

Converging evidence that common timing processes underlie temporal-order and simultaneity judgments: a model-based analysis

Miguel A. García-Pérez · Rocío Alcalá-Quintana

Published online: 27 March 2015
© The Psychonomic Society, Inc. 2015

Abstract Perception of simultaneity and temporal order is studied with simultaneity judgment (SJ) and temporal-order judgment (TOJ) tasks. In the former, observers report whether presentation of two stimuli was subjectively simultaneous; in the latter, they report which stimulus was subjectively presented first. SJ and TOJ tasks typically give discrepant results, which has prompted the view that performance is mediated by different processes in each task. We looked at these discrepancies from a model that yields psychometric functions whose parameters characterize the timing, decisional, and response processes involved in SJ and TOJ tasks. We analyzed 12 data sets from published studies in which both tasks had been used in within-subjects designs, all of which had reported differences in performance across tasks. Fitting the model jointly to data from both tasks, we tested the hypothesis that common timing processes sustain simultaneity and temporal-order judgments, with differences in performance arising from task-dependent decisional and response processes. The results supported this hypothesis, also showing that model psychometric functions account for aspects of SJ and TOJ data that classical analyses overlook. Implications for research on perception of simultaneity and temporal order are discussed.

Keywords Simultaneity · Temporal order · Timing processes · Decisional processes

Electronic supplementary material The online version of this article (doi:10.3758/s13414-015-0869-6) contains supplementary material, which is available to authorized users.

M. A. García-Pérez (✉) · R. Alcalá-Quintana
Departamento de Metodología, Facultad de Psicología, Universidad Complutense, Campus de Somosaguas, 28223 Madrid, Spain
e-mail: miguel@psi.ucm.es

The timing processes sustaining perception of simultaneity and temporal order are respectively inferred from observers' performance in simultaneity judgment (SJ) and temporal-order judgment (TOJ) tasks. In both cases, each trial presents a pair of stimuli whose onsets (or offsets) occur at arbitrary times, resulting in a temporal delay regarded as positive when the stimulus designated "test" lags the stimulus designated "reference" and negative when the test stimulus leads the reference stimulus. In SJ tasks, observers report whether presentation (or extinction) of the two stimuli was subjectively simultaneous. Because observers directly report perception of simultaneity, the temporal delay at which simultaneous responses are maximally prevalent is interpreted as the point of subjective simultaneity (PSS). In contrast, in TOJ tasks observers report which stimulus was subjectively presented (or extinguished) first (or second), with no option to report subjective simultaneity. The temporal delay at which observers give equal numbers of "test first" and "reference first" (alternatively, "test second" and "reference second") responses is taken as the PSS, but this is only an indirect measure because the TOJ task avoids collecting evidence of perceived simultaneity. Strictly speaking, such analysis of TOJ data only identifies the temporal delay at which "test first" and "reference first" responses are equally prevalent, which is synonymous with the PSS only under the strong additional assumption that observers give "test first" and "reference first" responses equally often when they perceive simultaneity. Unfortunately, the validity of such assumption cannot be tested with TOJ data for lack of evidence as to when observers perceived simultaneity and, then, guessed a response.

PSS estimates obtained with SJ and TOJ tasks have systematically been reported to differ in within-subjects studies using both tasks with the same stimuli and conditions (Barnett-Cowan & Harris, 2009, 2011; Capa, Duval, Blaison, & Giersch, 2014; Donohue, Woldorff, & Mitroff,

2010; Fujisaki & Nishida, 2009; Li & Cai, 2014; Linares & Holcombe, 2014; Love, Petrini, Cheng, & Pollick, 2013; Schneider & Bavelier, 2003; Stevenson & Wallace, 2013; van Eijk, Kohlrausch, Juola, & van de Par, 2008; van Eijk et al. 2010; Sanders, Chang, Hiss, Uchanski, & Hullar, 2011; Vatakis, Navarra, Soto-Faraco, & Spence, 2008). This has prompted the notion that simultaneity and temporal-order judgments are governed by different processes (Spence & Parise, 2010; Vroomen & Keetels, 2010). It should be noted that the PSS estimates supporting this conclusion are obtained by fitting arbitrary psychometric functions to the data, so that the SJ estimate of the PSS is the abscissa at which the fitted function peaks and the TOJ estimate is the abscissa at which the fitted function evaluates to 0.5. The fitted functions accommodate the path of the data to obtain these estimates but without links to the timing, decisional, and response processes that presumably produced the data. The absence of these links only permits a description of observed performance, precluding an evaluation of the underlying processes.

This paper looks at performance in SJ and TOJ tasks under an independent-channels model that makes explicit assumptions about the operation of timing, decisional, and response processes (García-Pérez & Alcalá-Quintana, 2012a). The resultant model-based psychometric functions thus include parameters relating to each of these processes and permit a direct test of differences in each of them across SJ and TOJ tasks. In search for these differences, the model was used to analyze 12 data sets from published studies in which both tasks defined a within-subjects factor. The analysis of data from three of the studies will be described in full detail. In the first one, Linares and Holcombe (2014) reported different PSS estimates from SJ and TOJ tasks with audiovisual pairs and concluded that perceptual latencies differ across tasks. In the second, Capa et al. (2014) had schizophrenia patients and normal controls perform SJ and TOJ tasks with visuo-visual pairs and compared PSS estimates across tasks as well as error rates in the TOJ task with error rates predicted from performance in the SJ task, concluding that PSSs as well as actual vs. predicted error rates differ, further showing that patients display a selective impairment at long temporal delays only in TOJ tasks. In the third study, Li and Cai (2014) reported different PSS estimates from SJ and TOJ tasks with visuo-visual pairs involving digits of different numerical magnitudes, concluding that under both tasks large numbers are perceived later than small numbers. Analysis of the remaining data sets is succinctly described in the [Electronic Supplementary Material](#) and involved visuo-visual pairs under a diversity of attentional cuing conditions (Schneider & Bavelier, 2003): audio-visual, visuo-tactile, audio-tactile, and tactile-tactile pairs (Fujisaki & Nishida, 2009), and audio-vestibular pairs (Sanders et al., 2011).

All analyses corroborate that discrepancies in performance across SJ and TOJ tasks result from task-dependent decisional

and response processes operating on the outcome of timing processes that are identical under both tasks. The next section briefly describes the model used in our analyses, including a variant that holds under the conditions of some of the data sets. Subsequent sections describe relevant details of the experiments whose data are analyzed, specific details of the way in which the model was fitted to each data set, and the results. The *Discussion* provides a final overview and also comments on some methodological issues arising in research on perception of temporal order and simultaneity.

Independent-channels model

The independent-channels (IC) model was presented in full detail in García-Pérez and Alcalá-Quintana (2012a) but a brief description will be given here. Before doing so, a broad overview of its components and assumptions is worth giving in plain language. In IC models, sensory signals from each stimulus component (e.g., auditory and visual) are independently processed through the corresponding sensory pathway to render a *perceived onset time* for each component. These perceived onsets vary across trials according to some distribution. A *timing judgment* (e.g., audio first, video first, or simultaneous) then results by application of a *decision rule* to the perceived onset of each component. Response processes subsequently map this judgment onto an observable response (among the options given by the task), although the observer may make *response errors*.

The model is formally presented next in terms adapted to the experimental conditions of Linares and Holcombe (2014), where the auditory (instead of the visual) component was the reference. For all practical purposes, this only implies that subscripts “a” and “v” (denoting the auditory and visual components) are swapped in model equations with respect to equations when the visual component is the reference. A variant of the model will subsequently be presented that applies to the conditions of Capa et al. (2014) and Li and Cai (2014), in which temporal delays were delivered via pairs of visual stimuli presented side by side on a monitor.

The *perceived onsets* (or perceived latencies) T_v and T_a of the visual and auditory components of a stimulus are random variables with densities g_v and g_a given by the shifted exponential distributions

$$g_i(t) = \lambda_i \exp[-\lambda_i(t - (\Delta t_i + \tau_i))], \quad t \geq \Delta t_i + \tau_i, \quad i \in \{v, a\} \quad (1)$$

where Δt_i is the physical onset of component i and λ_i and τ_i reflect peripheral processing and neural transmission times. If the auditory component is the reference, $\Delta t_a = 0$ by definition and $\Delta t \equiv \Delta t_v$ is the visual delay with which the pair is

presented, where $\Delta t < 0$ ($\Delta t > 0$) reflects that the visual onset precedes (lags) the auditory onset. Figure 1a shows sample distributions when $\Delta t = 0$ (i.e., no delay), reflecting the variability with which visual and auditory onsets are perceived across trials.

On a given trial, perceived onsets are drawn from these distributions and judgments arise from a decision rule applied to the perceived-onset difference $D = T_v - T_a$, which has the asymmetric Laplace distribution

$$f(d; \Delta t) = \begin{cases} \frac{\lambda_a \lambda_v}{\lambda_a + \lambda_v} \exp[\lambda_a(d - \Delta t - \tau)] & \text{if } d \leq \Delta t + \tau \\ \frac{\lambda_a \lambda_v}{\lambda_a + \lambda_v} \exp[-\lambda_v(d - \Delta t - \tau)] & \text{if } d > \Delta t + \tau \end{cases} \tag{2}$$

where $\tau = \tau_v - \tau_a$. Figure 1b shows the distribution of D for the case in Fig. 1a. A resolution parameter δ (Fig. 1b) limits the observer’s ability to tell small differences in perceived onset. Then, “video-first” (VF) judgments arise when D is large and negative ($D < -\delta$), “audio-first” (AF) judgments arise when D is large and positive ($D > \delta$), and “simultaneous” (S) judgments arise when D is below the resolution limit ($-\delta \leq D \leq \delta$). The probability of each judgment varies with Δt , as this shifts the distribution of perceived-onset differences. Figure 1c shows the functions p_{VF} , p_S , and p_{AF} describing how the probability of each judgment varies with Δt . These functions are given by

$$p_{VF}(\Delta t) = F(-\delta; \Delta t) \tag{3}$$

$$p_S(\Delta t) = F(\delta; \Delta t) - F(-\delta; \Delta t) \tag{4}$$

$$p_{AF}(\Delta t) = 1 - F(\delta; \Delta t) \tag{5}$$

with

$$F(d; \Delta t) = \int_{-\infty}^d f(z; \Delta t) dz = \begin{cases} \frac{\lambda_a}{\lambda_a + \lambda_v} \exp[\lambda_a(d - \Delta t - \tau)] & \text{if } d \leq \Delta t + \tau \\ 1 - \frac{\lambda_v}{\lambda_a + \lambda_v} \exp[-\lambda_v(d - \Delta t - \tau)] & \text{if } d > \Delta t + \tau \end{cases} \tag{6}$$

Under the model, these judgments precede and are independent of the question that observers answer at the end of the trial (i.e., the task that is administered). Also, the distribution of D in Eq. (2) must be invariant across SJ and TOJ tasks for any given observer when stimuli and conditions are identical. Nevertheless, the observed psychometric functions in SJ and TOJ tasks will differ because they express the probabilities in Eqs (3)–(5) differently, as discussed next.

In SJ tasks, AF and VF judgments are aggregated into nonsimultaneous (NS) judgments and the psychometric function for S responses in SJ tasks is

$$\Psi_{SJ}(\Delta t) = p_S(\Delta t) \tag{7}$$

(Fig. 2a). In TOJ tasks, however, observers must give VF or AF responses even when they judge simultaneity. This calls for a guessing parameter ξ , describing the probability with which an observer gives an AF response in these cases. The psychometric function for AF responses in TOJ tasks is, then,

$$\Psi_{TOJ}(\Delta t) = \xi p_S(\Delta t) + p_{AF}(\Delta t) \tag{8}$$

which may seem inconsistent with Ψ_{SJ} depending on the observer’s response bias (Fig. 2b).

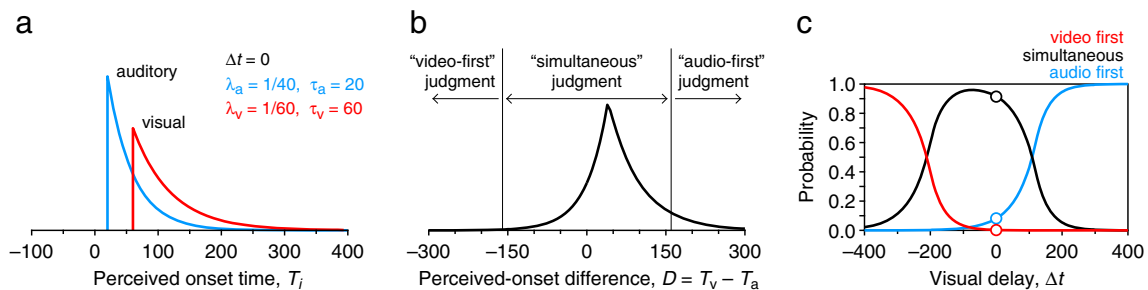


Fig. 1 IC model of timing judgments. (a) Distributions of perceived visual onset (red curve) and perceived auditory onset (blue curve) for simultaneously presented stimuli (i.e., with a visual delay $\Delta t = 0$ ms). The parameters of each distribution are indicated in the inset. (b) Distribution of perceived-onset differences (curve) and boundaries in decision space (vertical lines at $D = \pm\delta$ with $\delta = 160$), determining the

probability of each type of judgment. The distribution of perceived-onset differences is asymmetric unless $\lambda_a = \lambda_v$ and peaks at $D = \Delta t + \tau$. (c) Curves describing the probability of each type of judgment as a function of visual delay Δt . Circles denote the probabilities when $\Delta t = 0$ ms, coming from the partition illustrated in (b)

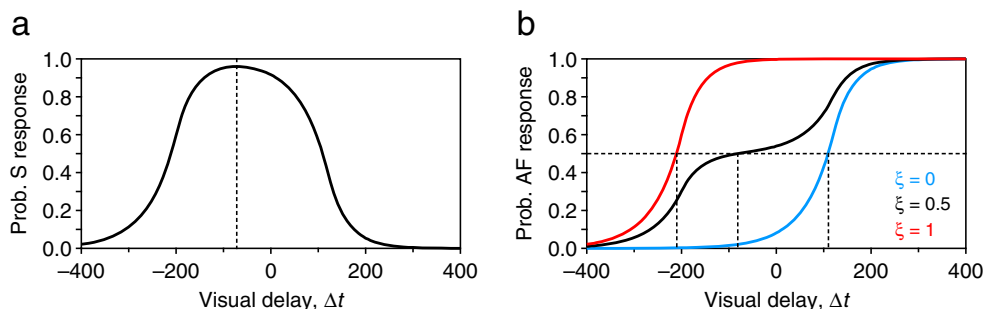


Fig. 2 Psychometric functions in SJ (a) and TOJ (b) tasks. In the absence of response errors, the psychometric function in SJ tasks equals the curve describing the probability of “simultaneous” judgments illustrated in Fig. 1c. In TOJ tasks, and also in the absence of response errors, the form of the psychometric function depends on how observers transform judgments of simultaneity into audio-first (AF) or video-first (VF) responses, which is determined by the response bias parameter ξ . When $\xi = 1$ (red curve), judgments of simultaneity are always reported as AF

The IC model just presented must be amended to cover realistically the mapping of judgments onto responses, a process during which errors can be made (see García-Pérez & Alcalá-Quintana, 2012a, 2012b). With this extension, the final form of the psychometric functions in SJ and TOJ tasks are

$$\Psi_{SJ}(\Delta t) = \epsilon_{VF}^{(SJ)} p_{VF}(\Delta t) + (1 - \epsilon_S^{(SJ)}) p_S(\Delta t) + \epsilon_{AF}^{(SJ)} p_{AF}(\Delta t) \quad (9)$$

$$\Psi_{TOJ}(\Delta t) = \epsilon_{VF}^{(TOJ)} p_{VF}(\Delta t) + \xi p_S(\Delta t) + (1 - \epsilon_{AF}^{(TOJ)}) p_{AF}(\Delta t) \quad (10)$$

where each ϵ reflects the probability of misreporting the judgment indicated by its subscript under the task indicated by its parenthetical superscript. Note that Eqs. (7) and (8) obtain when all ϵ s are zero (i.e., when response errors are never made).

It is useful to consider at this point how IC model parameters affect the (theoretical) shapes of Ψ_{SJ} and Ψ_{TOJ} , although we will mostly discuss these effects with respect to Ψ_{SJ} because the same effects just carry over to Ψ_{TOJ} . Parameters λ_a and λ_v come from the distributions of perceived onsets (Fig. 1a), making Ψ_{SJ} bilaterally asymmetric about its peak when $\lambda_a \neq \lambda_v$ and being separately responsible for the different drop-off rates of Ψ_{SJ} on either side (Fig. 2a). Parameter τ is responsible for the location of the peak of the distribution of perceived-onset differences (Fig. 1b) which, in turn, determines the location of the peak of Ψ_{SJ} (Fig. 2a). Parameter δ determines the width of the central region in decision space (Fig. 1b), which has direct effects on the height of Ψ_{SJ} at its peak and also on its overall breadth: Ψ_{SJ} is short and narrow when δ is small, whereas it is tall and broad when δ is large (not illustrated in the preceding figures). Because these

parameters produce different effects on Ψ_{SJ} , they are not confounded and, thus, can be recovered (within sampling error) from empirical SJ data showing various degrees of asymmetry, lateral shifts, or breadths and heights. Some examples of the empirical diversity of these characteristics will be seen in our forthcoming analyses of data. The effects of the additional response bias parameter ξ , unique to the TOJ task can be appreciated in the three sample psychometric functions in Fig. 2b, which also incorporate the effects of the parameters discussed above. Finally, the ϵ parameters only affect the asymptotic behavior of either type of psychometric function at large (positive or negative) delays, although the effect can be as large as shown in some cases (e.g., SJ and TOJ data and fitted functions for observer DL at large negative delays in Fig. 4a below, or analogous effects at large delays of both signs for observer P07 in Fig. 6a below).

The IC model thus describes performance in SJ and TOJ tasks with timing parameters (λ_a , λ_v , and τ) that have common values as long as stimuli and conditions are identical for both tasks and with decisional and response parameters (δ , ξ , and the various ϵ s) that may vary across tasks. These features permit direct tests of the model in within-subjects conditions, checking out the feasibility of accounting for the data under the assumption of identical timing processes across tasks. It should nevertheless be noted that timing parameters may vary greatly across sensory modalities (i.e., they will reasonably vary for visual, auditory, tactile, or vestibular stimuli due to specificities of peripheral processing and neural transmission along the corresponding pathways), across stimuli within the same sensory modality (e.g., due to the place–frequency map of the basilar membrane, low and high auditory frequencies reach their associated locations at different times; Uppenkamp, Fobel, & Patterson, 2001; Wojtczak, Beim, Michey, & Oxenham, 2012, Wojtczak et al. 2013) or across conditions involving the same stimuli (e.g., when different types of attentional cue are used in separate experimental

conditions; see our analysis of Scheinder and Bavelier’s, 2003, data in the [Electronic Supplementary Material](#)). What the IC model posits is that, all else being equal, timing parameters do not vary across SJ and TOJ tasks.

The model simplifies when temporal delays are delivered via identical stimuli as in the experiment of Capa et al. (2014), under the reasonable assumption that timing processes do not operate differently when visual stimuli are presented on the left or on the right visual field. Even though both stimuli are identical, it is still necessary to designate one of them as the reference to differentiate negative from positive delays. Without loss of generality, we will regard the stimulus presented on the left side of the monitor as the reference so that subscripts “a” and “v” in the previous equations are respectively replaced with “l” and “r.” Nevertheless, because the two stimuli are identical, perceived onsets must be identically distributed for both of them, yielding $\lambda_l = \lambda_r = \lambda$ and $\tau = \tau_r - \tau_l = 0$. Carrying over these substitutions, Eq. (2) reduces to the Laplace distribution

$$f(d; \Delta t) = \frac{\lambda}{2} \exp[-\lambda|d - \Delta t|] \tag{11}$$

and Eq. (6) becomes

$$F(d; \Delta t) = \begin{cases} \frac{1}{2} \exp[\lambda(d - \Delta t)] & \text{if } d \leq \Delta t \\ 1 - \frac{1}{2} \exp[-\lambda(d - \Delta t)] & \text{if } d > \Delta t \end{cases} \tag{12}$$

A further consequence of using pairs of identical stimuli is that the resultant removal of parameter τ permits relaxing the assumption of a symmetric central region in decision space. Amending the model to allow for non-symmetric boundaries respectively placed at δ_1 and δ_2 (with $\delta_1 \neq -\delta_2$) incorporates the notion of decisional bias towards one of the response options (Fig. 3). A symmetric central region (and, hence, the

assumption of lack of decisional bias) is needed in the general version of the model because SJ and TOJ tasks are instances of the method of single stimuli (MSS) and share its limitations. Specifically, and with reference to the diagram in Fig. 1b, a rigid displacement of the decision boundaries (by shifting the boundary currently at $-\delta$ to $\delta_1 = -\delta + k$ and similarly shifting the boundary at δ to $\delta_2 = \delta + k$, with positive or negative k) produces the same psychometric functions that would result by leaving the boundaries at $-\delta$ and δ and replacing τ with $\tau + k$. This non-identifiability is not a property of this model but a general shortcoming of MSS (García-Pérez & Alcalá-Quintana, 2013; Schneider & Bavelier, 2003; Yarrow, Jahn, Durant, & Arnold, 2011). Yet, under the conditions discussed now, where $\tau = 0$, this non-identifiability is removed and potentially asymmetric boundaries at δ_1 and δ_2 can be estimated. Then, in place of Eqs. (3)–(5) above, the probabilities of “right-first” (RF), “simultaneous” (S), and “left-first” (LF) judgments are given by

$$p_{RF}(\Delta t) = F(\delta_1; \Delta t) \tag{13}$$

$$p_S(\Delta t) = F(\delta_2; \Delta t) - F(\delta_1; \Delta t) \tag{14}$$

$$p_{LF}(\Delta t) = 1 - F(\delta_2; \Delta t) \tag{15}$$

which are subsequently combined with appropriately subscripted error parameters as in Eqs. (9) and (10) to render the psychometric functions in SJ and TOJ tasks under these conditions. Thus, the IC model again describes performance in SJ and TOJ tasks involving identical stimuli and conditions with common timing parameters (now only λ) and with decisional and response parameters (δ_1 , δ_2 , ξ , and the various ϵ s) that may vary across tasks.

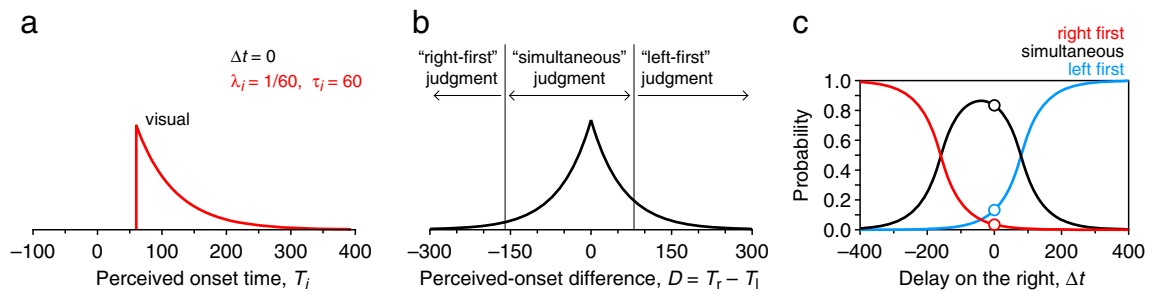


Fig. 3 IC model of timing judgments when temporal delays are delivered via identical visual stimuli presented on the left and on the right of a fixation point on a monitor. **(a)** Distribution of perceived onset for the only stimulus implied in this case. **(b)** Distribution of perceived-onset differences (curve) and potentially asymmetric boundaries in decision space (vertical lines at $D = \delta_1 = -160$ and $D = \delta_2 = 80$). Because perceived onsets are identically distributed for both stimuli, the

distribution of perceived-onset differences is symmetric and peaks at $D = \Delta t$. **(c)** Curves describing the probability of each type of judgment as a function of the temporal delay Δt with which the stimulus on the right is presented. Circles denote the probabilities when $\Delta t = 0$ ms, coming from the partition illustrated in **(b)**. Because the decision boundaries δ_1 and δ_2 are not symmetrically placed around $D = 0$, the vertical axis of bilateral symmetry of the ensemble of curves is at $\Delta t = (\delta_1 + \delta_2)/2 = -40$ ms

Data and model fit

Linares and Holcombe (2014)

In the study by Linares and Holcombe (2014), seven observers performed SJ and TOJ tasks with the same stimuli and conditions. The (reference) auditory stimulus was a 10-ms white noise burst; the visual (test) stimulus was a 10-ms change from black to white and back to black in the luminance of a circle displayed on a grey background. Test onset times were manipulated to occur with a delay Δt ranging from -250 ms to 250 ms in steps of 50 ms and 20 trials were given at each delay.

We downloaded the data made available by the authors from <http://www.dlinares.org> and fitted the IC model for each observer jointly across tasks using a FORTRAN version of the software described in Alcalá-Quintana and García-Pérez (2013). A joint fit implies that the timing parameters λ_a , λ_v , and τ are estimated to have common values for both tasks, whereas decisional and response parameters (δ , ξ , and the ϵ s) vary across tasks. If the data cannot be properly accounted for on the implied assumption that timing parameters have common values in both tasks, the fit will be poor and fitted curves will not follow the path of the data. Non-null values for the ϵ s were not pre-assumed; rather, the software option model = “best” (Alcalá-Quintana & García-Pérez, 2013) was used to determine which ϵ s ought to be non-null. The joint fit involves somewhere between 6 parameters (when all ϵ s are null) and 11 parameters (when all ϵ s are non-null) per observer. For comparison, we also fitted the data using the approach of Linares and Holcombe (2014), which separately fits a four-parameter cumulative Gaussian to TOJ data and a three-parameter scaled Gaussian to SJ data. This yields seven parameters (two of which are error rates) per observer. The interpretation of those parameters in terms of underlying timing processes is impossible, because the descriptive approach only produces curves whose location and slope best delineate the path of the data. For both fitting approaches, we computed the likelihood-ratio statistic G^2 as a measure of goodness of fit. The outcomes of the two approaches were compared with the Bayesian Information Criterion (BIC), which assesses economy of description for each approach by combining the goodness-of-fit measure and the number of parameters used to attain that fit.

Capa et al. (2014)

In the study by Capa et al. (2014), 20 schizophrenia patients and 20 normal controls performed SJ and TOJ tasks with identical stimuli and conditions. The reference stimulus for our analysis was a gray rectangle presented on the left of a central fixation point on the monitor; the test stimulus was an identical rectangle presented on the right of the fixation point

with a temporal delay Δt that ranged from -96 ms to 96 ms in steps of 24 ms. Offset times did not differ, as both stimuli remained on display until observers responded. At each non-null temporal delay, 20 TOJ trials and 54 SJ trials were given; the number of trials was doubled at the null temporal delay. Some observers missed some trials in the SJ task, so that the effective number of SJ trials was usually smaller and also varied across temporal delays. In the TOJ task, observers reported which stimulus was presented second (rather than first), which is largely inconsequential for our analyses: Judgments LF and RF referred to in Eqs. (13)–(15) straightforwardly translate into “right second” (RS) and “left second” (LS), respectively.

We obtained raw data from Anne Giersch and fitted the adapted version of the IC model also jointly across tasks. The same software was used for this purpose, once modified to implement the constraints and extensions discussed above. The joint fit again implies that the timing parameter λ is estimated to have the same value in both tasks, whereas decisional and response parameters (δ_1 , δ_2 , ξ , and the ϵ s) vary across tasks. Evidence that timing processes are identical in both tasks will again show in the form of a good fit (i.e., nonsignificant G^2) and fitted curves that follow the path of the data from each task. Decisions about which of the ϵ s ought to be non-null were made as with the previous data set. In this case, the joint fit also involves somewhere between 6 parameters (when all ϵ s are null) and 11 parameters (when all are non-null) per observer. Descriptive psychometric functions were not fitted to the data because the analyses reported by Capa et al. (2014) were not based on fitted functions.

Li and Cai (2014)

In the study of Li and Cai (2014), 16 observers carried out SJ and TOJ tasks with pairs of visual stimuli flashed for 50 ms on the left and on the right of a fixation point on a monitor under identical conditions. Stimuli were black digits presented on a uniform gray background and the reference versus test distinction was based on the numerical magnitude of the digits. The study investigated whether the numerical magnitude of a digit affects the speed with which it is processed. Small digits were “1” or “2” and large digits were “8” or “9.” A large and a small digit were selected at random for presentation in each trial and, in turn, the positions (left and right of the fixation point) of the two digits also were decided at random in each trial. The small digit was regarded as the reference irrespective of the position at which it was displayed and the large (test) digit was presented at the other position with a temporal delay Δt of 0 , ± 10 , ± 20 , ± 30 , or ± 50 ms. Data were aggregated across presentation positions and choice of digit, yielding 48 trials at each temporal delay under each task. Observers kept their gaze at the fixation point throughout the trial and reported whether or not the two presentations were subjectively simultaneous (SJ task) or the

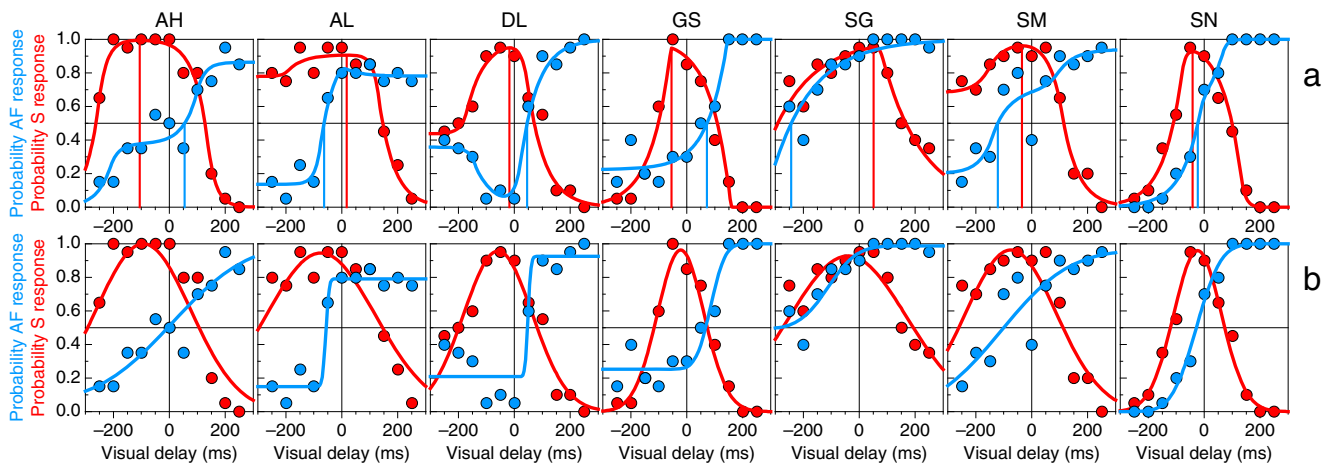


Fig. 4 Data and fitted psychometric functions for the study of Linares and Holcombe (2014) under the IC model (a) and under a descriptive approach (b). Results for the SJ task are shown in red; results for the TOJ

task are shown in blue. Each column pertains to the observer indicated at the top. Estimated PSSs for each observer from the SJ and TOJ tasks are shown by vertical lines of the appropriate color in the upper panels

side on which a digit had subjectively appeared first (TOJ task; with “left first” and “right first” responses translated accordingly into “small number first” or “large number first”).

We fitted the IC model to raw data provided by Shuang-Xia Li, using the simplified model used for the data from Capa et al. (2014), because the distributions of perceived onsets are unlikely to differ for reference (small number) and test (large number) stimuli. The numerical magnitude of a digit can only be extracted through cognitive processing once the stimulus image has been peripherally processed and transmitted up the visual pathway, and such late processing cannot retroactively affect perceived onsets. Then, at the stage in which stimulus onsets are perceived, test and reference stimuli might only differ in their overall luminance (given by the spatial extent of the black area making up each digit), but these were reportedly equated by making all digits contain the same number of black pixels. Our analyses thus proceeded as described in the preceding section for the data from Capa et al. (2014), although we also fitted the general model in the unlikely event that it provided a better account of the data (which was not the case; those results will not be presented in detail).

Results

Linares and Holcombe (2014)

Figure 4a shows data and fitted IC model psychometric functions for each observer in each task (Table 1 for parameter estimates). The G^2 statistic was not significant for any observer¹ and model curves follow the path of the data as closely as

the occasionally noisy data permit. Thus, performance in both tasks can be understood as emanating from identical timing processes governed by parameters λ_a , λ_v , and τ with the participation of task-dependent decisional and response processes that vary across tasks. Despite common timing parameters, the peak of Ψ_{SJ} (red vertical line in each panel) does not occur at the same Δt as the 50% point on Ψ_{TOJ} does (blue vertical line in each panel), mostly because the latter is determined by the guessing parameter ξ , with which observers give AF or VF responses when they judge simultaneity (Fig. 2b). Indeed, individual estimates of ξ (Table 1) differ meaningfully from 0.5 except for observer SN, revealing that observers tend to give imbalanced AF and VF responses upon S judgments. For this reason, the 50% point on Ψ_{TOJ} does not reflect the true location of the (sensory) PSS, and a comparison with a PSS defined at the location of the peak of Ψ_{SJ} is inadequate. The lack of an S response option in TOJ tasks seems to prompt attentive observers to push their resolution limit to minimize the need for guessing. This results in a functionally narrower central region in decision space or, in other words, a smaller value for parameter δ . Evidence to this effect has been found in other studies (García-Pérez & Alcalá-Quintana, 2012a) and is apparent in Table 1: estimated $\delta^{(SJ)}$ is invariably larger than estimated $\delta^{(TOJ)}$ for each observer. A two-tailed, paired-samples t test for means revealed that these differences were statistically significant ($t_6 = 3.36$; $p = 0.015$).

Figure 4b shows the results of independently fitting a scaled Gaussian to SJ data and a cumulative Gaussian to TOJ data. This approach describes the paths of the data without links to underlying processes and only serves the goal of identifying landmarks (peaks or percent points) on the fitted functions. The fitted functions nevertheless do not do justice to the data. They cannot capture conspicuous asymmetries or flat plateaus in SJ data, and nonmonotonic patterns or reductions in slope around the null delay in TOJ

¹ We used the 0.05 significance level in all cases, but note that Table 1 (and subsequent tables) includes the actual p value of each statistic, and that p values are reported for all of our statistical analyses.

Table 1 Estimated IC model parameters in the joint fit to SJ and TOJ data for each observer in the experiment by Linares and Holcombe (2014)

Obs.	G^2	p	$1/\lambda_v$	$1/\lambda_a$	τ	$\delta^{(SJ)}$	$\epsilon_{VF}^{(SJ)}$	$\epsilon_S^{(SJ)}$	$\epsilon_{AF}^{(SJ)}$	$\delta^{(TOJ)}$	$\epsilon_{VF}^{(TOJ)}$	ξ	$\epsilon_{AF}^{(TOJ)}$
AH	18.54	0.235	47.66	28.42	56.83	195.03	—	—	—	154.89	—	0.376	0.136
AL	10.47	0.574	19.65	54.89	30.85	161.25	0.779	0.091	—	46.09	0.137	1.000	0.219
DL	15.09	0.372	21.73	62.58	60.14	107.74	0.438	—	—	83.00	0.358	0.000	—
GS	12.86	0.613	71.47	1.00	-50.26	107.40	—	—	—	94.89	0.225	0.223	—
SG	18.62	0.289	10.31	144.03	149.16	231.61	—	—	—	183.82	—	1.000	—
SM	20.94	0.074	43.64	57.04	24.13	131.59	0.683	—	—	117.20	0.201	0.698	0.060
SN	6.63	0.980	65.72	15.20	-25.16	106.24	—	—	—	40.90	—	0.554	—

The columns labeled G^2 and p give the value and p value of the likelihood-ratio goodness-of-fit statistic. Parenthetical superscripts for δ s and ϵ s indicate the task (SJ or TOJ) to which the parameter applies. Long dashes in cells for ϵ s indicate that the parameter is excluded from the fitted model

data. Surely as a result of this failure to accommodate the data, the G^2 statistic (with 8 and 7 degrees of freedom respectively for SJ and TOJ data) was significant in 5 of 14 cases (35.7%): in the SJ task, for observers AH ($G^2 = 35.60$; $p < 0.001$), AL ($G^2 = 21.11$; $p = 0.007$), and SM ($G^2 = 21.62$; $p = 0.006$); in the TOJ task, for observers DL ($G^2 = 21.45$; $p = 0.003$) and SM ($G^2 = 16.88$; $p = 0.018$).

Note also in Table 1 that the good fit of the IC model was attained with only six parameters for observers SG and SN, with seven parameters for observers AH and GS, and with eight, nine, and ten parameters respectively for observers DL, SM, and AL. In contrast, the descriptive approach estimates a fixed number of seven parameters per observer. Although it seems clear in Fig. 4 that the better fit of the IC model is structural and not attained simply by recourse to more parameters, Fig. 5 shows a scatter plot of the BIC of each model across observers. A model with a smaller BIC accounts for the data more efficiently, provided the models under comparison fit the data similarly well in the first place

(which is not the case here, as the descriptive approach failed to fit the data properly in a number of cases). By the BIC, the IC model also outperforms the descriptive approach for all observers, even in cases in which it uses more parameters to account for the data.

Finally, related to the number of parameters needed to account for the data, note in Table 1 that error parameters were rarely needed to account for SJ data, whereas they were needed more often to account for TOJ data. This seems to be caused by a further difference between the tasks in terms of the responses they request. Specifically, some observers seem to identify non-simultaneity easily, but they find difficulties indicating which of the two stimuli came first. Then, errors in the TOJ task (defined as failures to correctly report which stimulus was presented first) are observed for delays at which no errors are observed in the SJ task (e.g., data at large negative or large positive visual delays for observers AH, AL, and GS in Fig. 4).

It is interesting to look at the picture that the IC model fit offers with respect to the issue investigated by Linares and Holcombe (2014), namely, differences in perceptual latency across tasks. First, note that the location of the peak of Ψ_{SJ} (the classical estimate of the PSS) is not interpretable as the mean difference between visual and auditory latencies (more generally, between the latencies of reference and test stimuli). Hence, the location of the peak of Ψ_{SJ} does not reflect the true PSS. As discussed in detail two paragraphs down, the mean difference in perceptual latency is $1/\lambda_v - 1/\lambda_a + \tau$ under the IC model, whereas Ψ_{SJ} peaks instead at $\Delta t = \delta(\lambda_v - \lambda_a)/(\lambda_v + \lambda_a) - \tau$ also under the model (García-Pérez & Alcalá-Quintana, 2012a). Second, even if $\xi = 0.5$, the 50% point on Ψ_{TOJ} occurs at $\Delta t = \delta(\lambda_v - \lambda_a)/(\lambda_v + \lambda_a) - \tau + (\log \lambda_v - \log \lambda_a)/(\lambda_v + \lambda_a)$ (García-Pérez & Alcalá-Quintana, 2012a), matching the location of the peak of Ψ_{SJ} only when $\lambda_v = \lambda_a$. Then, (classical) PSS estimates from SJ and TOJ tasks have diverse influences and their comparison cannot reveal differences in perceptual latency.

In contrast, some of the IC model parameters describe the assumed distributions of perceived onsets and could thus be

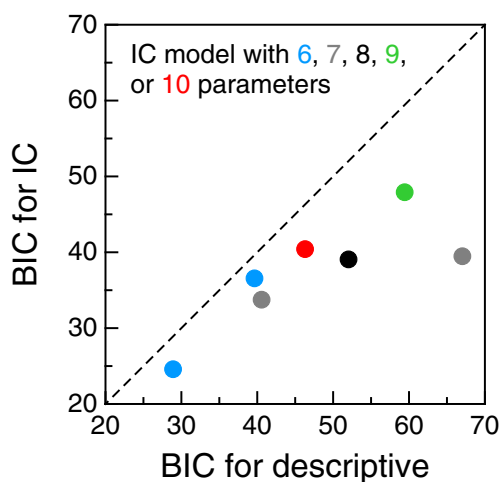


Fig. 5 Plot of the BIC of the IC model against the BIC of the descriptive model for fitting data from the study of Linares and Holcombe (2014). Each symbol denotes an observer. The color of each symbol indicates the number of parameters in the IC model fit (see legend); the descriptive model fit always included seven parameters

used to refer performance back to perceptual latencies. Specifically, from Eq. (2) with $\Delta t_i = 0$ for $i \in \{v, a\}$, the mean and variance of the distributions of perceived onsets (perceptual latencies) are respectively $\mu_i = 1/\lambda_i + \tau_i$ and $\sigma_i^2 = 1/\lambda_i^2$. Estimating the distributions of visual and auditory latencies could be possible under a model-based approach, if it was not because SJ and TOJ tasks are instances of MSS and, thus, have the two inescapable shortcomings discussed next.

Even though observers rely on separate perceived latencies to make their judgment in each trial, they are requested to respond in a way that only the relative difference in latency (not the individual values) manifests under either task. This became clear during model development at the point in which parameters τ_a and τ_v combine additively into τ , which is the only parameter that can be estimated with these tasks. This is not a property of this model, but the consequence of a task requiring the comparison of two sensory inputs (two perceptual latencies in this case). Then, the mean perceptual latency of visual or auditory signals cannot be estimated in absolute terms and only the difference between them can be estimated at $\mu_v - \mu_a = 1/\lambda_v - 1/\lambda_a + \tau$, identical for SJ and TOJ tasks when timing parameters are common. From parameter estimates in Table 1, the difference is negative at -4.4 only for observer AL and ranges from 10.7 for observer SM to 76.1 for observer AH. Estimated visual latencies that are longer than estimated auditory latencies are consistent with the slower peripheral processing and neural transmission of visual compared to auditory signals (see references to this effect in Vroomen & Keetels, 2010). Despite consistency with other sources of evidence, these estimated differences between visual and auditory latencies must be incorrect because of the second shortcoming of MSS.

In the general model, the central region in decision space is assumed to be symmetrically placed around the origin for reasons discussed above (Fig. 1b). Yet, this region might well be asymmetrically placed with boundaries at δ_1 and δ_2 (with $\delta_1 \neq -\delta_2$) as a result of decisional bias (Fig. 3b). Because this asymmetry is confounded with parameter τ in the general model, fitting the model under the necessary assumption that $|\delta_1| = |\delta_2| = \delta$ converts nominal estimates of τ into estimates of $\tau + (\delta_1 + \delta_2)/2$, with unknown δ_1 and δ_2 . Then, estimates of $\mu_v - \mu_a$ presented in the preceding paragraph are contaminated by a decisional bias whose magnitude cannot be assessed with SJ or TOJ tasks involving test and reference stimuli of different types.

In sum, the good fit of the IC model with common timing parameters to the data of Linares and Holcombe (2014) implies that identical perceptual latencies under both tasks cannot be ruled out, although estimating their values (or the difference between them) is hampered by the inevitable shortcomings of MSS. Some ways around these difficulties will be presented in the Discussion.

Capa et al. (2014)

Figure 6 shows data and fitted psychometric functions for each observer in each task, separately for schizophrenia patients (Fig. 6a) and normal controls (Fig. 6b), and with a final panel for average data and average fitted functions in each group. (Some observers were excluded from these averages; see the caption to Fig. 6.) Parameter estimates are listed in Tables 2 and 3. Compared with the analysis carried out by Capa et al. (2014), we did not collapse data at negative and positive delays of the same absolute magnitude. In fact, TOJ data often are asymmetric about the null delay (e.g., TOJ data for patients P01, P08, and P09 or controls C01, C06, C08, and C10). This asymmetry is captured by the fitted model, as also is the approximate symmetry of other data (e.g., TOJ data for patients P14, P19, and P20 or controls C03, C13, and C20). Asymmetric TOJ data are mostly determined by the response bias parameter ξ , but an asymmetric central region in decision space also plays a role. Several other peculiar aspects of the data are well captured by the model, as discussed next.

First, note that some observers seemed to respond to the TOJ task in reverse, as the proportion of RS responses decreases as right delay increases (see TOJ data for patients P04, P05, P15, and P18 or for control C19). The model nevertheless fits the data well by assuming common timing parameters under both tasks, because this response pattern is taken to reflect response errors (Alcalá-Quintana & García-Pérez, 2013) that do not hamper the estimation of timing and decisional parameters determining other aspects of observed performance. Second, note that patient P03 responded to the TOJ task as if he/she were performing an SJ task instead, using the designated response key for RS judgments to report asynchrony. The model also accommodates these data via response error parameters, adequately describing other aspects of observed performance via timing and decisional parameters. Third, note that patient P17 did not give a single RS response in the TOJ task, providing uninformative data that are nevertheless compatible with the model via a broad central region in decision space coupled with a strong bias against the RS response. Fourth, the path of TOJ data often display a remarkable reduction in slope around the null delay (e.g., data from patients P02, P06, P07, P08, P10, P16, and P19 or from controls C01, C02, C05, C11, C12, C13, C15, C16, C17, C18, and C20). This characteristic is expected from the response bias parameter ξ (Fig. 2b) and is well captured by the model curves. Fifth, note that SJ data show bilateral symmetry, which is expected when the same stimulus is used to deliver delays and, hence, the distribution of perceived onsets is the same for both stimuli (Fig. 3). Model curves incorporating this assumption fit the data well, although the vertical axis of bilateral symmetry is not always at the null delay (e.g., data from patients P15 and P17) due to asymmetric placement of the

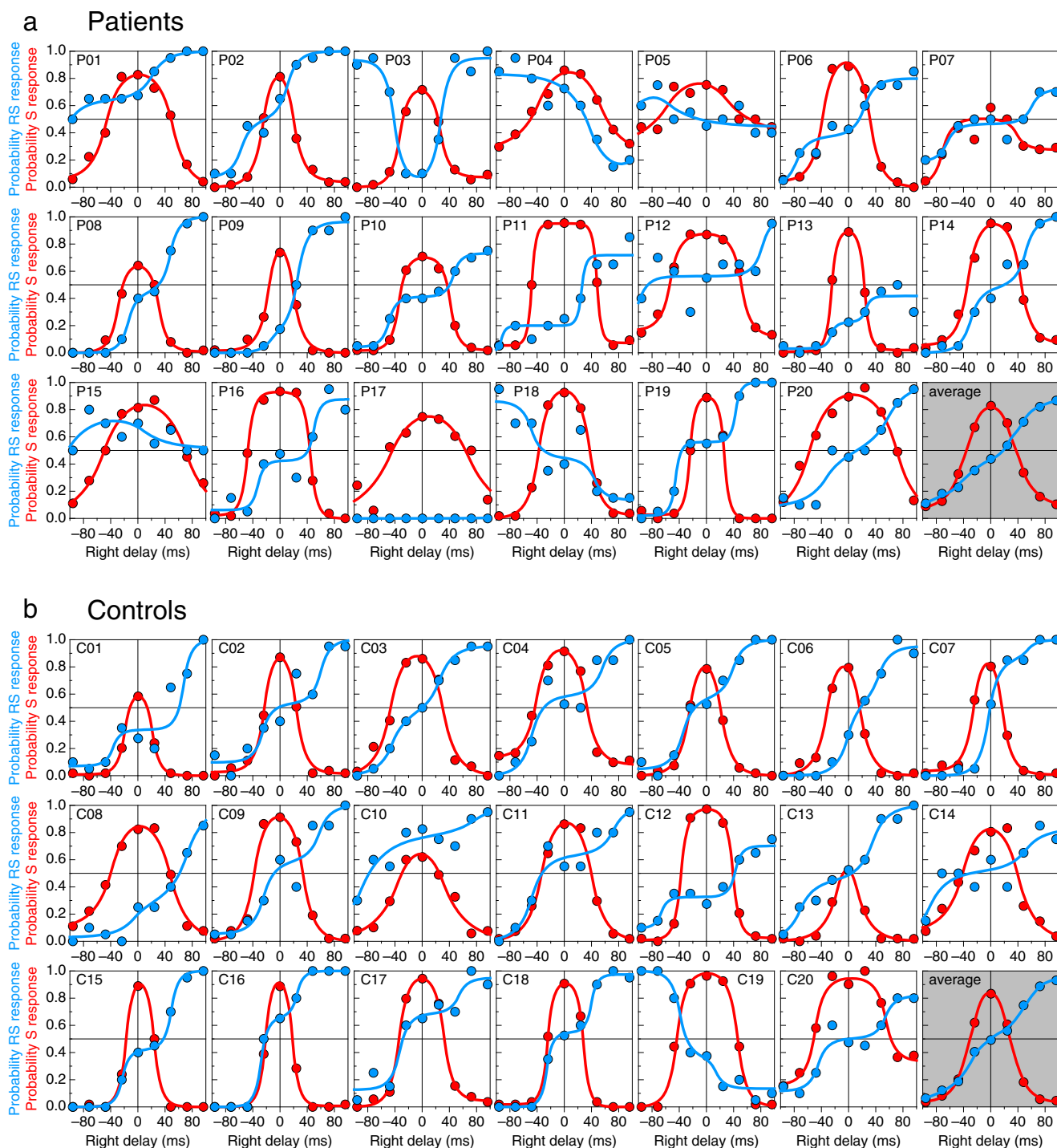


Fig. 6 Data and fitted psychometric functions for the study of Capa et al. (2014) under the IC model in the sample of patients (a) and normal controls (b). Results for the SJ task are shown in red; results for the TOJ task are shown in blue. Each panel pertains to an observer; the grayed panel at the bottom right in each part plots data and

psychometric functions aggregated across observers. Due to their peculiar responses in the TOJ task, data and fitted curves for patients P03, P04, P05, P15, P17, and P18 and control C19 are excluded from these averages

decision boundaries (Fig. 3c). Bilateral symmetry sometimes breaks at the tails due to response errors (e.g., SJ data from patients P03, P06, P07, P11, and P14 or from control C20). Finally, note that TOJ performance again indicates errors of

identification of temporal order when SJ performance reveals errorless identification of asynchrony (e.g., SJ and TOJ data points from patients P01 and P15 or controls C12 and C13 at large negative delays, or from patients P06, P10, and P13 or

Table 2 Estimated IC model parameters in the joint fit to SJ and TOJ data for each schizophrenia patient in the experiment by Capa et al. (2014)

Obs.	G^2	p	$1/\lambda$	$\delta_1^{(SJ)}$	$\delta_2^{(SJ)}$	$\epsilon_{VF}^{(SJ)}$	$\epsilon_S^{(SJ)}$	$\epsilon_{AF}^{(SJ)}$	$\delta_1^{(TOJ)}$	$\delta_2^{(TOJ)}$	$\epsilon_{VF}^{(TOJ)}$	ξ	$\epsilon_{AF}^{(TOJ)}$
P01	4.67	0.913	19.60	-46.85	51.31	0.047	0.100	—	-114.21	20.83	—	0.641	—
P02	4.14	0.941	13.65	-24.70	19.83	—	—	0.039	-57.22	10.61	0.069	0.493	—
P03	8.12	0.422	11.36	-32.86	26.51	—	0.236	0.073	-41.23	27.15	0.936	0.033	0.050
P04	8.83	0.453	28.80	-37.65	52.80	0.272	—	0.238	37.34	133.79	0.832	0.000	—
P05	9.08	0.430	41.83	-56.96	29.20	0.281	—	0.385	-123.03	-57.63	—	1.000	0.556
P06	5.35	0.866	14.20	-38.19	31.52	0.039	—	—	-76.76	22.94	—	0.373	0.200
P07	5.03	0.754	12.52	-71.38	38.08	—	0.493	0.275	-64.61	60.35	0.184	0.465	0.276
P08	4.81	0.903	11.14	-29.03	30.50	—	0.316	0.009	-14.35	47.41	—	0.445	—
P09	9.38	0.497	12.83	-16.55	18.92	0.015	—	—	-14.45	25.76	—	0.143	0.037
P10	0.71	0.998	9.64	-34.42	41.33	0.015	0.287	0.020	-49.12	46.06	0.042	0.411	0.268
P11	5.46	0.707	5.10	-47.97	48.06	0.054	0.049	0.073	-92.48	25.31	—	0.199	0.282
P12	10.46	0.315	12.21	-53.28	51.32	0.163	0.117	0.119	-104.38	82.01	—	0.563	—
P13	7.75	0.355	6.61	-25.22	23.69	0.008	0.089	0.020	-25.95	25.48	0.032	0.224	0.583
P14	9.10	0.523	13.70	-33.70	44.06	0.053	—	0.070	-29.38	51.17	—	0.480	—
P15	7.01	0.798	33.62	-50.47	70.31	—	—	—	-114.47	8.86	—	0.791	0.486
P16	13.99	0.082	7.92	-47.82	44.53	0.023	0.068	—	-32.68	49.72	0.062	0.425	0.124
P17	16.38	0.174	40.87	-46.78	66.56	—	—	—	—	—	—	—	—
P18	13.25	0.152	13.56	-37.58	37.17	—	—	0.022	-48.53	44.10	0.864	0.445	0.865
P19	7.39	0.597	6.30	-24.50	26.46	0.027	0.094	—	-45.53	42.85	0.021	0.562	—
P20	13.66	0.252	26.76	-57.69	70.95	—	—	—	-33.37	57.76	0.081	0.492	—

The columns labeled G^2 and p give the value and p value of the likelihood-ratio goodness-of-fit statistic. Parenthetical superscripts for δ s and ϵ s indicate the task (SJ or TOJ) to which the parameter applies. Long dashes in cells for ϵ s indicate that the parameter is excluded from the fitted model. Long dashes in all cells pertaining to decisional and response TOJ parameters for observer P17 indicate that TOJ data were uninformative (see the panel for this observer in Fig. 6a)

control C12 at large positive delays). All of these aspects are adequately captured through suitable model parameters, thus permitting an interpretable account of the timing, decisional, and response processes underlying observed performance within and across tasks.

Note in Tables 2 and 3 that the G^2 statistic was not significant for any patient and was significant for only two controls (C02 and C20). A statistically good fit is not surprising given that model curves follow the path of the data very closely for each observer, even in the two cases in which the G^2 statistic was significant. There are nevertheless meaningful differences between the location of the peak of Ψ_{SJ} and the location of the 50% point on Ψ_{TOJ} (when the latter exits) in the panel for each observer in Fig. 6, differences that are also patent at the average level among patients (see the grayed panel in Fig. 6a). These differences occur despite the fact that parameter λ is estimated to have the same value in both tasks under our joint fit, which reveals again that the different PSSs obtained with SJ and TOJ tasks are compatible with timing process operating identically under both tasks.

Given the manifest performance differences across patients and controls at the average level (grayed panels in Figs. 6a, b), differences in parameter estimates across groups are worth looking at. The average $1/\lambda$ was 17.11 in the group of patients

and 14.72 in the group of controls. This difference was not significant by a two-tailed, separate-variance two-sample t test for means ($z = 0.83$; $p = 0.404$), revealing that timing processes do not differ between patients and controls. Differences between groups in decisional parameters also were sought by analyzing the width $\delta_2 - \delta_1$ and the center $(\delta_1 + \delta_2)/2$ of the central region in decision space (see scatter plots of these measures across tasks for each group in Fig. 7). Parameter estimates from patient P17 were excluded from these analyses, as this observer did not provide usable data for the estimation of task-dependent parameters under the TOJ task. The average width $\delta_2^{(SJ)} - \delta_1^{(SJ)}$ was 80.18 among patients and 62.44 among controls, whereas the average width $\delta_2^{(TOJ)} - \delta_1^{(TOJ)}$ was 93.10 among patients and 88.53 among controls. A 2×2 ANOVA with task (SJ vs. TOJ) as a repeated-measures factor, group (patients vs. controls) as a grouping factor, and $\delta_2 - \delta_1$ as the dependent variable revealed significant main effects of task ($F_{1,37} = 17.587$, $p < 0.001$) with no main effects of group ($F_{1,37} = 2.072$, $p = 0.158$) and no interaction ($F_{1,37} = 2.003$, $p = 0.165$). This shows the absence of differences between patients and controls as to the width of the central region in decision space, although both groups display a higher temporal resolution (narrower central region) under the SJ task. A similar ANOVA using $(\delta_1 + \delta_2)/2$ as the dependent variable

Table 3 Estimated IC model parameters in the joint fit to SJ and TOJ data for each normal control in the experiment by Capa et al. (2014)

Obs.	G^2	p	$1/\lambda$	$\delta_1^{(SJ)}$	$\delta_2^{(SJ)}$	$\epsilon_{VF}^{(SJ)}$	$\epsilon_S^{(SJ)}$	$\epsilon_{AF}^{(SJ)}$	$\delta_1^{(TOJ)}$	$\delta_2^{(TOJ)}$	$\epsilon_{VF}^{(TOJ)}$	ξ	$\epsilon_{AF}^{(TOJ)}$
C01	11.86	0.221	9.40	-19.14	21.10	0.008	0.339	—	-36.81	67.28	0.071	0.338	—
C02	21.58	0.010	11.86	-22.97	23.76	0.025	—	0.015	-25.97	59.76	0.098	0.530	—
C03	6.47	0.840	18.23	-46.76	30.36	—	—	—	-43.89	25.52	—	0.461	0.045
C04	12.75	0.238	14.93	-40.88	31.95	0.132	—	0.080	-48.11	56.15	—	0.588	—
C05	7.79	0.732	14.34	-24.88	21.03	—	—	—	-34.75	38.83	0.047	0.572	—
C06	16.64	0.119	15.06	-31.64	16.92	—	—	—	-1.62	44.36	—	0.543	0.047
C07	7.36	0.691	12.30	-25.20	16.89	0.034	—	0.008	-3.32	53.93	—	0.856	—
C08	11.19	0.343	23.56	-38.14	47.50	0.093	—	—	-5.67	72.20	0.029	0.283	—
C09	14.87	0.188	14.62	-38.11	33.64	—	—	—	-24.58	57.46	0.054	0.587	—
C10	10.38	0.583	32.09	-35.33	30.52	—	—	—	-89.05	75.31	—	0.775	—
C11	8.83	0.717	17.92	-33.69	39.95	—	—	—	-48.29	72.03	—	0.632	—
C12	4.86	0.677	8.21	-37.44	38.94	0.009	0.021	0.025	-66.55	42.26	0.095	0.326	0.298
C13	10.95	0.533	19.01	-16.20	12.04	—	—	—	-71.74	34.67	—	0.449	—
C14	16.34	0.090	23.57	-40.39	39.96	0.066	—	—	-89.37	51.33	—	0.515	0.174
C15	10.72	0.553	8.86	-17.38	23.71	—	—	—	-22.98	49.10	—	0.422	—
C16	8.80	0.640	8.12	-21.46	18.22	—	—	0.006	-28.10	24.95	—	0.659	—
C17	12.39	0.192	11.67	-34.15	32.04	—	—	0.044	-33.26	53.82	0.126	0.687	0.051
C18	2.39	0.984	7.23	-24.74	28.24	0.018	0.069	—	-25.84	39.71	—	0.547	0.026
C19	15.38	0.052	11.20	-43.59	45.86	—	—	0.005	-38.95	14.13	1.000	0.371	0.865
C20	16.31	0.022	12.23	-49.44	54.59	0.162	0.045	0.335	-44.84	54.05	0.116	0.503	0.182

The columns labeled G^2 and p give the value and p -value of the likelihood-ratio goodness-of-fit statistic. Parenthetical superscripts for δ s and ϵ s indicate the task (SJ or TOJ) to which the parameter applies. Long dashes in cells for ϵ s indicate that the parameter is excluded from the fitted model

revealed no significant main effects of task ($F_{1,37} = 0.074, p = 0.788$) or group ($F_{1,37} = 1.946, p = 0.171$) and no significant interaction ($F_{1,37} = 2.984, p = 0.092$). This shows that patients and controls do not differ either as to decisional bias.

Having ruled out differences of timing or decisional processes between patients and controls by these analyses, the groups may still differ as to response errors. Average errors were the main focus of the analyses performed by Capa et al. (2014), and the following discussion attests to the fact that the groups differ mostly in their propensity to make errors. Gross evidence to this effect comes from the error parameters that turned out to be needed to account for the data. Note by a birds' eye view of Tables 2 and 3 that error parameters were needed more often for patients than for controls under either task and, through a closer look at the tables, that estimated values of the required error parameters are generally smaller for controls than they are for patients. In addition, for patients, estimated values were larger under the TOJ task than under the SJ task. This outcome reflects characteristics that can be readily observed in the data, as discussed next.

Consider first the SJ task. For controls (red symbols and curves in Fig. 6b), data and fitted curves generally decrease down to an ordinate of zero as the absolute value of right delay increases, with scarce evidence of response errors at the outer extremes; in contrast, patients (red symbols and curves in Fig. 6a) display more evidence of response errors at the outer

extremes. This shows clearly for average data (red symbols and curves in the grayed panels of Fig. 6) in that the left and right tails of the average curve follow the path of the raw data towards lower ordinate values in the group of controls than in the group of patients. Results are analogous in the TOJ task (blue symbols and curves in Fig. 6), even when observers are excluded whose inverted responses artificially increase the average error rate. (It should nevertheless be noted that the number of patients giving inverted responses was larger than the number of controls giving them, which might point to a further difference between the groups if these samples were representative.) While data and fitted curves for controls generally increase from zero at large negative delays to unity at large positive delays (Fig. 6b), data and fitted curves for patients often increase instead from above-zero levels (e.g., patients P01, P07, P12, and P20) and/or grow only to below-unity levels (e.g., patients P06, P07, P10, P11, P13, and P16) due to response errors. As a result, the average psychometric function for TOJ data is shallower for patients than it is for controls (Fig. 6, blue symbols and curves in the gray panels). Thus, the difference between the two average TOJ curves progressively increases as absolute delay increases, in line with the results reported by Capa et al. (2014) and their conclusion. It is nevertheless unclear why patients are more prone to errors than controls or why patients seem less able than controls to identify temporal order at relatively large delays

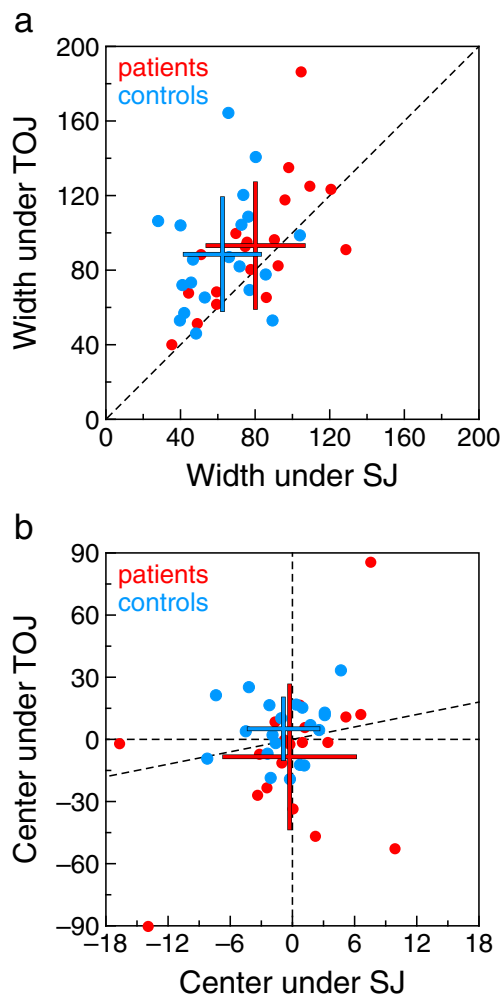


Fig. 7 Scatter plot of estimated width (a) and center (b) of the central region in decision space for each patient (red symbols) and control (blue symbols) in the study of Capa et al. (2014). In each panel, the oblique dashed line is the identity line (note that the horizontal and vertical ranges and scales are identical in panel a, but they are very different in panel b). Overlaid crosses sketch the distribution of the data in each group of observers. The two arms of each cross meet at the coordinates of the mean of each variable and the length of each arm spans the range from one standard deviation below the mean to one standard deviation above it

where their SJ data suggest that they actually perceive asynchrony.

The good fit of the IC model with common timing parameters to the data of Capa et al. (2014) also permits further inferences analogous to those discussed at the end of our analyses of the data of Linares and Holcombe (2014). Capa et al. did not address these issues, but it is useful to discuss what the IC model can say about perceptual latencies in the simplified conditions that arise when the two stimuli used to deliver temporal delays are identically processed up the sensory pathway. We will use the subscript “v” (for visual) in parameters, but the discussion clearly holds regardless of what type of stimulus is involved. In these conditions, the mean and variance of perceptual latencies are respectively $\mu_v = 1/\lambda_v + \tau_v$ and

$\sigma_v^2 = 1/\lambda_v^2 = 1/\lambda^2$ so that only the variance of perceptual latencies can again be estimated. The impossibility of a complete characterization of perceptual latencies in these simplified conditions is again a consequence of the limitations of MSS, and ways around them will be presented in the *Discussion*. On the other hand, the mean difference of perceptual latencies is trivially $\mu = \mu_v - \mu_v = 0$, and $\tau = \tau_v - \tau_v = 0$. This is what permits the estimation of potentially asymmetric boundaries δ_1 and δ_2 in decision space. In any case, the peak of Ψ_{SJ} does not occur at $\Delta t = 0$ when $\delta_1 \neq -\delta_2$ (Fig. 3c), so the classical PSS estimate in these conditions is only a measure of decisional bias, not of the average difference in perceptual latencies that is known to be 0 by the preceding argument. Similarly, the location of the 50% point on Ψ_{TOJ} (the classical TOJ estimate of the PSS) is affected by this decisional bias, besides the larger influence that the response bias parameter ξ has on its location. Then, also under these simplified conditions, classical estimates of the PSS from SJ and TOJ tasks (which will rarely match) do not portray perceptual latencies.

Li and Cai (2014)

Figure 8 shows data and fitted psychometric functions for each observer in each task, including a final panel for average data and average fitted curves. Parameter estimates are given in Table 4. As before, psychometric functions with common timing parameters under both tasks account remarkably well for the different paths described by the data and also for the large differences in classical estimates of the PSS (i.e., the location of the peak of Ψ_{SJ} and the location of the 50% point on Ψ_{TOJ}). The G^2 statistic was significant only for observer #4, for whom TOJ data were remarkably noisy (Fig. 8, corresponding panel). For all observers, SJ data show the bilateral symmetry expected when perceived onsets have the same distribution for both stimuli, and TOJ data also show clear evidence of the intermediate region of reduced slope that reflects guesses and is responsible for the spurious shift of the 50% point away from the null delay. A comparison of the SJ and TOJ performance of each observer again reveals errors in the TOJ identification of temporal order when SJ identification of asynchrony is errorless: See SJ and TOJ data points at the largest (negative and positive) delays for virtually all observers, but especially for #4, #6, #8, #12, and #16.

The average $1/\lambda$ across observers was 8.16. When we fitted the general model in which separate timing parameters are estimated for reference and test stimuli, the average $1/\lambda_s$ (for the reference, “small” stimulus) was 7.74 and the average $1/\lambda_l$ (for the test, “large” stimulus) was 8.79, with standard deviations of 2.64 and 2.39, respectively. We submitted these values to a simultaneous test for equality of means and variances (Bradley & Blackwood, 1989), which is more efficient and powerful than separate tests for equality of means and equality of variances (García-Pérez, 2013). The result showed that

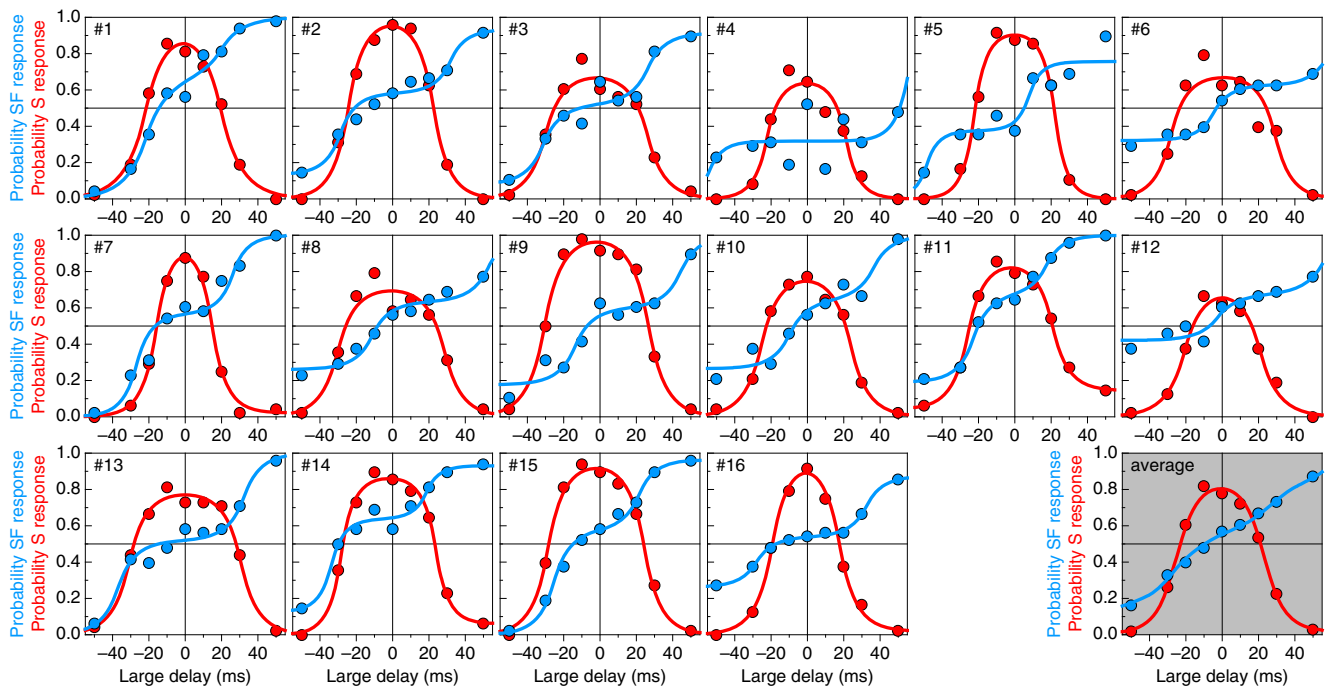


Fig. 8 Data and fitted psychometric functions for the study of Li and Cai (2014) under the IC model. Results for the SJ task are shown in red; results for the TOJ task are shown in blue. Each panel pertains to an

observer; the grayed panel at the bottom right plots data and psychometric functions aggregated across observers

means and variances do not differ significantly ($F_{2,14} = 0.745$; $p = 0.492$), as expected from the fact that perceived onsets cannot be affected by features determined cognitively once the stimuli have been processed upstream the sensory pathway. This leaves open the questions of how decisional or response

processes produce the different PSSs reported for SJ and TOJ tasks by Li and Cai (2014) and how the PSS under both tasks suggests that large numbers are perceived later than small numbers despite the fact that images of the numbers must be perceived before their numerical magnitudes can be assessed.

Table 4 Estimated IC model parameters in the joint fit to SJ and TOJ data for each observer in the experiment of Li and Cai (2014). The columns labeled G^2 and p give the value and p value of the likelihood-

ratio goodness-of-fit statistic. Parenthetical superscripts for δ s and ϵ s indicate the task (SJ or TOJ) to which the parameter applies. Long dashes in cells for ϵ s indicate that the parameter is excluded from the fitted model

Obs.	G^2	p	$1/\lambda$	$\delta_1^{(SJ)}$	$\delta_2^{(SJ)}$	$\epsilon_{VF}^{(SJ)}$	$\epsilon_S^{(SJ)}$	$\epsilon_{AF}^{(SJ)}$	$\delta_1^{(TOJ)}$	$\delta_2^{(TOJ)}$	$\epsilon_{VF}^{(TOJ)}$	ξ	$\epsilon_{AF}^{(TOJ)}$
1	10.37	0.584	10.71	-21.72	19.63	—	—	—	-21.64	20.00	—	0.663	—
2	9.20	0.513	7.79	-24.44	22.60	—	—	—	-28.80	31.79	0.133	0.583	0.066
3	11.25	0.259	8.88	-30.29	26.27	—	0.305	—	-31.28	26.22	0.077	0.520	0.085
4	28.18	0.003	6.60	-22.38	21.17	—	0.340	—	-53.77	54.93	—	0.318	—
5	15.69	0.109	5.93	-22.32	22.13	—	0.076	—	-48.49	7.74	—	0.375	0.244
6	13.66	0.189	8.09	-28.43	29.31	—	0.313	—	-4.74	59.11	0.321	0.627	—
7	13.20	0.280	7.27	-15.79	14.72	—	—	0.021	-26.91	26.09	—	0.567	—
8	8.80	0.551	8.75	-30.27	28.60	—	0.282	—	-11.16	52.36	0.261	0.635	—
9	12.11	0.355	8.75	-30.54	26.75	—	—	—	-13.67	44.45	0.177	0.600	—
10	7.20	0.706	8.86	-25.27	23.76	—	0.203	—	-9.80	36.44	0.265	0.643	—
11	2.79	0.947	8.01	-25.92	20.72	0.040	0.138	0.144	-22.82	17.45	0.191	0.673	—
12	6.07	0.809	9.50	-20.61	20.82	—	0.262	—	-2.50	54.84	0.420	0.674	—
13	5.26	0.918	8.58	-30.95	30.75	—	0.208	—	-36.94	32.35	—	0.518	—
14	4.95	0.763	6.36	-28.39	23.81	—	0.127	0.060	-33.78	18.59	0.126	0.638	0.069
15	6.90	0.735	8.11	-28.61	24.49	—	0.048	—	-25.44	21.32	—	0.576	0.036
16	3.14	0.958	8.40	-19.31	17.55	—	—	0.018	-28.09	32.46	0.261	0.539	0.125

To look at differences in decisional processes between SJ and TOJ tasks, we computed the width and center of the central region in decision space as we did with the data from Capa et al. (2014). Scatter plots are shown in Fig. 9. The average width $\delta_2^{(SJ)} - \delta_1^{(SJ)}$ was 48.65 and the average width $\delta_2^{(TOJ)} - \delta_1^{(TOJ)}$ was 58.50. A two-tailed, paired-samples t test for means showed that the difference is statistically significant ($t_{15} = -2.34$; $p = 0.034$), although an outlier with a large value for $\delta_2^{(TOJ)} - \delta_1^{(TOJ)}$ has certainly contributed to this outcome (Fig. 9a, stray data point). As a result of this stray data point, a two-tailed, paired-samples t test for variances showed that the difference between the variability of $\delta_2^{(TOJ)} - \delta_1^{(TOJ)}$ and the variability of $\delta_2^{(SJ)} - \delta_1^{(SJ)}$ is statistically significant ($t_{14} = -2.47$; $p = 0.027$). If the outlier is excluded, the central region

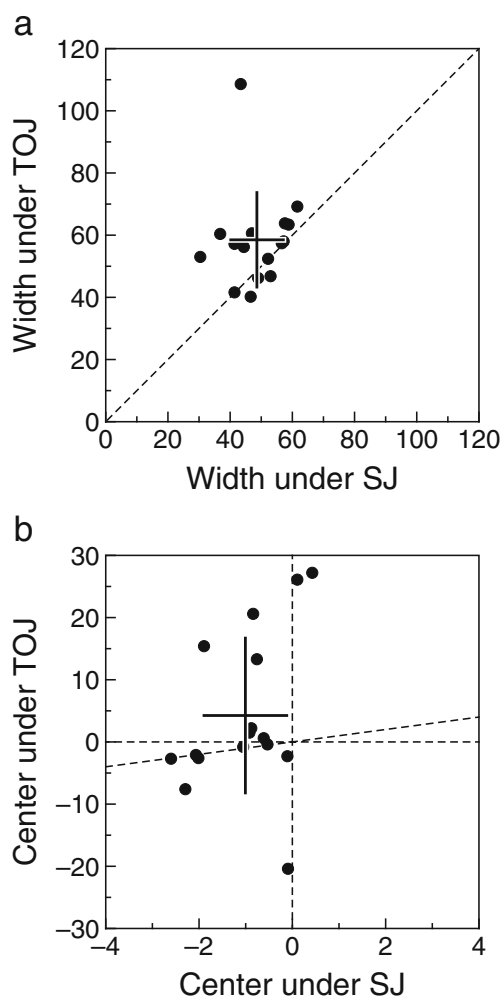


Fig. 9 Scatter plot of estimated width (a) and center (b) of the central region in decision space for each observer in the study of Li and Cai (2014). In each panel, the oblique dashed line is the identity line (note that the horizontal and vertical ranges and scales are identical in panel a, but they are very different in panel b). Overlaid crosses sketch the distribution of the data in each panel. The two arms of the cross meet at the coordinates of the mean of each variable and the length of each arm spans the range from one standard deviation below the mean to one standard deviation above it

in decision space is still significantly larger under the TOJ task ($t_{14} = -2.59$; $p = 0.021$) but intersubject variability is analogous in both tasks ($t_{13} = 0.37$; $p = 0.714$). On the other hand, the average center $(\delta_1^{(SJ)} + \delta_2^{(SJ)})/2$ was -1.00 and the average center $(\delta_1^{(TOJ)} + \delta_2^{(TOJ)})/2$ was 4.26 , a difference that was not statistically significant by a two-tailed, paired-samples t test for means ($t_{15} = -1.73$; $p = 0.104$). It is nevertheless clear from Fig. 9b that the variability of center locations across observers is much larger in the TOJ task than it is in the SJ task. A two-tailed, paired-samples t test for variances revealed that this difference is statistically significant ($t_{14} = -39.58$; $p = 9.0e-16$). Removing the outlier observation for width did not alter these results for center. Altogether, these analyses reveal that, compared with the SJ task, temporal resolution is functionally smaller in the TOJ task and that decisional bias is subject to larger individual variability in the TOJ task. A nonsymmetric placement of the central region in decision space under the SJ task is responsible for the shift of the classical PSS away from the null delay (Fig. 3c), whereas asymmetric placement combined with response bias in the TOJ task produces a shift of the classical PSS that is not identical to that observed in the SJ task.

As for response errors, they were rare under the SJ task in the experiment of Li and Cai (2014), as is evident from the widespread absence of error parameters for this task in Table 4 and also from the fact that data points (and, accordingly, fitted curves) fall down to an ordinate of zero at the most extreme (positive or negative) delays (Fig. 8). In contrast, response errors were more prevalent in the TOJ task, as manifested in TOJ data paths that do not cover the full range of the ordinate (Fig. 8; see also the frequent occurrence of non-zero error parameters for TOJ data in Table 4). These aspects of the data seem to reveal again a difficulty to identify temporal order for delays at which asynchrony is clearly perceived.

Discussion

At the surface level, the 50% point on a psychometric function fitted to TOJ data rarely matches the location of the peak of a psychometric function fitted to SJ data in within-subjects studies involving the same stimuli and conditions. The same is true about just-noticeable differences or other performance measures extracted from SJ and TOJ data, although these additional measures have not been discussed here (García-Pérez & Alcalá-Quintana, 2012a). These empirically observable and undeniable facts reflect a mixture of timing, decisional, and response processes that can only be separated out by fitting model-based psychometric functions with parameters that capture these influences. Assessing whether timing processes differ across SJ and TOJ tasks requires consideration of these separate influences in an analysis that allows for explicit tests of identity.

Our model-based analyses supported the identity hypothesis by showing that observed differences in performance across SJ and TOJ tasks in the experiments by Linares and Holcombe (2014), Capa et al. (2014), and Li and Cai (2014) can be accounted for when common timing processes sustain both tasks. Further evidence to this effect arises from the analyses of data sets from other published studies (see the [Electronic Supplementary Material](#)). As discussed in the *Introduction*, different PSSs in each task arise naturally because TOJ estimates are indirect and greatly affected by the observer's bias towards one of the response options when they perceive simultaneity. As shown in Fig. 2, this bias can shift the 50% point on Ψ_{TOJ} far away from the location of the peak of Ψ_{SJ} . The implicit assumption underlying classical analyses is that observers are unbiased and, thus, that they evenly choose both response options when they perceive simultaneity, but this assumption seems untenable in the light of the data.

The identity hypothesis is supported by other analyses that indicate that temporal-order and simultaneity judgments are mediated by identical timing processes (García-Pérez & Alcalá-Quintana, 2012a; Matsuzaki et al., 2014; Regener, Love, Petrini, & Pollick, 2014). Analogous model-based analyses in other areas of psychophysics also concur to indicate that differences in observed outcomes across tasks are a consequence of task-dependent decisional and response processes operating on the outcome of sensory processes that are invariant across tasks (Allik, Toom, & Rauk, 2014; García-Pérez & Peli, 2014).

We also have shown that IC model parameters relate to the timing, decisional, and response processes underlying performance in SJ and TOJ tasks and that their estimation is feasible. However, because SJ and TOJ tasks are both instances of MSS, the tasks themselves hide or combine some of the parameters that would additionally be needed for a complete characterization of timing and decisional processes. This shortcoming does not lessen validity to our conclusion regarding identical timing processes in SJ and TOJ tasks. Yet, they reveal that model-based analyses cannot overcome the inherent limitations of the task used to collect data when it comes to making quantitative inferences about processes that are confounded or hidden under the task. One way to overcome these limitations is to collect response times (RTs) separately and integrate RT data with SJ/TOJ data via a suitable model (Diederich & Colonius, 2015).

Our model-based approach suggests another way around these limitations. Specifically, the confound involving τ , δ_1 , and δ_2 in the general model can be solved by abandoning MSS (i.e., SJ and TOJ tasks) and using instead a paired-comparison method of the type used by Allan and Kristofferson (1974) or Fouriez, Capstick, Monette, Bellemare, Parkinson, and Dumoulin (2007), where each trial presents two temporal delays (i.e., two pairs of stimuli) that observers must compare.

The theoretical advantages of paired-comparison methods over MSS have been discussed elsewhere (García-Pérez, 2014; García-Pérez & Alcalá-Quintana, 2013) and their validity has been empirically demonstrated (García-Pérez & Peli, 2014). A model of performance in simultaneity and temporal-order judgments under paired-comparison tasks has been developed and tested empirically, but these results have not been reported yet. Use of the paired-comparison task followed by a model-based analysis of the data permits estimating all the relevant sensory and decisional parameters that are needed to test hypotheses or make inferences about perceptual latencies for different types of stimuli or about differences in timing, decisional, and response processes across stimulus types (e.g., in studies on cross-modal perception of synchrony), experimental conditions (e.g., in studies on prior entry), or groups of subjects (e.g., in studies assessing differences between patients and controls).

Acknowledgments This research was supported by grant PSI2012-32903 from Ministerio de Economía y Competitividad (Spain). Parts of the computations were carried out on EOLO, the MECD- and MICINN-funded HPC of Climate Change at Moncloa Campus of International Excellence, Universidad Complutense. The authors thank Waka Fujisaki, Anne Giersch, Timothy Hullar, Shuang-Xia Li, Daniel Linares, Keith Schneider, and their co-authors for making their data available for our analyses.

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(4), 972–998. doi:10.3758/s13428-013-0325-2
- Allan, L. G., & Kristofferson, A. B. (1974). Successiveness discrimination: Two models. *Perception and Psychophysics*, 15(1), 37–46. doi:10.3758/BF03205825
- Allik, J., Toom, M., & Rauk, M. (2014). Detection and identification of spatial offset: Double-judgment psychophysics revisited. *Attention, Perception, and Psychophysics*, 76(8), 2575–2583. doi:10.3758/s13414-014-0717-0
- Barnett-Cowan, M., & Harris, L. R. (2009). Perceived timing of vestibular stimulation relative to touch, light and sound. *Experimental Brain Research*, 198(2–3), 221–231. doi:10.1007/s00221-009-1779-4
- Barnett-Cowan, M., & Harris, L. R. (2011). Temporal processing of active and passive head movement. *Experimental Brain Research*, 214(1), 27–35. doi:10.1007/s00221-011-2802-0
- Bradley, E. L., & Blackwood, L. G. (1989). Comparing paired data: A simultaneous test for means and variances. *American Statistician*, 43(4), 234–235. doi:10.1080/00031305.1989.10475665
- Capa, R. L., Duval, C. Z., Blaison, D., & Giersch, A. (2014). Patients with schizophrenia selectively impaired in temporal order judgments. *Schizophrenia Research*, 156(1), 51–55. doi:10.1016/j.schres.2014.04.001
- Diederich, A., & Colonius, H. (2015). The time window of multisensory integration: Relating reaction times and judgments of temporal order. *Psychological Review*. doi:10.1037/a0038696
- Donohue, S. E., Woldorff, M. G., & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing

- abilities. *Attention, Perception, and Psychophysics*, 72(4), 1120–1129. doi:10.3758/APP.72.4.1120
- Fouriezos, G., Capstick, G., Monette, F., Bellemare, C., Parkinson, M., & Dumoulin, A. (2007). Judgments of synchrony between auditory and moving or still visual stimuli. *Canadian Journal of Experimental Psychology*, 61(4), 277–292. doi:10.1037/cjep2007028
- Fujisaki, W., & Nishida, S. (2009). Audio–tactile superiority over visuo–tactile and audio–visual combinations in the temporal resolution of synchrony perception. *Experimental Brain Research*, 198(2–3), 245–259. doi:10.1007/s00221-009-1870-x
- García-Pérez, M. A. (2013). Statistical criteria for parallel tests: A comparison of accuracy and power. *Behavior Research Methods*, 45(4), 999–1010. doi:10.3758/s13428-013-0328-z
- García-Pérez, M. A. (2014). Does time ever fly or slow down? The difficult interpretation of psychophysical data on time perception. *Frontiers in Human Neuroscience*, 8, 415. doi:10.3389/fnhum.2014.00415
- 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 and Review*, 19(5), 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*, 3, 94. doi:10.3389/fpsyg.2012.00094
- García-Pérez, M. A., & Alcalá-Quintana, R. (2013). Shifts of the psychometric function: Distinguishing bias from perceptual effects. *The Quarterly Journal of Experimental Psychology*, 66(2), 319–337. doi:10.1080/17470218.2012.708761
- García-Pérez, M. A., & Peli, E. (2014). The bisection point across variants of the task. *Attention, Perception, and Psychophysics*, 76(6), 1671–1697. doi:10.3758/s13414-014-0672-9
- Li, S.-X., & Cai, Y.-C. (2014). The effect of numerical magnitude on the perceptual processing speed of a digit. *Journal of Vision*, 14(12), 18. doi:10.1167/14.12.18.1–9.
- 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
- 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
- Matsuzaki, K. S., Kadota, H., Aoyama, T., Takeuchi, S., Sekiguchi, H., Kochiyama, T., & Miyazaki, M. (2014). Distinction between neural correlates of audiovisual temporal order and simultaneity judgments. *International Journal of Psychophysiology*, 94(2), 193. doi:10.1016/j.ijpsycho.2014.08.801
- Regener, P., Love, S., Petrini, K., & Pollick, F. (2014). Audiovisual processing differences in autism spectrum disorder revealed by a model-based analysis of simultaneity and temporal order judgments. *Journal of Vision*, 14(10), 429. doi:10.1167/14.10.429
- 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), 333–336. doi:10.1016/S0010-0285(03)00035-5
- Spence, C., & Parise, C. (2010). Prior-entry: a review. *Consciousness and Cognition*, 19(1), 364–379. doi:10.1016/j.concog.2009.12.001
- Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: Task and stimulus dependencies. *Experimental Brain Research*, 227(2), 249–261. doi:10.1007/s00221-013-3507-3
- Uppenkamp, S., Fobel, S., & Patterson, R. D. (2001). The effects of temporal asymmetry on the detection and perception of short chirps. *Hearing Research*, 158(1–2), 71–83. doi:10.1016/S0378-5955(01)00299-4
- 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. *Perception and Psychophysics*, 70(6), 955–968. doi:10.3758/PP.70.6.955
- van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & van de Par, S. (2010). Temporal order judgment criteria are affected by synchrony judgment sensitivity. *Attention, Perception, and Psychophysics*, 72(8), 2227–2235. doi:10.3758/APP.72.8.2227
- 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
- Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, and Psychophysics*, 72(4), 871–884. doi:10.3758/APP.72.4.871
- Wojtczak, M., Beim, J. A., Michey, C., & Oxenham, A. J. (2012). Perception of across-frequency asynchrony and the role of cochlear delays. *Journal of the Acoustical Society of America*, 131(1), 363–377. doi:10.1121/1.3665995
- Wojtczak, M., Beim, J. A., Michey, C., & Oxenham, A. J. (2013). Effects of temporal stimulus properties on the perception of across-frequency asynchrony. *Journal of the Acoustical Society of America*, 133(2), 982–997. doi:10.1121/1.4773350
- 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