Perceived Temporal Order and Simultaneity: Beyond Psychometric Functions

Miguel A. García-Pérez and Rocío Alcalá-Quintana

1 Introduction

Perception via the traditional senses of vision, audition, touch, gustation, or olfaction implies mechanisms (the sense organs) and neural structures (the sensory pathways) that transduce, transmit, and process physical energy (in vision, audition, and touch) or molecules (in gustation and olfaction). The same holds for the non-traditional senses of nociception, thermoception, equilibrioception, and others. In contrast, time does not emanate from a physical source and we do not have a sense organ for time, yet we have a vivid experience of it. Perception of time (chronoception) for brief events manifests in two remarkable abilities arguably subserved by separate processes. One is the ability to discriminate whether or not two punctate (instantaneous) events occurred simultaneously; the other is the ability to discriminate whether or not two brief events lasted the same duration. These punctate or brief events are delivered by presenting stimuli that can be perceived with our senses. Those stimuli are the occasion for some elusive machinery in the brain to extract the signals that render our perception of the time of occurrence of punctate events and our perception of temporal durations.

The duration of a stimulus is defined as the time elapsed between its onset and its offset. Then, perception of the duration of a stimulus presentation requires a second-stage process based on the output of first-stage processes determining the perceived onset and offset of the stimulus. This chapter focuses only on the first-stage processes and, specifically, on the methods used to assess their functioning and the utility of such methods to characterize timing processes. First-stage processes imply capture and transduction at the corresponding sense organ, followed by transmission of the sensory signals up the applicable pathway onto a central mechanism in the brain. Transduction, transmission, and processing of stimulus signals incur temporal delays that differ across sensory modalities but such delays also vary across stimulus types within the same modality and across repeated presentations of the exact same stimulus. When two punctate events occur simultaneously, the arrival times of

their signals at the central mechanism reveal differences in speed of processing. Investigating such differences across and within sensory modalities would be simple if arrival times were accessible, that is, if the location of the target center in the brain were identified and the arrival times of sensory signals at that center were recorded electrophysiologically. Because this is currently impossible, indirect behavioral data must be used instead.

The behavioral (psychophysical) methods more widely used for this purpose are the *binary simultaneity judgment* (SJ2) task and the also binary *temporal-order judgment* (TOJ) task, which consist of presenting two stimuli (A and B) with a *temporal offset*, *temporal delay*, or *stimulus onset asynchrony* (SOA) that varies across trials. Both tasks are variants of the *single-presentation method*, whose defining characteristic is that each trial delivers a single stimulus magnitude (an SOA here) and requests a categorical response from the observer. In an SJ2 trial, observers report whether or not the stimuli A and B were subjectively presented simultaneously; in a TOJ trial, instead, observers report which of the two stimuli appeared to be presented first, with no option to report subjective simultaneity even if that were what the observer perceived. The *ternary simultaneity judgment* (SJ3) task blends SJ2 and TOJ tasks by allowing observers to report the three judgments: A first, B first, or A and B simultaneous (Ulrich, 1987). A further response option has occasionally been allowed for observers to report that presentation was subjectively non-simultaneous but order was impossible to identify (e.g., Weiß & Scharlau, 2011), making up a *4-ary simultaneity judgment* (SJ4) task. Discussion of the SJ4 task will be deferred to a later section of this chapter.

A second set of psychophysical methods has also been used that belongs in the category of *dual-presentation* or *multiple-presentation* methods. In these cases, two or more SOAs are sequentially presented in a trial and the observer is asked to indicate which of them satisfies some condition. For instance, in the so-called *two-alternative forced-choice* (2AFC) task, each trial presents two pairs of stimuli (i.e., a pair of SOAs) and the observer reports in which pair the presentation was more (or less) synchronous (Allan & Kristofferson, 1974; Fouriez et al., 2007; Grant et al., 2004; Pastore & Farrington, 1996; van de Par & Kohlrausch, 2000; Yarrow et al., 2016); in the *match-to-sample* or *ABX* task, each trial presents a sample pair (i.e., a sample SOA) followed by two other pairs (two more SOAs) and the observer reports in which of the latter two pairs the (a)synchrony was more similar to (or more different from) that of the sample (Hillenbrand, 1984; Liberman et al., 1961; McGrath & Summerfield, 1985; van Eijk et al., 2009). Dual- and multiple-presentation methods are used less often than single-presentation methods and they will not be covered in detail.

Whichever task is used to collect data, observers' responses are tallied to compute the proportion of trials in which each judgment was reported at each of a set of SOAs. The most common method of analysis of data consists of fitting *psychometric functions*, continuous curves that match the path of the data. Figure 12.1 shows sample psychometric functions for SJ2, SJ3, and TOJ tasks without the data that might have given rise to them. When binary responses are involved (i.e., in SJ2 and TOJ tasks), only the psychometric function for one of the responses is needed (Figs. 12.1a and 12.1c); when ternary responses are involved (i.e., in SJ3 tasks), only the psychometric functions for two of the responses are needed although all three are often plotted (Fig. 12.1b).

The form of the function fitted to the data varies greatly in the literature. For TOJ data, the most common option is to fit a cumulative Gaussian or logistic function, which can be referred to an observer model with simple sensory and decisional components. For SJ2 data, the most common option is to fit a scaled Gaussian, which merely provides a description of the data because it cannot be referred to any observer model. Because SJ2 data are generally asymmetric (as illustrated in Fig. 12.1a) its rising and declining parts are sometimes fitted separately using monotonic functions, which describes the data better but again cannot be referred to an observer model. For SJ3 data, a mixture of these strategies is used across the implied psychometric functions (e.g., van Eijk et al., 2008), a strategy with which the probability that *some* response is given often falls short of (or exceeds) unity. The fitted functions are then used to summarize performance via measures such as the *point of subjective simultaneity* (PSS), the *synchrony range* (SR), the *synchrony boundaries* (SBs) or the *just noticeable difference* (JND). The PSS is defined as the SOA at which the psychometric function for “simultaneous” responses peaks in SJ2 or SJ3 tasks, or as the SOA at which the psychometric function evaluates to 0.5 in TOJ tasks (dotted vertical lines in Fig. 12.1). The SR is defined as the range of SOAs within which “simultaneous” responses are more prevalent than any other response in SJ2

FIGURE 12.1 Sample psychometric functions for SJ2 data (a), SJ3 data (b), and two sets of TOJ data (c). SOAs are assumed to be delivered via audiovisual stimulus pairs, with negative (positive) SOAs indicating that auditory onset preceded (lagged) visual onset.

or SJ3 tasks and the SBs are the SOAs at the endpoints of the SR. The JND, also known as the *difference limen* (DL), is half the range of SOAs over which the psychometric function in TOJ tasks increases from, say, 0.25 to 0.75.

The three tasks can be used to estimate all of these measures, but many within-subjects studies involving two or more of the tasks have shown that the estimates are inconsistent (Bedard & Barnett-Cowan, 2016; Binder, 2015; Capa et al., 2014; Donohue et al., 2010; Fujisaki & Nishida, 2009; Li & Cai, 2014; Linares & Holcombe, 2014; Love et al., 2013; Sanders et al., 2011; Schneider & Bavelier, 2003; van Eijk et al., 2008; Vatakis et al., 2008). These inconsistencies are difficult to understand considering that only the question that observers respond to at the end of each trial differs across tasks and, hence, that the underlying timing processes should be invariant. Analogous inconsistencies have been found in within-subjects studies in which performance on single- and dual-presentation tasks was compared (Stevenson & Wallace, 2013).

One reason for these inconsistencies is that conventional psychometric functions only capture observed performance, which is the final outcome of the interplay of sensory (timing), decisional, and response processes involved in any psychophysical task. Then, even on the reasonable assumption that the timing component of performance is identical in SJ2, SJ3, and TOJ tasks, differences in decisional and response components would produce differences in observed performance across them. These differences are somewhat misleading because they do not reflect actual differences in timing processes. Thus, separating out the components of observed performance is needed for a proper characterization of timing processes that is not tainted by the effects of other processes. Accomplishing this separation requires psychometric functions whose mathematical form is derived from a suitable observer model that explicitly represents all of the intervening components. As discussed earlier, cumulative Gaussian or logistic psychometric functions fitted to TOJ data can be referred to an observer model whereby “A first” responses are given when a decision variable exceeds a criterion and “B first” responses are given otherwise. Sensible as this observer model may seem, it makes an inconvenient prediction for SJ2 or SJ3 tasks, namely, that the observer will never give “simultaneous” responses. In other words, if observers had the ability that this model attributes them in the TOJ task (i.e., if they could *always* tell temporal order), why would they lose that ability in SJ2 or SJ3 tasks and report perceived simultaneity so often? Or, in reverse, if observers’ reports of perceived simultaneity in SJ2 and SJ3 tasks are genuine, by what mechanism do they give temporal-order responses in analogous occasions under the TOJ task? The TOJ task precludes reporting simultaneity, which does not imply that such judgments are never made. Then, TOJ data must necessarily be affected by how observers go about (mis)reporting temporal order when they perceive simultaneity.

The goal of this chapter is to describe an observer model that explicitly represents the timing, decisional, and response processes involved in single-presentation timing tasks, thus offering an integrated and coherent account of performance across tasks. The model elaborates on ideas put together by Sternberg and Knoll (1973) in their *independent channels* models, which were further developed by Schneider and Bavelier (2003). The resultant psychometric functions include parameters that characterize the underlying timing, decisional, and response processes. The model has been empirically validated extensively and the focus of this chapter is instead on its underpinnings and on practical aspects of its use for testing hypothesis or making inferences about how timing processes differ across manipulations (e.g., cuing) or across groups of observers (e.g., patients vs. normal controls).

2 The Observer Model

The observer model makes explicit assumptions about each of the processes whose participation is needed to respond on each trial. These processes include a *sensory component* that provides the evidence on which timing judgments are based, a *decisional component* that makes a judgment based on the sensory evidence, and a *response component* by which the judgment is expressed. The next three subsections describe each of these components and their referents, also characterizing them formally. The fourth subsection illustrates how the interplay of components shapes the psychometric functions in SJ2, SJ3, and TOJ tasks.

2.1 The Sensory Component

A punctate temporal event is signaled physically by, for example, the onset of a stimulus. Each of the two stimuli used to deliver an SOA must be peripherally processed by the corresponding sense organ and their neural signals must be transmitted up the corresponding sensory pathway onto a central mechanism. These operations incur delays that vary across sensory modalities and across stimuli within the same modality, resulting in differences in *arrival time* at the central mechanism and providing the evidence for timing judgments. Such delays are not fixed so that variability occurs across repeated presentations of the same stimulus.

The concept of arrival time is broader than what the previous description suggests, which only referred to physiological components. This may suffice in a description for simple stimuli such as a flash of light or a sound beep, for which the arrival time of signals at a central mechanism is perhaps also the time at which the presence of the stimulus is identified by the observer,

which is what actually yields the arrival time referred to in the model. With vestibular stimuli such as yaw rotation, signals are surely reaching the brain as soon as the movement starts but the observer will not identify the motion until some aspect of it (amplitude, speed, acceleration, etc.) attains a necessary magnitude determined by the observer's sensitivity. Similarly, for time-varying stimuli such as single-syllable utterances in audiovisual speech, auditory and visual signals reach the brain continually from the nominal onset of the stimulus but the referents for auditory and visual arrival times in the model are the auditory identification of certain sounds and the visual identification of their articulations, respectively.

Across repeated presentations, arrival times will have a distribution that is impossible to determine empirically. Nevertheless, some distributions may be hypothesized that satisfy the physical constraint of causality: The arrival time of a stimulus signal cannot precede the onset of the stimulus itself. If stimulus onset is regarded as the origin of time, suitable distributions for arrival times must have all their probability mass on the positive real line. This constraint rules out the normal distributions included in most observer models in psychophysics. Figure 12.2 shows three plausible candidates, namely, a shifted exponential distribution given by

$$g(t) = \lambda \exp[-\lambda(t - \tau)], \quad t \geq \tau, \quad (1)$$

a shifted gamma distribution given by

$$g(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} (t - \tau)^{\alpha-1} \exp[-\lambda(t - \tau)], \quad t \geq \tau, \quad (2)$$

where Γ is the gamma function, and a log-normal distribution given by

$$g(t) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\frac{(\log(t) - \mu)^2}{2\sigma^2}\right], \quad t \geq 0. \quad (3)$$

FIGURE 12.2 Sample distributions of arrival time. (a) Shifted exponential distribution given by Eq. 1. (b) Shifted gamma distribution given by Eq. 2. (c) Log-normal distribution given by Eq. 3. Parameter values are printed in the insets.

Parameter values are printed in Fig. 12.2 for each distribution, which were chosen so that arrival times have the same mean and variance in all cases. The form of the distribution of arrival times cannot be determined empirically but any particular choice is largely immaterial if it meets the above constraints. The shifted exponential distribution in Eq. 1 is simple and easily tractable mathematically, and it has often been used to model arrival latencies and peripheral processing times (e.g., Colonius & Diederich, 2011; Heath, 1984). In addition, this distribution has proven empirically adequate to account for observed performance in timing tasks (see García-Pérez & Alcalá-Quintana, 2012a, 2012b, 2015a, 2015b, 2015c). Thus, and without loss of generality, these are the distributions that will be used here.

Because SOAs are typically delivered via two different stimuli (and, generally, from different sensory modalities), the parameters of the corresponding distributions of arrival times will differ. In empirical studies, one of the stimuli is regarded as the reference so that SOA is defined as the relative delay with which the other (test) stimulus is presented. Thus, negative (positive) SOAs indicate that the onset of the test precedes (lags) the onset of the reference. Two distributions must thus be considered. The distribution of arrival times for the reference stimulus is subject to the constraint illustrated in Fig. 12.2, because the origin of time is set at its physical onset; in contrast, the (generally different) distribution of arrival times for the test stimulus will be shifted to the right or to the left according to the applicable SOA, potentially encroaching into the negative region of a time line whose origin is not at its own onset but at the onset of the reference. Formally, the arrival times (or *perceived onsets*, or *perceived latencies*) T_r and T_t of reference and test stimuli are random variables with densities g_r and g_t given by

$$g_i(t) = \lambda_i \exp[-\lambda_i(t - (\Delta t_i + \tau_i))], \quad t \geq \Delta t_i + \tau_i, \quad i \in \{r, t\} \quad (4)$$

where Δt_i is the onset time of stimulus i and λ_i and τ_i are the parameters of each distribution, which surely vary across observers and maybe also across experimental conditions. By the convention that sets the origin of time at the onset of the reference, $\Delta t_r = 0$ by definition and $\Delta t \equiv \Delta t_t$ is the SOA with which the pair is presented. The top row in Fig. 12.3 shows sample distributions at three different SOAs.

As discussed later, parameters λ_i and τ_i affect the shape of the observed psychometric function and provide information about speed of processing and variability of arrival times: The mean and standard deviation of arrival times for stimulus i are $1/\lambda_i + \tau_i$ and $1/\lambda_i$, respectively. These parameters thus inform of the sensory limits for perception of temporal order. If the standard deviations of arrival times are small, temporal order can be correctly perceived at

FIGURE 12.3 Effects of SOA on the distribution of arrival times for reference and test stimuli (top row) and on the distribution of arrival-time differences (bottom row). (a) SOA of -50 ms. (b) SOA of 0 ms. (c) SOA of 50 ms. In the top row, the distribution for the reference stimulus is invariant with SOA because the origin of time is at its onset; the distribution for the test stimulus instead shifts with SOA. In the bottom row, the distribution of arrival-time differences shifts to the right as SOA increases. Vertical lines at $D = \delta_1$ and $D = \delta_2$ demarcate regions associated with each type of judgment (labels at the top of each region); the probability of each judgment (numeral under each label) is the area under the distribution within the applicable region.

smaller SOAs than when standard deviations are large. Furthermore, large differences in the standard deviation of arrival times across stimuli produce large differences in the accuracy with which temporal order can be perceived at positive vs. negative SOAs (i.e., when the test stimulus precedes or lags the reference stimulus). Finally, large differences in average arrival time across stimuli produce discrepancies between physical and perceptual synchrony.

2.2 The Decisional Component

On any given trial, the perceived onset of each stimulus is a random value drawn from the corresponding distribution. These perceived onsets provide the evidence for a timing judgment, which the observer makes by application of a decision rule. Sternberg and Knoll (1973) and Schneider and Bavelier (2003) discussed several decision rules tailored to the response format of particular variants of the task. To detach the decisional component from the response component imposed by the task, we will consider a decision rule by which observers only make judgments at this stage, irrespective of the type of response later requested by the task. Observers' spontaneous reports to the effect that sometimes they guessed a response in TOJ trials because they could not tell which stimulus was presented first provides evidence that judgments precede responses and that the three types considered explicitly only in SJ3 tasks are made in all tasks. To better understand why judgments and responses must be separated, consider an experiment in which observers are given the SJ2, SJ3, or TOJ response options at random at the end of each trial. Because response options are revealed only after the stimuli were extinguished, a judgment must have been made before it can be expressed as a response. Such random mixture of trials from several tasks does not seem to have ever been used in timing perception, but Schneider and Komlos (2008) used it in a different context to make a similar point.

The decision variable is the *arrival-time difference* (or *perceived-onset difference*, or *latency difference*) $D = T_t - T_r$, which has the asymmetric Laplace distribution

$$f(d; \Delta t) = \begin{cases} \frac{\lambda_r \lambda_t}{\lambda_r + \lambda_t} \exp[\lambda_r (d - \Delta t - \tau)] & \text{if } d \leq \Delta t + \tau, \\ \frac{\lambda_r \lambda_t}{\lambda_r + \lambda_t} \exp[-\lambda_t (d - \Delta t - \tau)] & \text{if } d > \Delta t + \tau \end{cases} \quad (5)$$

where $\tau = \tau_t - \tau_r$. Combination of τ_t and τ_r into an aggregate parameter τ is a trivial consequence of tasks that involve differencing: The precise arrival time of each stimulus is immaterial for the judgment and only the difference

matters. The unfortunate consequence is that neither τ_t nor τ_r can be estimated separately, thus precluding the estimation of average arrival times for each individual stimulus although their variability (i.e., $1/\lambda_i$) can still be estimated. In addition, arrival-time differences are well characterized, with mean $1/\lambda_t - 1/\lambda_r + \tau + \Delta t$ and variance $1/\lambda_t^2 + 1/\lambda_r^2$. The bottom row in Fig. 12.3 shows the distribution of D for each case in the top row, which only shifts location with SOA.

Before an amendment that will be introduced later, the decision rule partitions the domain of D into three regions with boundaries at δ_1 and δ_2 (see the bottom row of Fig. 12.3). Then, “test-first” (TF) judgments arise when D is large and negative ($D < \delta_1$), “reference-first” (RF) judgments arise when D is large and positive ($D > \delta_2$), and “simultaneous” (S) judgments arise when D is small ($\delta_1 \leq D \leq \delta_2$). The probability of each judgment varies with SOA (Δt), as this shifts the distribution of arrival-time differences (see the bottom row of Fig. 12.3). Formally, the probabilities p_{TF} , p_S , and p_{RF} of each judgment vary with Δt as

$$p_{TF}(\Delta t) = F(\delta_1; \Delta t), \quad (6a)$$

$$p_S(\Delta t) = F(\delta_2; \Delta t) - F(\delta_1; \Delta t), \quad (6b)$$

$$p_{RF}(\Delta t) = 1 - F(\delta_2; \Delta t), \quad (6c)$$

where

$$F(d; \Delta t) = \int_{-\infty}^d f(z; \Delta t) dz = \begin{cases} \frac{\lambda_r}{\lambda_r + \lambda_t} \exp[\lambda_r(d - \Delta t - \tau)] & \text{if } d \leq \Delta t + \tau \\ 1 - \frac{\lambda_t}{\lambda_r + \lambda_t} \exp[-\lambda_t(d - \Delta t - \tau)] & \text{if } d > \Delta t + \tau \end{cases}. \quad (7)$$

In principle, δ_1 and δ_2 could be placed anywhere. An asymmetric placement such that $\delta_1 \neq -\delta_2$ reflects a *decisional bias* whereby the absolute magnitude of the arrival-time difference required to make a TF judgment is not the same as that required to make an RF judgment. This may represent a natural bias of the observer but it may also be caused by experimental manipulations that favor one type of judgment over the other. In contrast, when $\delta_1 = -\delta_2$ (as in the bottom row of Fig. 12.3) the decision rule is unbiased. The width $\delta_2 - \delta_1$ reflects the operating resolution of the observer: The narrower this region, the smaller the arrival-time difference that allows the observer to judge temporal order (whether or not such judgment is physically correct; note in the bottom panel

of Fig. 12.3a that the probability of an RF judgment is 0.03 even though the test stimulus was presented 50 ms before the reference).

As discussed later, the width and placement of the central region in decision space has consequences on the shape of the psychometric function under all tasks. Hence, estimating the decisional parameters δ_1 and δ_2 allows separating out their influence for an assessment of timing processes. Schneider and Bavelier (2003) noted that multiple location parameters are confounded in SJ2 and TOJ tasks, something that generalizes to all variants of the single-presentation method (García-Pérez & Alcalá-Quintana, 2013; Yarrow et al., 2011). Inspection of Eqs. 6 and 7 reveals that the location parameters affected here by this confound are τ , δ_1 , and δ_2 , and that only the aggregates $\delta_1 - \tau$ and $\delta_2 - \tau$ can be estimated. A compromise (though suboptimal) solution is to enforce the assumption of no decisional bias, that is, to assume $\delta_1 = -\delta$ and $\delta_2 = \delta$, leaving δ as the only free parameter representing the half-width of a symmetric central region in decision space. The implications as well as other ways around this confound will be discussed later in this chapter.

2.3 The Response Component

In each trial, the observer must report the judgment arising from the timing and decisional components just discussed, using the response format that the task allows. In addition, response errors may occur by which observers misreport their judgment. Evidence of response errors is often found empirically in the form of, e.g., exceptional TF responses at large negative SOAs where RF responses have been given in most other trials. These errors may be caused by carelessness, by an unnatural arrangement of the response interface, or by insufficient practice to use it properly. The probability of a response error is generally small, but errors seem to affect some responses more often than others (for empirical examples, see García-Pérez & Alcalá-Quintana, 2012a, 2012b, 2015a, 2015b, 2015c). The response component thus comprises the mapping of judgments onto one of the responses allowed by the task, with a potential for misreporting such judgments due to errors. This component produces the final data that delineate the observed psychometric function.

Mapping judgments onto responses is straightforward in the SJ3 task because there is a distinct response option for each possible judgment. In the absence of response errors, the psychometric functions in the SJ3 task are directly given by the functions p_{TF} , p_S , and p_{RF} in Eqs. 6 above. Response errors imply that each type of judgment has some probability of being misreported and, in such event, that the two forms that the misreport may take also have different probabilities. Let $\epsilon_X^{(SJ3)}$ denote the probability of misreporting judgment X in the SJ3 task and let κ_{X-Y} denote the probability of misreporting it as

response Y. Then, the final psychometric functions for TF, S, and RF responses in SJ3 tasks are

$$\Psi_{\text{TF}}^{(\text{SJ}_3)}(\Delta t) = \left(1 - \varepsilon_{\text{TF}}^{(\text{SJ}_3)}\right) p_{\text{TF}}(\Delta t) + \varepsilon_{\text{S}}^{(\text{SJ}_3)} \kappa_{\text{S-TF}} p_{\text{S}}(\Delta t) + \varepsilon_{\text{RF}}^{(\text{SJ}_3)} \kappa_{\text{RF-TF}} p_{\text{RF}}(\Delta t), \quad (8a)$$

$$\Psi_{\text{S}}^{(\text{SJ}_3)}(\Delta t) = \varepsilon_{\text{TF}}^{(\text{SJ}_3)} \kappa_{\text{TF-S}} p_{\text{TF}}(\Delta t) + \left(1 - \varepsilon_{\text{S}}^{(\text{SJ}_3)}\right) p_{\text{S}}(\Delta t) + \varepsilon_{\text{RF}}^{(\text{SJ}_3)} (1 - \kappa_{\text{RF-TF}}) p_{\text{RF}}(\Delta t), \quad (8b)$$

$$\Psi_{\text{RF}}^{(\text{SJ}_3)}(\Delta t) = \varepsilon_{\text{TF}}^{(\text{SJ}_3)} (1 - \kappa_{\text{TF-S}}) p_{\text{TF}}(\Delta t) + \varepsilon_{\text{S}}^{(\text{SJ}_3)} (1 - \kappa_{\text{S-TF}}) p_{\text{S}}(\Delta t) + \left(1 - \varepsilon_{\text{RF}}^{(\text{SJ}_3)}\right) p_{\text{RF}}(\Delta t). \quad (8c)$$

These expressions are easily unpacked with the help of the tree diagram in figure 12.4b of García-Pérez and Alcalá-Quintana (2012a), which is not reproduced here. Note also that Eqs. 8 reduce to Eqs. 6 if all ε s are zero (i.e., in the absence of response errors). The psychometric functions in Fig. 12.1b for the SJ3 task arise from these equations (without response errors) when $1/\lambda_t = 40$, $1/\lambda_r = 90$, $\tau = 60$, and $\delta = 100$.

Mapping judgments onto responses is also straightforward in SJ2 tasks because TF and RF judgments are unambiguously aggregated into a category of non-simultaneous judgments. Under our notation for error parameters, the psychometric function for S responses in SJ2 tasks is

$$\Psi_{\text{S}}^{(\text{SJ}_2)}(\Delta t) = \varepsilon_{\text{TF}}^{(\text{SJ}_2)} p_{\text{TF}}(\Delta t) + \left(1 - \varepsilon_{\text{S}}^{(\text{SJ}_2)}\right) p_{\text{S}}(\Delta t) + \varepsilon_{\text{RF}}^{(\text{SJ}_2)} p_{\text{RF}}(\Delta t) \quad (9)$$

(see the tree diagram in figure 12.4a of García-Pérez & Alcalá-Quintana, 2012a), which reduces to Eq. 6b when all ε s are zero. The psychometric function in Fig. 12.1a arises from this equation (without response errors) also when $1/\lambda_t = 40$, $1/\lambda_r = 90$, $\tau = 60$, and $\delta = 100$, which explains why the psychometric function for S responses is identical in Figs. 12.1a and 12.1b.

Finally, TOJ tasks force observers to give TF or RF responses upon S judgments. The mapping thus requires an extra *response bias* parameter ξ reflecting the probability with which the observer gives RF responses in such cases. Incorporating response errors as before (see the tree diagram in figure 12.4c of García-Pérez & Alcalá-Quintana, 2012a), the psychometric function for RF responses in TOJ tasks is

$$\Psi_{\text{RF}}^{(\text{TOJ})}(\Delta t) = \varepsilon_{\text{TF}}^{(\text{TOJ})} p_{\text{TF}}(\Delta t) + \xi p_{\text{S}}(\Delta t) + \left(1 - \varepsilon_{\text{RF}}^{(\text{TOJ})}\right) p_{\text{RF}}(\Delta t), \quad (10)$$

which reduces to Eq. 6c plus a fraction of Eq. 6b when all ε s are zero. Also in the absence of errors, when $\xi = 0$ (i.e., when the observer has a strong response bias in the direction of never giving RF responses upon S judgments) Eq. 10 reduces to Eq. 6c and, thus, the psychometric function for RF responses in the TOJ task is identical to that for RF responses in the SJ3 task. Alternatively, when $\xi = 1$ (i.e., when the observer has a strong bias in the direction of always giving RF responses upon S judgments) Eq. 10 reduces to the sum of Eqs. 6b and 6c and, thus, the psychometric function for RF responses in the TOJ task is identical to the sum of the psychometric functions for RF and S responses in the SJ3 task. The psychometric functions shown in Fig. 12.1c for the TOJ task arise from Eq. 10 also without response errors and with parameter values as before ($1/\lambda_t = 40$, $1/\lambda_r = 90$, $\tau = 60$, and $\delta = 100$); additionally, $\xi = 0.85$ for the continuous curve whereas $\xi = 0.15$ for the dashed curve. For cases in which $\xi = 0.5$, see Fig. 12.4 below.

Error parameters also have an effect on the final shape of the psychometric functions under all tasks, as does the response bias parameter ξ in TOJ tasks. These effects will be described in the next section. Estimating the response bias parameter ξ that operates in TOJ tasks is crucial for an understanding of some nagging empirical results; in contrast, estimating error parameters (the ε s and κ s in Eqs. 8–10 above) is theoretically uninteresting but useful to remove contamination affecting estimates of the timing and decisional parameters described above.

2.4 *Overall Effects on the Shape of the Psychometric Function*

In an empirical study, the observed psychometric function reflects the joint influence of all the components just discussed. They affect the overall shape, location, and slope of the psychometric function so that changes in any of these characteristics across experimental conditions or across groups of observers cannot be arbitrarily attributed to any of the components. Recourse to model-based psychometric functions for the analysis of data is needed to separate out those influences. We mentioned earlier that this is not achievable in full when the data are collected with single-presentation methods because the task confounds some of the parameters of interest. Ways around this problem will be discussed later in this chapter, but it is useful at this point to consider several scenarios that illustrate the way in which all parameters contribute to shaping the observed psychometric functions. We will leave response errors aside in this presentation, which generally affect only the asymptotes of the psychometric functions (for a discussion and illustration of them in the context of SJ3 tasks, see García-Pérez & Alcalá-Quintana, 2012b).

Figure 12.4 shows the psychometric functions that arise under two different forms for the relative distributions of arrival times of test and reference stimuli (top row) and two different widths for the symmetric central region in

FIGURE 12.4 *Psychometric functions in SJ2, SJ3, and TOJ tasks in four different scenarios resulting from the combination of two cases for the distributions of arrival times (top row) and two cases for the width of the (symmetric) central region in decision space (left column). In each of the other four panels, solid red, black, and blue curves are the psychometric functions for TF, S, and RF responses, respectively, in the SJ3 task (see legend in the bottom-left panel). The solid black curve is also the psychometric function for S responses that would be observed in an SJ2 task. The solid blue curve is also the psychometric function for RF responses that would be observed in a TOJ task if $\xi = 0$; the dashed and dotted blue curves are those that would be observed in the TOJ task if $\xi = 0.5$ or $\xi = 1$, respectively.*

decision space (left column). Consider first the top-right panel in the 2×2 array of psychometric functions, for the case in which the arrival-time distributions for reference and test stimuli differ in offset (parameters τ_r and τ_t) but not in spread (parameters λ_r and λ_t). The psychometric function for S responses in SJ2 or SJ3 tasks (black curve) is symmetric because $\lambda_r = \lambda_t$, and peaks at $\tau_r - \tau_t$. In the panel underneath, when the central region in decision space is narrower, the psychometric function for S responses keeps these characteristics but it is narrower and shorter as a result of higher resolution to judge temporal order. In both cases, the psychometric function for RF responses in the TOJ task can span a broad range of shapes and locations due to the response-bias parameter ξ (blue curves), which decouples observed performance in SJ and TOJ tasks even when timing and decisional parameters are identical.

By comparison, $\lambda_r \neq \lambda_t$ (panels on the left side of the 2×2 array) renders asymmetric functions that are analogously affected by the width of the central region and by the mismatch between τ_r and τ_t . However, note that the peak of the psychometric function for S responses in SJ2 and SJ3 tasks (black curves) does not occur at $\tau_r - \tau_t$ in these conditions.

Given their multiple determinants, PSSs or JNDs are not dependable for a characterization of timing processes. The data ask instead for an alternative analysis that can separate out these influences. Fitting model-based psychometric functions in which these influences are captured by distinct parameters thus allows proper inferences about how experimental manipulations or group membership affect timing, decisional, or response processes. The next section summarizes the empirical evidence supporting this model of performance in timing tasks.

3 Empirical Evidence Supporting the Model

A theoretical model allows extracting latent aspects that are not directly accessible via PSSs, SRs, JNDs, or other indices based on observed performance. However, these benefits can only be gained if the model offers a satisfactory account of empirical data, not only in terms of fitting data adequately but also, and more importantly, when model predictions are supported by the data. If this is the case, estimated model parameters can then be subsequently analyzed in search for an interpretation of experimental outcomes in terms of timing, decisional, and response processes.

Testing the adequacy of a model requires assessing its success at fitting empirical data. Although a model can never be proven correct, one expects an adequate model to pass goodness-of-fit tests the stated percentage of times (García-Pérez, 2017): At, say, the 5% significance level, an adequate model will be rejected in about 5% of the occasions; a meaningfully larger number of rejections indicates inadequacy. The model gains further support under stringent goodness-of-fit tests involving predictions to the effect that some parameters must remain invariant across experimental manipulations.

This model has been subjected to the latter type of test extensively, in the two forms described next. Firstly, recall that timing parameters (i.e., λ_r , λ_t , τ_r , and τ_t) reflect delays in capture and transmission of sensory information that should not be affected by which question the observer responds to at the end of each trial (i.e., whether data are collected with SJ2, SJ3, or TOJ tasks). These parameters should thus be invariant across tasks, whereas response and decisional parameters would reasonably vary across tasks. If the model is adequate,

common values for these timing parameters should provide a satisfactory account of SJ and TOJ data collected in within-subjects studies under otherwise identical experimental conditions. This type of data have been reported in a number of independent studies (e.g., Capa et al., 2014; Fujisaki & Nishida, 2009; Li & Cai, 2014; Linares & Holcombe, 2014; Matthews & Welch, 2015; Sanders et al., 2011; Schneider & Bavelier, 2003; van Eijk et al., 2008). An analysis of the 455 data sets from those studies supported the expectation of common timing parameters across tasks: The model including common timing parameters for all tasks was rejected in 24 (5.27%) of the occasions at the 5% significance level (see García-Pérez & Alcalá-Quintana, 2012b, 2015a, 2015b).

Secondly, because the model includes parameters that separately characterize the distribution of arrival latencies for test and reference stimuli, manipulations that affect the sensory processing of one stimulus but not the other should result in data that can be accounted for with common parameter values for the non-manipulated stimulus along with parameter values that vary for the other stimulus across the conditions in which it is manipulated. The analysis of SJ2 data on asynchronous audiovisual speech in an experiment in which only the visual stimulus was manipulated in four different ways (Magnotti, Ma, & Beauchamp, 2013) supported this prediction: Fitted under the stated constraint, the model was rejected at the 5% significance level for only 1 of 16 observers (6.25%; see García-Pérez & Alcalá-Quintana, 2015c).

In all of the analyses just mentioned, model psychometric functions followed very closely the path of empirical data to capture characteristics that conventional psychometric functions (i.e., arbitrary logistic or Gaussian functions) could not accommodate. These include asymmetries in SJ2 and TOJ data, relatively broad plateaus in SJ2 data, and intermediate regions of reduced slope in TOJ data. This ability to account for subtle features of the data provides additional qualitative support for the model and indicates that observed performance reflects the interplay of timing, decisional, and response processes. As discussed in the next section, summary performance measures such as PSSs, SRs, or JNDs are insufficient (and misleading) to identify relevant differences in timing processes across groups or experimental conditions.

4 PSSs, JNDs, and SRs vs. Interpretation of Model Parameters

The following discussion will leave the TOJ task aside. As shown in Fig. 12.4, the irrelevant response-bias parameter ξ strongly affects the shape of the observed psychometric function so as to make PSSs and JNDs uninformative and uninterpretable. To keep things simple, the discussion will focus on the SJ2

task although consideration of the SJ3 task yields analogous outcomes. Also, we will consider the SR instead of the JND.

The top row in Fig. 12.5 shows two scenarios regarding timing processes, with latency distributions that are in both cases identical for test and reference stimuli except for a shorter average latency for the reference stimulus. Compared to the left side, the scenario on the right side depicts latency distributions that are narrower and closer to stimulus onset, as one might expect in an experimental condition in which (or in a group of observers for whom) latencies are shorter and subject to less variability. Thus, on the left side, arrival latencies for reference and test stimuli have means $1/\lambda_r + \tau_r = 100$ and $1/\lambda_t + \tau_t = 140$, respectively, whereas their standard deviations are identically valued at $1/\lambda_r = 1/\lambda_t = 50$; on the right side, arrival latencies have instead means $1/\lambda_r + \tau_r = 45$ and $1/\lambda_t + \tau_t = 85$ and standard deviations $1/\lambda_r = 1/\lambda_t = 25$. One would generally like to know whether differences between groups or experimental conditions occur at the level of timing processes (i.e., their speed and variability), and these parameters convey such information.

These two scenarios regarding timing processes can be combined with the two different scenarios regarding decisional processes that are illustrated on the left of Fig. 12.5. At the top, the central region for simultaneity judgments is broader than it is at the bottom (i.e., $\delta = 100$ at the top, compared to $\delta = 50$ at the bottom). A broader central region implies that observers need larger differences in arrival time to identify temporal order (i.e., they have less ability to tell small differences in arrival time). This is also another characteristic that may differentiate groups or experimental conditions, and one that a researcher would like to know about.

Because these aspects are captured by model parameters (i.e., δ , the λ s, and the τ s), fitting model-based psychometric functions to estimate them provides all the necessary information for inferences about the timing and decisional components of observed performance, and also about how they vary across groups or experimental conditions (for detailed examples, see García-Pérez & Alcalá-Quintana, 2015b, 2015c). Fitting the model and obtaining parameter estimates is straightforward with the software described in Alcalá-Quintana and García-Pérez (2013).

This method of analysis is in sharp contrast with the routine calculation of PSSs and SRs (or JNDs), whose values are printed in each panel in the 2×2 array of psychometric functions at the bottom right of Fig. 12.5. Note that the PSS is entirely immune to these differences in timing and decisional processes, sitting at an SOA of -40 ms in all cases. In turn, the SR is also almost identical across differences in timing processes (center and right columns) and it is only slightly affected by differences in decisional processes (center and bottom rows).

FIGURE 12.5 PSSs and SRs (or JNDs) are insufficient to assess performance differences across groups or experimental conditions. The 2x2 array of psychometric functions at the bottom right arise from the combination of two scenarios regarding timing processes (top row) and two scenarios regarding decisional processes (left column). Each panel shows the resultant psychometric function in an S/I2 task along with the value of the PSS and the SR in each case. Layout as in Fig. 12-4.

PSSs and SRs cannot portray differences in the underlying processes. Note that fitting arbitrary psychometric functions in these conditions will also not do justice to the data. The scaled Gaussian that is often fitted to SJ2 data cannot approximate the shapes described by the psychometric functions in Fig. 12.5, and their parameters cannot be referred to underlying processes either.

5 Shortcomings, Variants, and Extensions

This section discusses two variants of timing tasks that are useful for addressing some of the issues that arise in the use of single-presentation methods, namely, the likely presence of a fourth type of judgment and the problems arising from an inescapable confound.

5.1 *The SJ4 Task*

As presented thus far, the model assumes that only three judgments are possible: stimulus A subjectively first, stimulus B subjectively first, or A and B subjectively simultaneous. This assumption embodies the theoretical position that perception of non-simultaneity is necessary and sufficient for perception of temporal order: If observers perceive asynchrony, they also identify temporal order (Allan & Kristofferson, 1974; Baron, 1969). An alternative view (Hirsh & Sherrick, 1961) is that perception of non-simultaneity is a necessary but insufficient condition for perception of temporal order: Perception of asynchrony may still not allow observers to identify temporal order. This latter stance assumes a fourth type of judgment whose existence has been disputed for decades. There is, however, direct and indirect evidence of its presence.

Direct evidence comes from the only (to our knowledge) empirical study in which an SJ4 task was used to allow observers to report this fourth judgment (Weiß & Scharlau, 2011). Although this judgment was never reported by some of their observers, aggregated data in the two experimental conditions of their experiment 1 (see their figure 5) revealed that these judgments are maximally prevalent at SOAs around either end of the SR, with a notch around the point where simultaneous responses are maximally prevalent. These data suggest that the fourth judgment is associated with latency differences that are sufficiently large to judge non-simultaneity but not enough to identify temporal order. In other words, instead of the three regions in which decision space was partitioned in the model, extra regions flanking the central region for S judgments seem necessary. Figure 12.6 shows an extension of the model along these lines and illustrates how the probability of each judgment varies with SOA, namely,

FIGURE 12.6 *Model extension to cover the SJ4 task. The decision space (center panel) includes two regions for “uncertain order” (U) judgments flanking the region for S judgments, besides the outer regions for TF and RF judgments. Note that the extra regions may have different widths. The right panel shows the resultant psychometric function for each judgment category (colored as their labels are in the center panel) under the sample arrival-time distributions for test and reference in the left panel.*

$$p_{\text{TF}}(\Delta t) = F(\delta_1; \Delta t) \quad (11a)$$

$$p_{\text{S}}(\Delta t) = F(\delta_3; \Delta t) - F(\delta_2; \Delta t) \quad (11b)$$

$$p_{\text{RF}}(\Delta t) = 1 - F(\delta_4; \Delta t) \quad (11c)$$

$$p_{\text{U}}(\Delta t) = F(\delta_4; \Delta t) - F(\delta_3; \Delta t) + F(\delta_2; \Delta t) - F(\delta_1; \Delta t) \quad (11d)$$

where U stands for “uncertain about order, though not simultaneous”. Error parameters can also be introduced in this model.

Indirect evidence on the U judgment also exists that comes from the joint analysis of SJ2 and TOJ data in within-subjects studies. García-Pérez and Alcalá-Quintana (2015a, 2015b) reported substantial evidence to the effect that the central region in decision space is broader in TOJ tasks than it is in SJ2 tasks. This result can easily be interpreted under the partition illustrated in Fig. 12.6. In SJ2 tasks, only the boundaries δ_2 and δ_3 are operative because arrival-time differences within $[\delta_2, \delta_3]$ render S judgments and all others render non-simultaneous judgments (whether or not temporal order was additionally identified). In contrast, only δ_1 and δ_4 are operative in TOJ tasks because arrival-time differences lower than δ_1 render TF judgments, arrival-time differences greater than δ_4 render RF judgments, and all others force observers to misreport their inability to judge temporal order (whether they perceived non-simultaneity or simultaneity in such cases). Because $\delta_4 - \delta_1 \geq \delta_3 - \delta_2$, the empirical observation of a broader central region in TOJ tasks is consistent with the existence of the U judgment.

To look further into this issue, we conducted a study using the SJ4 task with identical visual stimuli (Gabor patches) presented on the left and on the right of a central fixation point on a monitor running at 60 Hz so that SOAs varied in steps of 16.7 ms. The patch on the right was regarded as the reference and data were collected in three consecutive 192-trial sessions with SOAs determined by an adaptive procedure (García-Pérez, 2014). On each trial, left and right stimuli were abruptly presented with the prescribed SOA and they were both removed simultaneously 750 ms after the onset of the stimulus presented second. Observers then used a keyboard to enter the response that described their judgment. Figure 12.7 shows data and fitted psychometric functions (incorporating error parameters) for each of 19 observers, plus a summary panel.¹

Most notably, fitted curves follow the path of the data very closely in all cases. At the 5% significance level, the G^2 statistic rejected the model for observer #3 only, but even in that case the agreement between data and fitted curves is visibly good. In general, U responses occurred at the small SOAs expected under the model, with an imbalance at negative vs. positive SOAs that is consistent with the allowance for decision ranges of different width at negative vs. positive arrival-time differences (see Fig. 12.6). Some observers gave U responses relatively frequently (e.g., #1 and #5). Other observers gave U responses sparingly (e.g., #9 and #14) or not at all (e.g., #2 and #7). Interestingly, the latter observers' data show bumps of "improper" temporal-order responses (i.e., RF responses at negative SOAs and TF responses at positive SOAs) where U responses would be expected (see red and blue data points and curves at small positive and negative SOAs for observers #2, #6, #7, #8, #9, #10, #12, #13, #14, #15, #18, and #19). It is unclear whether these bumps reflect misreports (i.e., temporal-order responses upon U judgments) or authentic reversals of subjective temporal order, but their presence here is consistent with previous evidence to the same effect, as discussed next.

Firstly, the SJ3 data of van Eijk et al. (2008) showed bumps of improper temporal-order responses that are also found in earlier data sets and which prompted Ulrich (1987) to dismiss independent-channels models. Yet, an analysis under our SJ3 model showed that misreports account for this empirical feature within the realm of independent-channels models (García-Pérez & Alcalá-Quintana, 2012b). The present SJ4 data suggest that U judgments—which must be misreported as TF or RF responses in SJ3 tasks—also contribute to these bumps.

1 The data, the MATLAB code used to estimate model parameters and assess goodness of fit, and other related Online Material are available in the book's GitHub repository.

FIGURE 12.7 *SJ4 data and fitted psychometric functions. The grayed panel at the bottom right shows aggregated data and averaged psychometric functions across observers, and note that U responses in that panel display a pattern analogous to that reported by Weiß and Scharlau (2011) also for their aggregated data. Color code as in Fig. 12.6.*

Secondly, improper temporal-order responses at the SOAs where U responses are expected (or given) may indicate authentic reversals of perceived temporal order, not guesses upon U judgments. The cause of such reversals is elusive but they have been reported in TOJ tasks involving tactile stimuli with arms crossed, resulting in N-shaped psychometric functions that are at odds with the sigmoidal shapes obtained without arm crossing and with the comparatively narrow and peaked shape of SJ2 data with or without arm crossing (see Cadieux, Barnett-Cowan, & Shore, 2010; Fujisaki & Nishida, 2009; Yamamoto & Kitazawa, 2001; see also Heed & Azañón, 2014).

We assessed whether N-shaped TOJ data could arise from the presence of a region for U judgments in decision space that, by a mechanism still to be unraveled, yields reversals of perceived temporal order. For this purpose, we derived a theoretical TOJ curve for each observer in Fig. 12.7 by making reasonable (but speculative) assumptions about how they would have responded in a TOJ task. Specifically, RF and TF responses in the SJ4 task would directly transfer into the same responses in the TOJ task, S responses in the SJ4 task would be evenly split into TF and RF responses in the TOJ task, and U responses in the SJ4 task would be split according to the imbalance of TF and RF responses at each SOA in the SJ4 task. If Ψ_{TF} , Ψ_S , Ψ_{RF} , and Ψ_U are the psychometric functions for TF, S, RF, and U responses in Fig. 12.7, the psychometric function for RF responses in the TOJ task would thus be $\Psi_{RF}^{(TOJ)}(\Delta t) = \Psi_{RF}(\Delta t) + \Psi_S(\Delta t)/2 + \Psi_U(\Delta t) \times \Psi_{RF}(\Delta t)/[\Psi_{RF}(\Delta t) + \Psi_{TF}(\Delta t)]$. The results (not shown) revealed a diversity of patterns similar to that reported by Cadieux et al. (2010, figure 2), including N-shaped functions. It is still unclear why reversals of perceived temporal order occur, why they seem absent in some observers, or why their prevalence varies across experimental conditions, but their relation to intermediate regions for U judgments in decision space makes the SJ4 task a useful tool to investigate these issues.

It is interesting to note that observer #11 gave very few U responses and did not give any S response. One might surmise that this arises from an exquisite ability to perceive temporal order and a lack of S and U regions in decision space (i.e., $\delta_1 \approx \delta_2 \approx \delta_3 \approx \delta_4$). However, the path of TF and RF responses speaks against this surmise and suggests instead that observer #11 approached the SJ4 task essentially as a TOJ task in which S and U judgments are instead misreported as TF or RF responses.

During debriefing, some observers reported effortless identification that the left stimulus had been presented first, even if only by a very short time. These spontaneous reports are consistent with the outcome (see the Online Material) that the interval $[\delta_1, \delta_2]$ was generally estimated to be narrower than the interval $[\delta_3, \delta_4]$, as illustrated in Fig. 12.6. Because the sequential display of two visual stimuli at nearby spatial positions induces beta apparent movement (Larsen, Farrel, & Bundesen, 1983), this outcome suggests a predisposition to perceive rightward beta movement, with (potential) leftward beta movement giving instead an impression of non-simultaneous presentations without clear identification that the right stimulus was presented first. This is speculative at present, but the SJ4 task also proves useful tool to investigate this issue.

It is important to stress that the SJ4 model was fitted with allowance for different arrival-time distributions of test and reference stimuli (i.e., $\lambda_r \neq \lambda_t$

and $\tau \neq 0$). Because of the confound mentioned earlier and further discussed in the next section, this fitting approach requires the constraint that $\delta_2 = -\delta_3$. Estimates of λ_r and λ_t nevertheless turned out to be very similar on a subject-by-subject basis (see the Online Material), which seems reasonable in retrospect given that (1) reference and test stimuli were identical Gabor patches differing only in their location in the visual field and (2) location in the visual field does not affect the processing speed of visual stimuli (García-Pérez & Alcalá-Quintana, 2015b). An alternative approach to fitting these data under the constraints that $\lambda_r = \lambda_t$ and $\tau_t = \tau_r$ is discussed in the next subsection.

5.2 *Decisional Bias and the Dual-presentation Task*

We mentioned that single-presentation methods always confound several parameters in any model of psychophysical performance. Unfortunately, they always confound sensory parameters with decisional parameters, thus precluding the characterization of timing processes and the identification of the cause of observed differences across conditions. Then, the confound affects studies aimed at assessing *prior entry* (a hypothetical sensory acceleration that hastens the processing of attended stimuli) or *temporal recalibration* (an adjustment of subjective synchrony due to prolonged exposure to asynchronous stimulation). To test these hypotheses, one needs to tell whether observed differences across conditions are due to timing or to decisional processes, something that turns out to be impossible with single-presentation methods. Figure 12.8 illustrates this confound for an SJ3 task in a hypothetical experiment on prior entry with audiovisual stimuli. The column labeled “natural processing” reflects a condition in which neither stimulus is favored so that visual and auditory arrival times have their natural distributions; the column labeled “sensory (visual) acceleration” shows what visual arrival times might be like if processing of the visual signal is speeded up: Arrival times advance, as determined by $\tau_v = 20$ compared to $\tau_v = 60$ in the former case.

The panels underneath show the psychometric functions that would be obtained in each condition. Whether the central region in decision space is centered (middle row) or displaced (bottom row), visual acceleration displaces the psychometric functions to the left. However, the experimental manipulation presumably inducing visual acceleration might instead induce only a decisional bias. Then, compared to the psychometric functions at the top-left of the 2×2 array, decisional bias will also produce a leftward shift of the psychometric functions (bottom-left panel). Then, sensory acceleration without decisional bias (top-right panel, grayed) produces the same observable effect as decisional bias without sensory acceleration (bottom-left panel, grayed). Given that single-presentation methods confound sensory and decisional processes

FIGURE 12.8 *Single-presentation methods cannot distinguish decisional bias from sensory acceleration. The grayed panels in the 2x2 array depict identical psychometric functions that can result either from pure sensory acceleration (top-right panel) or from pure decisional bias (bottom-left panel). Layout as in Fig. 12.4.*

and do not allow telling apart their respective influences, the prior-entry (or temporal-recalibration) hypotheses cannot be tested with them.

Methodological confounds that preclude testing for prior entry with single-presentation methods were pointed out by Spence and Parise (2010) but ways around this problem have not been devised thus far. We will show here that a dual-presentation method and a ternary response format allow separating out sensory and decisional contributions to observed performance (see also García-Pérez & Alcalá-Quintana, 2013).

Before discussing this issue, we should note that single-presentation methods are still useful when SOAs are delivered with stimuli that can be reasonably assumed to render identical arrival-time distributions. In these conditions, $\lambda_r = \lambda_t = \lambda$ and $\tau_t = \tau_r$ so that $\tau = 0$. Because the confound affects parameters τ , δ_1 , and δ_2 , the constraint $\tau = 0$ resolves it and allows estimating decisional bias. García-Pérez and Alcalá-Quintana (2015a, 2015b) illustrated these particular circumstances in several empirical cases in which test and reference stimuli were identical visual patterns presented at different spatial locations

on a monitor, and once it had been established that position in the visual field does not alter the processing speed of a visual stimulus. These conditions hold also for the SJ4 data presented in the previous subsection, something that allows an analysis without the constraint $\delta_2 = -\delta_3$ and permits an assessment of decisional bias. Presentation of the results of this alternative approach to fitting the above SJ4 data is deferred to the Online Material.

Nevertheless, this approach is inappropriate in prior entry studies in which the issue under investigation is whether attentional manipulations alter the distribution of arrival times and, thus, whether $\lambda_r \neq \lambda_t$ and/or $\tau_t \neq \tau_r$ even when test and reference stimuli are otherwise identical. In these conditions, separating out decisional and timing influences on observed performance is crucial. This can only be accomplished using a dual-presentation method coupled with a ternary response format. In this variant of the task, each trial presents two SOAs sequentially and the observer reports whether presentations were subjectively more synchronous in the first interval, in the second, or they were instead indistinguishable as to (a)synchrony. One of the intervals in each trial displays the standard SOA, which is not necessarily synchronous but has the same magnitude in all trials;² the other interval displays a test SOA whose magnitude varies across trials. Standard and test SOAs are also presented in both possible orders across trials. The observer's responses are then tallied to compute the proportion of trials in which each judgment (test more synchronous, standard more synchronous, or standard and test indistinguishable as to synchrony) was reported under each order of presentation (test SOA presented first or test SOA presented second) at each test SOA.

The model for this dual-presentation task is a straightforward extension of the model described earlier for single-presentation tasks, although derivation of model psychometric functions is somewhat more elaborate. A manuscript in preparation will present such extension in detail, including an empirical test of the validity of the model. Here we will only describe the model succinctly with an eye to illustrating how this task solves the inescapable problems of single-presentation methods and their inappropriateness for testing the prior-entry or temporal-recalibration hypotheses. In a nutshell, the model assumes that observers gather arrival-time differences D_1 and D_2 in each interval on each trial, whose individual distributions are given by Eq. 5 above for the

2 Sets of trials involving several standard SOAs can be interwoven in a session but the subsequent analysis is conducted separately for each standard SOA, just as if each of them had been used in a separate session (e.g., Allan & Kristofferson, 1974; Pastore & Farrington, 1996; Yarrow et al., 2016).

SOA used in each interval. Because observers are asked to report the interval in which presentation was subjectively more synchronous, the decision variable is the difference of *perceived offsets* (unsigned perceived asynchronies) $D = |D_2| - |D_1|$, whose probability distribution exists in closed form. The decision space also includes three regions so that the observer judges the first (alternatively, second) interval to be more synchronous when $D > \delta_2$ (alternatively, $D < \delta_1$) and judges both intervals to be indistinguishably (a)synchronous when $\delta_1 \leq D \leq \delta_2$. Leaving aside the extension that incorporates parameters for response errors, model psychometric functions for this task are shown in Fig. 12.9 under the same scenarios used in Fig. 12.8 to illustrate the inappropriateness of single-presentation methods.

It is immediately obvious by visual inspection of the 2×2 array of panels at the bottom right of Fig. 12.9 that lack of decisional bias results in psychometric functions that superimpose for both orders of presentation of test and standard SOA in each trial (middle row) whereas decisional bias renders psychometric functions that differ across presentation orders (bottom row). In either case, sensory acceleration (right column) produces psychometric functions that are rigidly shifted laterally compared to their location in the absence of sensory acceleration (middle column), regardless of whether or not the baseline condition involves decisional bias. In other words, sensory acceleration and decisional bias produce distinct effects on the observed psychometric functions, and these effects are captured by model parameters.

Then, the dual-presentation task with a ternary response format allows separating out timing and decisional processes while entirely removing the response bias that might also contaminate the data if observers were forced to report one of the intervals as more synchronous when they actually judge them to be equally (a)synchronous. This task thus allows estimating all the relevant parameters (λ_r , λ_v , τ , δ_1 , and δ_2) that are needed to express observed performance in terms of timing and decisional processes. Hence, prior-entry and temporal-recalibration hypotheses can be tested with this task. We should stress that timing and decisional components cannot be separated out if a binary response format is used instead (i.e., if the option to report indistinguishability is not given) or if responses are aggregated across orders of presentation of standard and test SOA.

6 Conclusion

Research on timing requires collecting data on observers' performance. Such data inform of the speed of sensory processing for each of the stimuli used to

deliver SOAs under the selected experimental conditions, but observed performance is also modulated by decisional and response processes. An analysis that separates the contribution of the three components of performance is needed for a proper assessment of timing processes. We have shown that a model-based analysis is useful for this purpose and that classical measures of observed performance (PSSs, JNDs, or SRs) mix up these contributions misleadingly.

We have also shown that the single-presentation tasks most often used in empirical studies (SJ2, SJ3, and TOJ tasks) confound timing and decisional components, lending them unsuitable for studies in which differences in

FIGURE 12.9 *Distinguishing decisional bias from sensory acceleration with a dual-presentation task and a ternary response format. Without loss of generality, the standard SOA is assumed to be a synchronous presentation of visual and auditory stimuli whose individual arrival-time distributions are indicated in the top row in the baseline (normal) condition and under an experimental manipulation that speeds up processing of the visual stimulus. The left column shows the decision space with boundaries that incorporate decisional bias (bottom row) or lack thereof (middle row). The 2×2 array of panels at the bottom right shows that sensory acceleration and decisional bias have distinguishable effects on the observed psychometric functions.*

timing processes across conditions are under scrutiny (e.g., prior entry) or for studies in which differences in decisional processes also across conditions are under scrutiny (e.g., temporal recalibration).

The above does not mean that single-presentation tasks are useless. In a number of situations, research interests focus on model parameters not affected by the confound (e.g., the variability of arrival latencies), experimental conditions constrain performance in a way that the confound is bypassed (e.g., test and reference stimuli are identical in all respects), or interest lies in qualitative aspects (e.g., whether the presence of U judgments warrants the use of an SJ4 task). In such cases, single-presentation tasks are useful, but model-based analyses are still needed to extract all the information that the data can provide. The software included in the Online Material accompanying this chapter (for model-based analyses of SJ4 data) supplements the software in Alcalá-Quintana and García-Pérez (2013; for model-based analysis of SJ2, SJ3, and TOJ data) to facilitate this task.

Acknowledgements

This work was supported by grants PSI2012-32903 and PSI2015-67162-P from Ministerio de Economía y Competitividad (Spain). Parts of the computations were carried out on EOLO, the MEC-D- and MICINN-funded HPC for climate change at Moncloa Campus of International Excellence, Universidad Complutense.

References

- Alcalá-Quintana, R., & M.A. García-Pérez (2013). Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. *Behavior Research Methods*, 45, 972–998.
- Allan, L.G., & A.B. Kristofferson (1974). Successiveness discrimination: Two models. *Perception & Psychophysics*, 15, 37–46.
- Baron, J. (1969). Temporal roc curves and the psychological moment. *Psychonomic Science*, 15, 299–300.
- Bedard, G., & M. Barnett-Cowan (2016). Impaired timing of audiovisual events in the elderly. *Experimental Brain Research*, 234, 331–340.
- Binder, M. (2015). Neural correlates of audiovisual temporal processing: Comparison of temporal order and simultaneity judgments. *Neuroscience*, 300, 432–447.
- Cadieux, M.L., M. Barnett-Cowan, & D.I. Shore (2010). Crossing the hands is more confusing for females than males. *Experimental Brain Research*, 204, 431–446.

- Capa, R.L., C.Z. Duval, D. Blaison, & A. Giersch (2014). Patients with schizophrenia selectively impaired in temporal order judgments. *Schizophrenia Research*, 156, 51–55.
- Colonius, H., & A. Diederich (2011). Computing an optimal time window of audiovisual integration in focused attention tasks: Illustrated by studies on effect of age and prior knowledge. *Experimental Brain Research*, 212, 327–337.
- Donohue, S.E., M.G. Woldorff, & S.R. Mitroff (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics*, 72, 1120–1129.
- Fouriez, G., G. Capstick, F. Monette, C. Bellemare, M. Parkinson, & A. Dumoulin (2007). Judgments of synchrony between auditory and moving visual stimuli. *Canadian Journal of Experimental Psychology*, 61, 277–292.
- Fujisaki, W., & S. Nishida (2009). Audio-tactile superiority over visuo-tactile and audiovisual combinations in the temporal resolution of synchrony perception. *Experimental Brain Research*, 198, 245–259.
- García-Pérez, M.A. (2014). Adaptive psychophysical methods for nonmonotonic psychometric functions. *Attention, Perception, & Psychophysics*, 76, 621–641.
- García-Pérez, M.A. (2017). Thou shalt not bear false witness against null hypothesis significance testing. *Educational and Psychological Measurement*, 77(4), 631–662.
- García-Pérez, M.A., & R. Alcalá-Quintana (2012a). On the discrepant results in synchrony judgment and temporal-order judgment tasks: A quantitative model. *Psychonomic Bulletin & Review*, 19, 820–846.
- García-Pérez, M.A., & R. Alcalá-Quintana (2012b). Response errors explain the failure of independent-channels models of perception of temporal order. *Frontiers in Psychology*, 3, 94.
- García-Pérez, M.A., & R. Alcalá-Quintana (2013). Shifts of the psychometric function: Distinguishing bias from perceptual effects. *Quarterly Journal of Experimental Psychology*, 66, 319–337.
- García-Pérez, M.A., & R. Alcalá-Quintana (2015a). Converging evidence that common timing processes underlie temporal-order and simultaneity judgments: A model-based analysis. *Attention, Perception, & Psychophysics*, 77, 1750–1766.
- García-Pérez, M.A., & R. Alcalá-Quintana (2015b). The left visual field attentional advantage: No evidence of different speeds of processing across visual hemifields. *Consciousness and Cognition*, 37, 16–26.
- García-Pérez, M.A., & R. Alcalá-Quintana (2015c). Visual and auditory components in the perception of asynchronous audiovisual speech. *i-Perception*, 6(6), 1–20.
- Grant, K.W., V. van Wassenhove, & D. Poeppel (2004). Detection of auditory (cross-spectral) and auditory-visual (cross-modal) synchrony. *Speech Communication*, 44, 43–53.
- Heath, R.A. (1984). Response time and temporal order judgement in vision. *Australian Journal of Psychology*, 36, 21–34.

- Heed, T., & E. Azañón (2014). Using time to investigate space: A review of tactile temporal order judgments as a window onto spatial processing in touch. *Frontiers in Psychology*, 5, 76.
- Hillenbrand, J. (1984). Perception of sine-wave analogs of voice onset time stimuli. *Journal of the Acoustical Society of America*, 75, 231–240.
- Hirsh, I.J., & C.E. Sherrick (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology*, 62, 423–432.
- Larsen, A., J.E. Farrell, & C. Bundesen (1983). Short- and long-range processes in visual apparent movement. *Psychological Research*, 45, 11–18.
- Li, S.-X., & Y.-C. Cai (2014). The effect of numerical magnitude on the perceptual processing speed of a digit. *Journal of Vision*, 14(12), 1–9.
- Lieberman, A.M., K.S. Harris, J.A. Kinney, & H. Lane (1961). The discrimination of relative onset-time of the components of certain speech and nonspeech patterns. *Journal of Experimental Psychology*, 61, 379–388.
- Linares, D., & A.O. Holcombe (2014). Differences in perceptual latency estimated from judgments of temporal order, simultaneity and duration are inconsistent. *i-Perception*, 5, 559–571.
- Love, S.A., K. Petrini, A. Cheng, & F.E. Pollick (2013). A psychophysical investigation of differences between synchrony and temporal order judgments. *PLoS ONE*, 8(1), e54798.
- Magnotti, J.F., W.J. Ma, & M.S. Beauchamp (2013). Causal inference of asynchronous audiovisual speech. *Frontiers in Psychology*, 4, 798.
- Matthews, N., & L. Welch (2015). Left visual field attentional advantage in judging simultaneity and temporal order. *Journal of Vision*, 15(2), 1–13.
- McGrath, M., & Q. Summerfield (1985). Intermodal timing relations and audio-visual speech recognition by normal-hearing adults. *Journal of the Acoustical Society of America*, 77, 678–685.
- Pastore, R.E., & S.M. Farrington (1996). Measuring the difference limen for identification of order of onset for complex auditory stimuli. *Perception & Psychophysics*, 58, 510–526.
- Sanders, M.C., N.-Y.N. Chang, M.M. Hiss, R.M. Uchanski, & T.E. Hullar (2011). Temporal binding of auditory and rotational stimuli. *Experimental Brain Research*, 210, 539–547.
- Schneider, K.A., & D. Bavelier (2003). Components of visual prior entry. *Cognitive Psychology*, 47, 333–336.
- Schneider, K.A., & M. Komlos (2008). Attention biases decisions but does not alter appearance. *Journal of Vision*, 8(15), 1–10.
- Spence, C., & C. Parise (2010). Prior-entry: A review. *Consciousness and Cognition*, 19, 364–379.

- Sternberg, S., & R.L. Knoll (1973). The perception of temporal order: Fundamental issues and a general model. In Kornblum, S. (Ed.), *Attention and Performance IV* (pp. 629–685). New York: Academic Press.
- Stevenson, R.A., & M.T. Wallace (2013). Multisensory temporal integration: Task and stimulus dependencies. *Experimental Brain Research*, 227, 249–261.
- Ulrich, R. (1987). Threshold models of temporal-order judgments evaluated by a ternary response task. *Perception & Psychophysics*, 42, 224–239.
- van Eijk, R.L.J., A. Kohlrausch, J.F. Juola, & S. van de Par (2008). Audiovisual synchrony and temporal order judgments: Effects of experimental method and stimulus type. *Perception & Psychophysics*, 70, 955–968.
- van Eijk, R.L.J., A. Kohlrausch, J.F. Juola, & S. van de Par (2009). Temporal interval discrimination thresholds depend on perceived synchrony for audio-visual stimulus pairs. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 1254–1263.
- van de Par, S., & A. Kohlrausch (2000). Sensitivity to auditory-visual asynchrony and to jitter in auditory-visual timing. In Rogowitz, B.E. & T.N. Pappas (Eds.), *Proceedings of SPIE: Human vision and electronic imaging V* (Vol. 3959, pp. 234–242). Bellingham, WA: SPIE Press.
- Vatakis, A., J. Navarra, S. Soto-Faraco, & C. Spence (2008). Audiovisual temporal adaptation of speech: Temporal order versus simultaneity judgments. *Experimental Brain Research*, 185, 521–529.
- Weiß, K., & I. Scharlau (2011). Simultaneity and temporal order perception: Different sides of the same coin? Evidence from a visual prior-entry study. *Quarterly Journal of Experimental Psychology*, 64, 394–416.
- Yamamoto, S., & S. Kitazawa (2001). Reversal of subjective temporal order due to arm crossing. *Nature Neuroscience*, 4, 759–765.
- Yarrow, K., N. Jahn, S. Durant, & D.H. Arnold (2011). Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Consciousness and Cognition*, 20, 1518–1531.
- Yarrow, K., S.E. Martin, S. Di Costa, J.A. Solomon, & D.H. Arnold (2016). A roving dual-presentation simultaneity-judgment task to estimate the point of subjective simultaneity. *Frontiers in Psychology*, 7:416.