with the two relative timing measures. This indicates that TOJ and SJ access, at least to some degree, the same internal response.

Further comparisons of the TOJ and SJ data reveals that the SJ and TOJ PSS distributions differ significantly in their means for the TV pairing but not for the AV and AT pairings (Table 2e)—the TOJ $PSS_{TV}$ are further removed from physical synchrony compared to the other PSS values. This suggests that the main cause for the found difference in transitivity between TOJ and SJ lies in the TV modality-pairing. The JND estimates do not differ significantly in their means across the two judgment tasks (Table 2f).

## Discussion

The goal of the current work was to examine whether perceived synchrony, as measured by the PSS, is consistent across the auditory, visual, and tactile domains. This was achieved by assessing the transitivity of the PSS estimates of crossmodally-paired stimuli. We used two different types of relative-timing judgments to determine perceived

simultaneity: temporal order judgments (TOJ) and simultaneity judgments (SJ). Thus, our experimental design also allows us to contribute evidence to a current debate in the literature on relative-timing measurement, namely, the relationship between TOJ and SJ. To the best of our knowledge, this study is the first to compare audiovisual, audiotactile, and visuotactile data across the two judgment tasks. This stands in contrast to previous comparisons of TOJ and SJ, which were largely restricted to audiovisual data. To summarize, the current study yields two main observations: (i) while estimates obtained with SJ do not violate the prediction of transitivity, estimates obtained with TOJ are intransitive. (ii) In contrast to several previous reports, we find that perceived synchrony (PSS) as well as judgment sensitivity (JND) estimates can show high levels of agreement across TOJ and SJ. In the following section, we will further discuss these findings and their implications.

**A model for the (in-)transitivity of timing judgments**

Our test of PSS transitivity suggests that SJ of crossmodally-paired stimuli are based on a global multimodal representation of relative timing that is consistent across visual, auditory, and tactile events. In contrast, the intransitivity of the PSS estimates obtained with TOJ reflects processing inconsistencies between different crossmodal pairings, which reflect modality-pairing-specific processes. In the following, we will explore where in the stimulus-response process the inconsistency may have originated (for a similar approach see Sternberg & Knoll, 1973). For this, we will rely on a simple model of how relative-timing judgments are generated. This model separates the stimulus-response process into two sequential stages: an initial sensory stage encompassing the neural response to the sensory stimulation, and a subsequent decision stage encompassing all decision processes that are related to the task that is to be performed (here, either SJ or TOJ). Inconsistencies

between different crossmodal judgments can originate at either stage. At the sensory stage, inconsistency can result if the processing of a signal varied according to the nature (e.g., modality) or the relative timing of accompanying signals. For example, the processing of signal *a* might be accelerated in the presence of signal *b* but decelerated in the presence of signal *c*. At the decision stage, inconsistency can result from differences in the setting of decision criteria or response strategies between the modality-pairings. Such modality-pairing-specific decision processes—if not coordinated by a higher-level instance—are likely to result in decisions that are not coherent among each other, e.g., *a* and *c* cannot reliably be ordered, even though *a* is judged to occur before *b*, and *b* before *c*.

If we assume these two successive stages of processing, Table 3 summarizes the four possible combinations of processing coherence between the stages and their individual predictions for data (in)transitivity. Inconsistent processing of modality-pairings at either stage will lead to intransitivity of judgments (Table 3, rows 2–4). In contrast, transitive judgments should result if both sensory processing and task-related decision processes are consistent (Table 3, row 1). In principle, it is also possible for judgments to be transitive if the inconsistencies at the two stages were equal in magnitude but opposite in sign, thus canceling each other (Table 3, row 4). However, we deem this last scenario unlikely. Such a cancelation could either result from a fortuitous coincidence, which we exclude as a useful explanation for our data, or it could indicate some form of compensation mechanism—namely, the compensation of incoherence at the sensory stage by the decisional processes of the second stage of SJ. We can think of two arguments against such a compensation mechanism. First, there is no obvious reason why compensation would only take place for SJ and not TOJ. Second, it is not clear how the degree of compensation within one modality-pairing ought to be appropriately determined

in the absence of the third stimulus. Inconsistency in processing, as described here, arises between three stimulus pairings. Compensation, on the other hand, would have to take place within the processing of single pairings, that is, in the absence of the respective third stimulus. To compensate for inconsistency, some form of knowledge of the absent third stimulus would be required during the processing of each pairing. Since the number and type of stimuli that could be added to the pair is infinite, compensation for all possible cases is unlikely. Therefore, row 1 of Table 3 presents what we deem the most likely case of how transitive data is generated, namely that processing at both the sensory and the decision stage is coherent across the tested modalities. Hence, we can conclude that for SJ processing is coherent at both stages.

With regards to the TOJ data, there remain three plausible options for where in the processing pathway the inconsistency could occur (Table 3, rows 2-4). To localize the source of intransitivity for TOJ, we must first examine the relationship between TOJ and SJ, a topic that has generated some recent debate.

**The relationship between Temporal Order Judgments and Synchrony Judgments**

When measuring synchrony perception, TOJ and SJ are often used interchangeably. This practice reflects a generally accepted assumption that both tasks are equivalent measures of the same perceptual phenomenon. However, experimental evidence concerning this assumption is mixed. On the one hand, several studies have reported that the results obtained with the two tasks can diverge. PSS and JND estimates (usually for the AV modality-pairing) have been found to differ and/or to be uncorrelated (Linares & Holcombe, 2014; Love et al., 2013; Maier et al., 2011; for a review see van Eijk et al., 2008), or to be unequally affected by manipulations of stimuli and paradigm (Mitrani et al., 1986; Mossbridge, Fitzgerald, O'Connor, & Wright, 2006; Vatakis et al., 2008). From such results,

it has commonly been concluded that the two types of judgments are mediated via separate and possibly independent processing systems (Jaskowski, 1991; Love et al., 2013; Mitrani et al., 1986; Mossbridge et al., 2006; Pöppel, 1997; van Eijk et al., 2008; Vatakis et al., 2008).

On the other hand, several suggestions have been made for how divergent PSS and JND estimates can be explained within a framework that assumes some commonality in processing between the TOJ and SJ task (Allan, 1975; García-Pérez & Alcalá-Quintana, 2012; García-Pérez & Alcalá-Quintana, 2015; Yarrow, Jahn, Durant, & Arnold, 2011). For example, estimates obtained with either measure are prone to decisional criterion shifts and strategic responding (Spence et al., 2001; Yarrow et al., 2011). Therefore, differences in criterion settings between TOJ and SJ may well account for the reported differences in the PSS and JND estimates while the two tasks could still be based on the same internal sensory signal. Additionally, PSS or JND estimates have occasionally been found to correlate across task types (Sanders, Chang, Hiss, Uchanski, & Hullar, 2011; Smeele, 1994; van Eijk et al., 2008). This should not be the case if the two types of judgments were based on independent processes. While it is possible that significant correlations result from statistical type-II error, it is far more likely for a failure to obtain correlated estimates (as previously reported) to result from type-I error, i.e., low statistical power. In conclusion, neither differences in estimates across TOJ and SJ nor uncorrelated estimates are unequivocal indicators for a separation in the processing of simultaneity and temporal order. Lastly, some studies have found similarities in how the SJ and TOJ PSS recalibrate after prolonged exposure to a crossmodal, temporal discrepancy (Fujisaki, Shimojo, Kashino, & Nishida, 2004; Heron, Hanson, & Whitaker, 2009; Vroomen, Keetels, de Gelder,

& Bertelson, 2004). This provides further reasons to question the postulation that separate processing systems exist for TOJ and SJ.

How does the present work fit in with this debate? In contrast to previous studies, which focused on the comparison of PSS and JND estimates for AV stimuli across response tasks, the present research is the first to also compare estimates for AT and VT stimuli. To reiterate our main result, we found that PSS estimates show a positive linear relationship across the TOJ and SJ tasks. This provides substantial evidence against a strict separation of simultaneity and order processing. In other words, TOJ and SJ are based, at least partly, on a common process rather than being entirely independent measures.

Three factors may have contributed to why we found a significant relationship between TOJ and SJ while previous work did not. First, we used a comparably large number of repetitions per stimulus pair, namely 60, while others used less (e.g., 12 repetitions in Love et al., 2013). The smaller the number of stimulus repetitions, the more likely it is to obtain estimates that are largely dominated by noise. This would greatly diminish the chance of detecting any true relationships between the data obtained via TOJ and SJ. Second, we fitted our data with a nonparametric method while most previous studies fitted their data with parametric functions that assume symmetry of sensitivity thresholds to both sides of the PSS. By using a non-parametric fit, we avoided making any such assumption and allowed data asymmetries around the PSS to be accounted for. A potential problem of the parametric approach is that fitting skewed data with a symmetric function will inadvertently result in a biased PSS estimate. The degree of fitting bias could vary between TOJ and SJ data, resulting in differences in fitted parameters of interest, thus masking the true relationship between the parameters. Finally, we interleaved the presentation of stimulus pairs from the three modality-pairings. In contrast, previous

studies investigated only one pairing of modalities per experiment. That is to say, each modality-pairing was tested in a blocked fashion. A blocked presentation of stimuli might have led to the cultivation of particular response strategies during experimentation, such as always choosing the visual stimulus when uncertain about order. In contrast, intermixed presentation, as in the present study, could prevent very strong preferences from developing during testing.

The fact that JND as well as PSS estimates obtained via TOJ and SJ are related has two interesting consequences beyond the fact that both tasks rely on a common internal process. First, we can conclude that under conditions, which are yet to be specified, PSS and JND estimates can be obtained with some reliability in spite of TOJ and SJ being bias-prone measures. Second, we can draw conclusions concerning the nature of interindividual differences in relative-timing perception. Participants with larger SJ PSS are likely to also have larger TOJ PSS. This lends credit to the existence of a “personal equation”, namely, the idea that individual deviations from the group mean are not merely due to noise but rather reflect systematic interindividual differences in the interval at which two events are perceived as synchronous (see also Stone et al., 2001). This was posited in the 19$^{th}$ century after it was observed that astronomers differed consistently in their report of a star’s location on the telescope image plane at a time point marked by the beat of a clock (Mollon & Perkins, 1996). Similarly, participants with larger SJ JND are likely to also have larger TOJ JND. This suggests that differences in sensitivity are observer-specific rather than arbitrary. The JND is sometimes used as an estimate of the temporal window for multisensory binding since multisensory integration scales with the temporal distance between two events (King & Palmer, 1985; Meredith, Nemitz, & Stein, 1987).

Consequently, the present consistency in sensitivity differences across the two tasks favors the interpretation that different observers vary in the size of their integration windows.

**Why are TOJ intransitive?**

Where in the stimulus-response process does the inconsistency of TOJ originates? Earlier, we listed three possible combinations of processing coherence at the two stages of the stimulus-response process that result in the intransitivity in TOJ (Table 3, rows 2–4). So far, we have established that SJ processing is coherent at both stages and TOJ and SJ share parts of their processing pathway. Hence, TOJ processing must be consistent at one of the two stages and inconsistent at the other. This leaves two possibilities. First, TOJ are inconsistent at the second stage. That is, the observed intransitivity is due to differences in criterion settings and/or response strategies between pairings (row 2, Table 3). In this case, TOJ and SJ are based on a common internal signal but rely on distinct and potentially uncorrelated decision processes. Alternatively, TOJ are inconsistent at the first stage, e.g., due to intransitive nonlinearities in early sensory processes, such as modality-pairing-dependent accelerations in neural processing speed (row 3, Table 3). In this case, TOJ and SJ are based on distinct internal responses but are subject to correlated decision processes. Of these two interpretations, we favor the first. There are two reasons for this. First, the suggestion that TOJ and SJ share sensory processes is more conservative. The alternative account would not only require decision processes to correlate between two measures that are likely prone to different types of bias but also would these biases have to be consistent across modality-pairings, such that SJ are transitive. This seems very unlikely. Also, it is reasonable to assume that much of the sensory processing should be shared between the two tasks. After all, the sensory stimulation is the same in both tasks — it is the decision process that distinguishes a temporal order judgment and a simultaneity

judgment. Second, we find that the differences between individual observers' TOJ PSS and SJ PSS for the same modality-pairing can be as large as 160 ms. It is very unlikely for such large discrepancies to be due to differences in sensory processing alone. Therefore, it is reasonable to assume that the observed differences between TOJ and SJ responses originate in the decision stage, that is, in the second stage of processing. Several recent publications with an emphasis on a model-based analysis of TOJ and SJ data conform with this conclusion (García-Pérez & Alcalá-Quintana, 2012; García-Pérez & Alcalá-Quintana, 2015; Yarrow et al., 2011)

It is possible to further specify the source of TOJ intransitivity by taking the relationship between the PSS estimates obtained with TOJ and SJ into account. Of the six obtained PSS measures, five are consistent with each other: For SJ, the three PSS estimates are indistinguishable from being transitive and for TOJ, two of the PSS estimates ($PSS_{AV}$ and $PSS_{AT}$) do not differ significantly from their SJ counterparts. Only the $PSS_{TV}$ estimates differ between the two tasks. For this comparison, the PSS estimates obtained with TOJ are substantially further away from 0. This shift of the $PSS_{TV}$ estimates is likely the biggest factor contributing to the present intransitivity of TOJ.

Here, we can only speculate why the $PSS_{TV}$ estimates alone should be affected by a criterion shift or a response strategy. It is possible that relative timing judgments are made by relying on only one modality—for example, the auditory since it is often the most reliable with regard to temporal information—as a reference or temporal anchor that decreases the probability of biased judgments. TV judgments by themselves would not have access to this reference frame and therefore might be more easily biased. Also, TV stimuli that are not accompanied by auditory information are quite rare in natural environments. Almost every tactile stimulation or haptic interaction produces a sound. On

the other hand, AV stimulation without accompanying tactile information or AT stimulation without accompanying visual information are much more common and provide the observer with more opportunities to observe and learn to properly judge the temporal relationship between the involved modalities.

Finally, we need to discuss why inconsistency at the decision level occurred for TOJ but not for SJ. One possible explanation is that the two tasks differed in their level of difficulty. For the present study, we chose to present the three different modality-pairings in an interleaved rather than blocked fashion (i.e., a separate experimental block for each pairing). Both presentation modes have their drawbacks. Blocked presentation may lead to particular attentional distribution patterns that result in inconsistent prior entry effects (e.g., attending to audition in AV and AT pairings and vision in TV pairings; for a review on the prior entry effect see Spence & Parise, 2010), thus causing artifactual intransitivity of judgments. On the other hand, interleaving modality-pairings may raise the level of difficulty for both tasks but potentially more so for the TOJ. First, interleaved presentation requires the observer to simultaneously monitor three different sensory channels. At this point, it is not clear whether this demand exceeds the attentional capacity of the average observer. Arguably, under high processing demands it may be harder to determine the order of stimuli across three modalities than simply stating whether stimuli were synchronous or not. Second, the interleaving of modality-pairings increases the number of response options for TOJ but not for SJ. In our case, only 1% of answers involved the sense that was not presented on any given trial (e.g., “auditory-first” answer for a VT stimulus pair). However, we cannot exclude the possibility that the increased difficulty of three answer options resulted in participants repeatedly pressing the wrong button (“auditory-first” answer when “visual-first” was perceived). Despite these drawbacks, we chose

interleaved over blocked presentation with the following reasoning in mind: a criterion-independent perceptual threshold model is commonly applied when interpreting data obtained with TOJ and SJ. That is, it is implicitly assumed that the PSS and JND estimates are unbiased measures of a perceptual state (namely, perception of synchrony). According to such a model, task difficulty should only affect the JND — an increase in difficulty should result in a decrease in judgment sensitivity but not in a shift of the PSS. To date, there is little experimental evidence to explicitly expect this assumption to be violated (of course, the violation cannot be excluded). In contrast, there is substantial empirical evidence in favor of prior entry effects (Spence & Parise, 2010). Therefore, we chose the experimental design that, in our opinion, was less likely to lead to intransitive judgments in the case that bimodal percepts are consistent.

**Inconsistent percepts or inconsistent judgments?**

The present study set out to investigate whether the perceived simultaneity of crossmodally-paired stimuli is coherent across all pair-wise combinations of the visual, auditory, and tactile modalities. Unfortunately, it is not known to date in how far the most commonly used tasks used to assess perceived synchrony—the TOJ and the SJ task—provide veridical measures of the underlying percept. Both have been shown to be vulnerable to shifts in decisional criterion and strategic responding (Schneider & Bavelier, 2003; Shore, Spence, & Klein, 2001; Spence et al., 2001; Yarrow et al., 2011). In particular, response strategies of post-perceptual (i.e., cognitive) nature may cause the PSS estimate to substantially deviate from the percept. To illustrate, a major assumption of the TOJ task holds that whenever participants are maximally unsure about stimulus order (namely, whenever stimuli are perceptually synchronous), they divide their responses equally between the two available options. Nonetheless, it is feasible that participants might,

instead, demonstrate a preference for one of the two options in perceptually indeterminate situations (e.g., “vision-first” responses for AV stimulus pairings). Such a preference would introduce a systematic bias in both the PSS and the JND estimate. Due to this, the use of TOJ and SJ can only reveal the consistency of participants’ *answer patterns* across several judgments, rather than the desired aim of measuring the coherence of several bimodal *percepts*. In spite of this, our current findings allow us to make an educated guess concerning the coherence of percepts. For the SJ task, we believe that the fact that PSS estimates did not deviate from transitivity strongly suggests a coherence of the underlying percepts. Incidental transitivity resulting from consistent post-perceptual answer strategies is unlikely. In contrast, the intransitivity of PSS estimates obtained with the TOJ task does not provide sufficient support in favor of or against the coherence of the associated bimodal percepts. Here, intransitivity may have resulted from perceptual incoherence or it may have resulted from incoherence in post-perceptual answer strategies. Of these two accounts, the latter is far more likely given the substantial linear relationship between the results obtained with TOJ and SJ. This suggests that TOJ access the same perceptually coherent representation of relative timing as do SJ. Accordingly, any differences in PSS estimates between the TOJ and the SJ task can be ascribed to differences in post-perceptual decision processes rather than differences in the perceptual estimates.

To summarize the previous two sections of this discussion, we conclude that the intransitivity of TOJ results from inconsistent criterion shifts between the three different modality-pairings. We believe the origin of these shifts to be cognitive rather than perceptual.

Finally, an important point needs to be clarified. We do not intend to make a general statement such as ‘TOJ are always intransitive, while SJ are not’. Under different

experimental conditions, the intransitivity of PSS estimates obtained with TOJ may not be found or PSS estimates obtained with SJ may be intransitive (in fact, the analysis on the level of individual participants revealed that the SJ PSS data for one participant was indeed intransitive). Rather, we argue that, due to the susceptibility of TOJ and SJ to criterion shifts, it cannot be known with 100% confidence whether or not non-perceptual biases are present in any particular experimental situation. Non-violation of transitivity between different judgments may provide some indication to the absence of such biases, or, at the very least, provide justification for the generalization of experimental results from bimodal to multimodal scenarios.

**Conclusion**

In conclusion, the current study contributes to the literature on relative-timing perception in two ways. First, our data does not support the view that TOJ and SJ are based on two independent processing systems. On the contrary, we conclude that both tasks access the same internal representation of perceptual synchrony. We believe that the differences in synchrony estimates between TOJ and SJ, which are sometimes reported, can be succinctly explained by differences in task-related decision processes of post-perceptual origin. Second, the processing of crossmodal relative timing of visual, auditory, and tactile events is mediated via a neural mechanism that is globally consistent, rather than by a number of independent subprocesses specific to each pair-wise combination of modalities. Once again, differences in post-perceptual decision processes is the most likely explanation for inconsistent (i.e., intransitive) response patterns.

References

Alais, D., & Carlile, S. (2006). Synchronizing to real events: Subjective audiovisual alignment scales with perceived auditory depth and speed of sound. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(6), 2244–2247. doi:http://dx.doi.org/10.1073/pnas.0407034102

Allan, L. G. (1975). The relationship between judgments of successiveness and judgments of order. *Perception & Psychophysics*, *18*(1), 29–36. doi:10.3758/BF03199363

Cairney, P. T. (1975). Bisensory order judgement and the prior entry hypothesis. *Acta Psychologica*, *39*(5), 329–340. doi:10.1016/0001-6918(75)90025-6

Cardoso-Leite, P., Gorea, A., & Mamassian, P. (2007). Temporal order judgment and simple reaction times: Evidence for a common processing system. *Journal of Vision*, *7(6):11*, 1–14. doi:10.1167/7.6.11

Corwin, T. R., & Boynton, R. M. (1968). Transitivity of visual judgments of simultaneity. *Journal of Experimental Psychology*, *78*(4), 560–568. doi:10.1037/h0026617

Di Luca, M., Machulla, T.-K., & Ernst, M. O. (2009). Recalibration of multisensory simultaneity: Cross-modal transfer coincides with a change in perceptual latency. *Journal of Vision*, *9(12):7*, 1–16. doi:10.1167/9.12.7

Dinnerstein, A. J., & Zlotogura, P. (1968). Intermodal perception of temporal order and motor skills: effects of age. *Perceptual and Motor Skills*, *26*, 987–1000. doi:10.2466/pms.1968.26.3.987

Dixon, N., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, *9*, 719–721. doi:10.1068/p090719

Eagleman, D. M. (2010). How does the timing of neural signals map onto the timing of perception? In R. Nijhawan & B. Khurana (Eds.), *Space and time in perception and action* (pp. 216–231). UK: Cambridge University Press. doi:10.1017/CBO9780511750540.014

Efron, R. (1963). The effect of handedness on the perception of simultaneity and temporal order. *Brain*, *86*(2), 261–284. doi:10.1093/brain/86.2.261

Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Processe. *Archiv für die gesamte Physiologie des Menschen und der Tiere*, *11*(1), 403–432. doi:10.1007/BF01659319

Fink, M., Ulbrich, P., Churan, J., & Wittmann, M. (2006). Stimulus-dependent processing of temporal order. *Behavioural Processes*, *71*(2-3), 344–52. doi:10.1016/j.beproc.2005.12.007

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, *7*(7), 773–778. doi:10.1038/nn1268

García-Pérez, M. A., & Alcalá-Quintana, R. (2012). On the discrepant results in synchrony judgment and temporal-order judgment tasks: a quantitative model. *Psychonomic Bulletin & Review*, *19*(5), 820–846. doi:10.3758/s13423-012-0278-y

García-Pérez, M. A., & Alcalá-Quintana, R. (2015). Converging evidence that common timing processes underlie temporal-order and simultaneity judgments: a model-based analysis. Attention, Perception, & Psychophysics, 77(5), 1750-1766. doi: 10.3758/s13414-015-0869-6

Hanson, J. V. M., Heron, J., & Whitaker, D. (2008). Recalibration of perceived time across sensory modalities. *Experimental Brain Research*, *185*(2), 347–352. doi:10.1007/s00221-008-1282-3

Harrar, V., & Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Experimental Brain Research*, *186*(4), 517–524. doi:10.1007/s00221-007-1253-0

Heron, J., Hanson, J. V. M., & Whitaker, D. (2009). Effect before cause: supramodal recalibration of sensorimotor timing. *PloS One*, *4*(11), e7681. doi:10.1371/journal.pone.0007681

Hirsh, I. J., & Sherrick, C. E. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology*, *62*(5), 423–32. doi:10.1037/h0045283

Jaskowski, P. (1991). Two-stage model for order discrimination. *Perception & Psychophysics*, *50*(1), 76–82. doi:10.3758/BF03212206

Jaskowski, P. (1993). Selective attention and temporal-order judgment. *Perception*, *22*, 681–689.

Keetels, M., & Vroomen, J. (2005). The role of spatial disparity and hemifields in audio-visual temporal order judgments. *Experimental Brain Research*, *167*(4), 635–640. doi:10.1007/s00221-005-0067-1

King, A. J. (2005). Multisensory integration: Strategies for synchronization. *Current Biology*, *15*(9), R339–341. doi:10.1016/j.cub.2005.04.025

King, A. J., & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research*, *60*, 492–500. doi:10.1007/BF00236934

Linares, D., & Holcombe, A. O. (2014). Differences in perceptual latency estimated from judgments of temporal order, simultaneity and duration are inconsistent. I-Perception, 5(6), 559-571. doi: 10.1068/i0675

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. doi:10.1016/S0926-6410(03)00074-0

Love, S. A., Petrini, K., Cheng, A., & Pollick, F. E. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments. *PloS One*, *8*(1), e54798. doi:10.1371/journal.pone.0054798

Luce, R. D., & Suppes, P. (1965). Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology III* (pp. 249–410). New York: Wiley.

Machulla, T.-K., Di Luca, M., Froehlich, E., & Ernst, M. O. (2012). Multisensory simultaneity recalibration: storage of the aftereffect in the absence of counterevidence. *Experimental Brain Research*, *217*(1), 89–97. doi:10.1007/s00221-011-2976-5

Maier, J. X., Di Luca, M., & Noppeney, U. (2011). Audiovisual asynchrony detection in human speech. *Journal of Experimental Psychology. Human Perception and Performance*, *37*(1), 245–256. doi:10.1037/a0019952

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration neurons. I. Temporal factors in superior colliculus. *The Journal of Neuroscience*, *7*(10), 3215–3229.

Mitrani, L., Shekerdjiiski, S., & Yakimof, N. (1986). Mechanisms and asymmetries in visual perception of simultaneity and temporal order. *Biological Cybernetics*, *54*, 159–165. doi:10.1007/BF00356854

Mollon, J., & Perkins, A. (1996). Errors of judgement at Greenwich in 1796. *Nature*, *380*, 101–102. doi:10.1038/380101a0

Mossbridge, J. A., Fitzgerald, M. B., O'Connor, E. S., & Wright, B. A. (2006). Perceptual-learning evidence for separate processing of asynchrony and order tasks. *The Journal of Neuroscience*, *26*(49), 12708–12716. doi:10.1523/JNEUROSCI.2254-06.2006

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining $R^2$ from generalized linear mixed-effects models. *Methods in Ecology and Evolution 4*(2), 133-142. doi: 10.1111/j.2041-210x.2012.00261.x

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences*, *1*(2), 56–61. doi:10.1016/S1364-6613(97)01008-5

Roseboom, W., Nishida, S., & Arnold, D. H. (2009). The sliding window of audio–visual simultaneity. *Journal of Vision*, *9(12):4*, 1–8. doi:10.1167/9.12.4

Rutschmann, J., & Link, R. (1964). Perception of temporal order of stimuli differing in sense mode and simple reaction time. *Perceptual and Motor Skills*, *18*(2), 345–352. doi:10.2466/pms.1964.18.2.345

Rutschmann, R. (1966). Perception of temporal order and relative visual latency. *Science*, *152*(3725), 1099–1101. doi:10.1126/science.152.3725.1099

Sanders, M. C., Chang, N.-Y. N., Hiss, M. M., Uchanski, R. M., & Hullar, T. E. (2011). Temporal binding of auditory and rotational stimuli. *Experimental Brain Research*, *210*(3-4), 539–547. doi:10.1007/s00221-011-2554-x

Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive Psychology*, *47*(4), 333–366. doi:10.1016/S0010-0285(03)00035-5

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychological Science*, *12*(3), 205–212. doi:10.1111/1467-9280.00337

Smeele, P. M. T. (1994). *Perceiving speech: integrating auditory and visual speech* (Unpublished doctoral dissertation). Delft University of Technology, Delft. Retrieved from http://repository.tudelft.nl/view/ir/uuid:218c06bd-dde7-4f0d-bb6a-22c090289508/

Spence, C., Baddeley, R., Zampini, M., James, R., & Shore, D. I. (2003). Multisensory temporal order judgments: when two locations are better than one. *Perception & Psychophysics*, *65*(2), 318–28.

Spence, C., & Parise, C. (2010). Prior-entry: a review. *Consciousness and Cognition*, *19*(1), 364–79. doi:10.1016/j.concog.2009.12.001

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*(4), 799–832. doi:10.1037/0096-3445.130.4.799

Spence, C., & Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. *Current Biology*, *13*(13), R519–R521. doi:10.1016/S0960-9822(03)00445-7

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: Fundamental issues and a general model. In S. Kornblum (Ed.), *Attention and Performance IV* (pp. 629–685). New York: Academic Press.

Stetson, C., Cui, X., Montague, P. R., & Eagleman, D. M. (2006). Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron*, *51*, 651–659. doi:10.1016/j.neuron.2006.08.006

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 of the Royal Society B: Biological Sciences*, *268*(1462), 31–38. doi:10.1098/rspb.2000.1326

Teatini, G., Farne, M., Verzella, F., & Berruecos, P. jr. (1976). Perception of temporal order: Visual and auditory stimuli. *Giornale Italiano Di Psicologia*, *3*, 157–164.

Van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Attention, Perception, & Psychophysics*, *70*(6), 955–968. doi:10.3758/PP.70.6.955

Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2008). Audiovisual temporal adaptation of speech: Temporal order versus simultaneity judgments. *Experimental Brain Research*, *185*(3), 521–529. doi:10.1007/s00221-007-1168-9

Von Békésy, G. (1963). Interaction of paired sensory stimuli and conduction in peripheral nerves. *Journal of Applied Physiology*, *18*, 1276–1284.

Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research*, *22*(1), 32–35. doi:10.1016/j.cogbrainres.2004.07.003

Yarrow, K., Jahn, N., Durant, S., & Arnold, D. H. (2011). Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Consciousness and Cognition*, *20*(4), 1518–1531. doi:10.1016/j.concog.2011.07.003

Zampini, M. (2003). Multisensory temporal order judgments: The role of hemispheric redundancy. *International Journal of Psychophysiology*, *50*(1-2), 165–180. doi:10.1016/S0167-8760(03)00132-6

Zampini, M., Brown, T., Shore, D. I., Maravita, A., Röder, B., & Spence, C. (2005). Audiotactile temporal order judgments. *Acta Psychologica*, *118*(3), 277–91. doi:10.3758/BF03193329

Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audio-visual simultaneity judgments. *Perception & Psychophysics*, *67*(3), 531–544. doi:10.3758/BF03193329

Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental Brain Research*, *152*(2), 198–210. doi:10.1007/s00221-003-1536-z

Zampini, M., Shore, D. I., & Spence, C. (2005). Audiovisual prior entry. *Neuroscience Letters*, *381*(3), 217–222. doi:10.1016/j.neulet.2005.01.085

Zychaluk, K., & Foster, D. (2009). Model-free estimation of the psychometric function. *Attention, Perception, & Psychophysics*, *71*(6), 1414–1425. doi:10.3758/APP

Table 1

*Mean PSS and JND values in milliseconds for each modality-pairing as well as mean distance from the transitivity plane, with corresponding standard errors of the mean (SEM) in brackets, are shown for both temporal order judgments (TOJ) and synchrony judgments (SJ).*

| **Task type** | **Modality-pairings** | | | | | | **Transitivity distance** |
|---|---|---|---|---|---|---|---|
| | **Audiovisual** | | **Audiotactile** | | **Visuotactile** | | |
| | PSS (SEM) | JND (SEM) | PSS (SEM) | JND (SEM) | PSS (SEM) | JND (SEM) | Mean (SEM) |
| **TOJ** | −30 (23) | 75 (9) | −50 (12) | 66 (10) | −46 (16) | 55 (6) | −38 (6) |
| **SJ** | −4 (12) | 81 (7) | −28 (4) | 56 (9) | 10 (10) | 65 (6) | −8 (6) |

Table 2

*Test statistics and corresponding p-values for the tests described in the 'Results' section.*

| | **Modality-pairings** | | |
|---|---|---|---|
| **Test type** | **AV** | **AT** | **TV** |
| a. One-tailed t-test of the TOJ PSS against 0 | t(10) = −1.29 p = 0.23 | t(10) = −4.31, p = 0.00# | t(10) = −2.87, p = 0.017# |
| b. One-tailed t-test of the SJ PSS against 0 | t(10) = −0.35, p = 0.73 | t(10) = −7.29, p = 0.00# | t(10) = 0.98, p = 0.35 |
| **Comparison of TOJ and SJ** | | | |
| c. Correlation of JNDs | r(9) = 0.46, p = 0.16 | r(9) = 0.54, p = 0.09 | r(9) = 0.6, p = 0.05 |
| d. Correlation of PSS | r(9) = 0.73, p = 0.01 | r(9) = 0.75, p = 0.01* | r(9) = 0.4, p = 0.23 |
| e. Two-tailed paired-sample t-test of the PSS | t(10) = 1.55, p = 0.15 | t(10) = 2.36, p = 0.08# | t(10) = 3.68, p = 0.00# |
| f. Two-tailed paired-sample t-test of the JND | t(10) = −0.64, p = 0.54 | t(10) = 1.1, p = 0.3 | t(10) = −1.99, p = 0.07 |

# a Bonferroni correction has been applied to p-values below 0.1 to account for the two t-tests performed on each dataset;
* the correlation is no longer significant after one influential data point is removed from the data set (r(8) = 0.5, p = 0.14)

Table 3

*All possible combinations of processing coherence/incoherence at two processing stages and the resulting transitivity/intransitivity of data.*

| **Coherence in** | | **Transitivity** |
|---|---|---|
| **Sensory Processes** | **Decision Processes** | |
| Yes | Yes | Yes |
| Yes | No | No |
| No | Yes | No |
| No | No | No* |

*Note.* The transitivity of PSS depends on the processing coherence between the modality-pairings at two processing stages: the sensory stage and the decision stage. The table shows all possible combinations of processing coherence at each stage and their individual predictions for data (in)transitivity.

*In principle, transitive data can result. However, we deem this possibility unlikely (see text for detailed explanation).

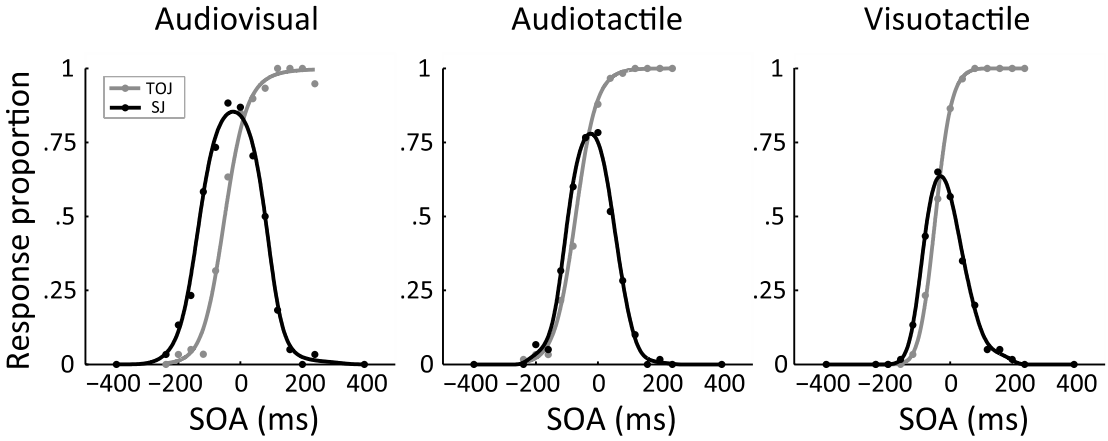


*Figure 1.* TOJ and SJ data (grey and black circles, respectively) as well as the fitted psychometric functions (continuous lines) for one participant. For TOJ, data is plotted as the proportion of trials in which the auditory stimulus was reported to have occurred before the visual or tactile stimulus (for audiovisual and audiotactile pairings, respectively), or in which the tactile stimulus was reported to have occurred before the visual (for visuotactile pairings), as a function of the SOA. For SJ, data is plotted as the proportion of trials in which the two stimuli were judged as synchronous, as a function of SOA. Positive SOAs indicate a lead of the auditory stimulus for audiovisual and audiotactile pairings, and a lead of the tactile stimulus for visuotactile pairings.

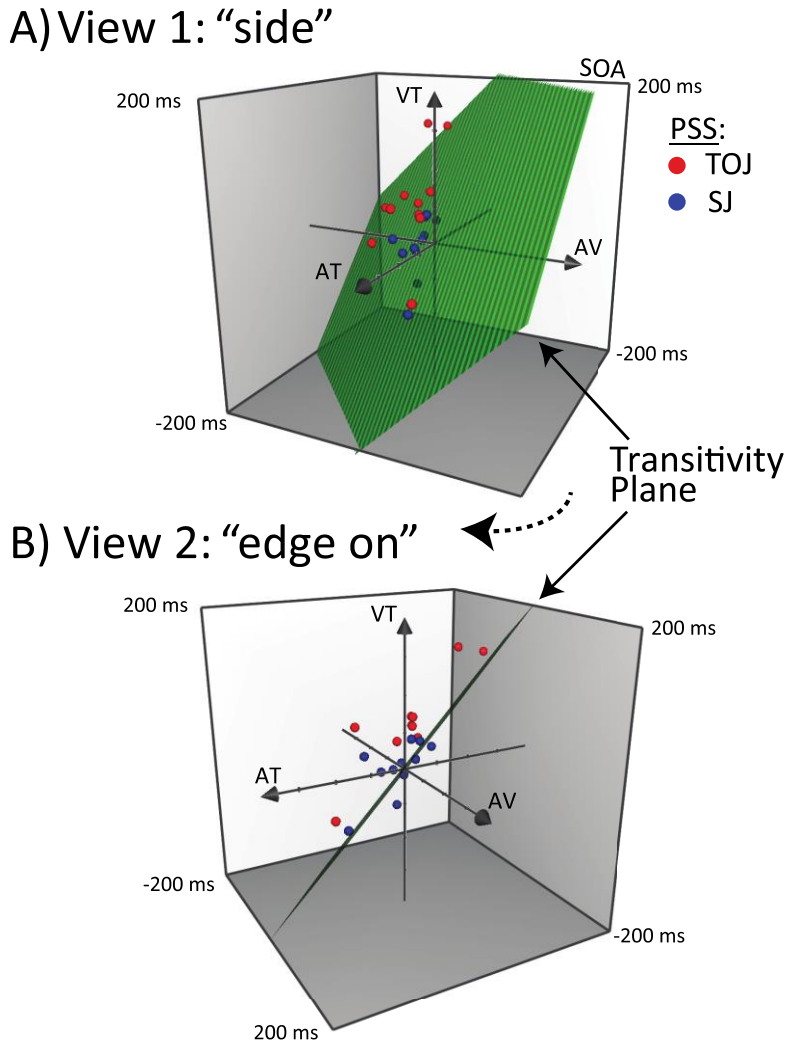


*Figure 2.* Each participant’s PSS data in relation to the transitivity plane. Each participants’ three PSS estimates (for the AV, AT, and TV modality-pairings) can be represented as a point in a three-dimensional coordinate system, which axes are defined by the SOAs of the three modality-pairings. If the PSS data were transitive, all data points should fall onto a plane (*transitivity plane*). (A) The transitivity plane is shown as a shaded area. (B) The same representation as (A) rotated such that the transitivity plane is viewed ‘edge on’.

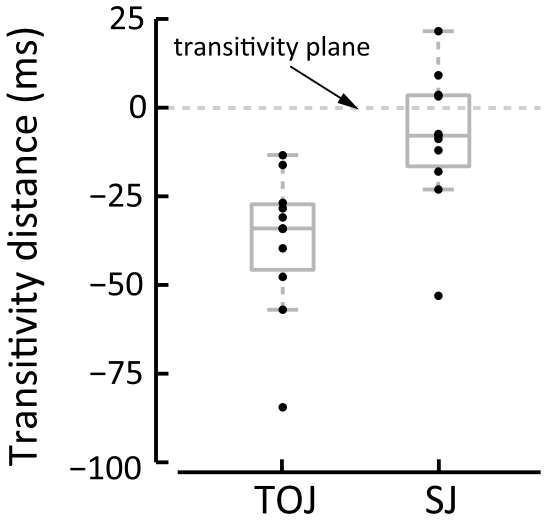


*Figure 3.* Participants' distances from the transitivity plane are shown as single data points for both TOJ and SJ. Overlaid boxplots indicate the distribution mean, interquartile range, and data range (excluding outliers).

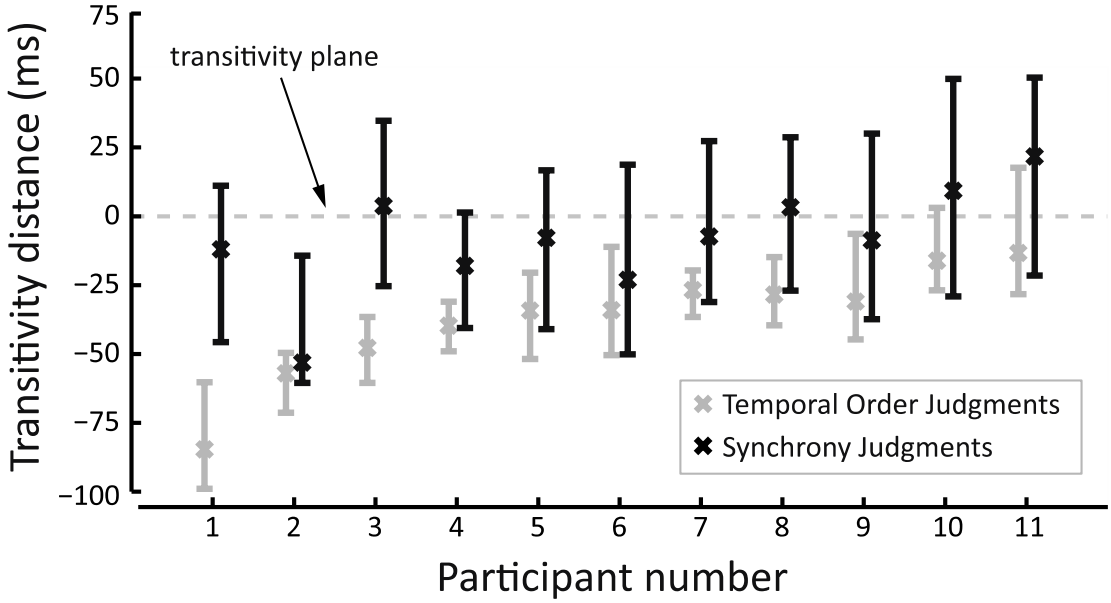


*Figure 4.* Each participant's distance from the transitivity plane for both TOJ and SJ, with 95% confidence intervals that were estimated by bootstrapping.

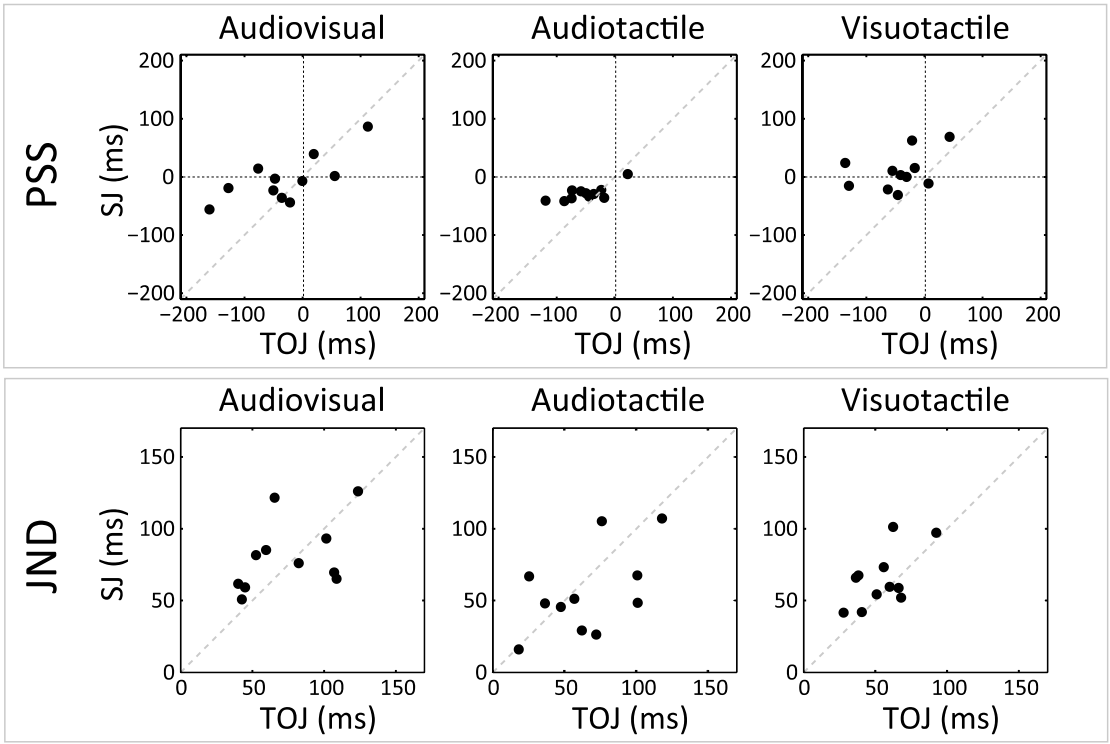


*Figure 5.* SJ PSS and JND estimates (upper and lower panel, respectively) of each participant plotted against the corresponding TOJ estimates, for each of the three modality-pairings. The diagonal lines indicate the identity of TOJ and SJ estimates. The correlation coefficient *r* between estimates is reported in Table 2.