Relative timing: from behaviour to neurons

S. Mehdi Aghdaee1,2, Lorella Battelli3,4 and John A. Assad1,3

1Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA
2Vision Sciences Laboratory, Department of Psychology, Harvard University, Cambridge, MA 02138, USA
3Center for Neuroscience and Cognitive Systems@UniTn, Istituto Italiano di Tecnologia, Rovereto 38068, Italy
4Berenson – Allen Center for Noninvasive Brain Stimulation and Department of Neurology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215, USA

Processing of temporal information is critical to behaviour. Here, we review the phenomenology and mechanism of relative timing, ordinal comparisons between the timing of occurrence of events. Relative timing can be an implicit component of particular brain computations or can be an explicit, conscious judgement. Psychophysical measurements of explicit relative timing have revealed clues about the interaction of sensory signals in the brain as well as in the influence of internal states, such as attention, on those interactions. Evidence from human neurophysiological and functional imaging studies, neuropsychological examination in brain-lesioned patients, and temporary disruptive interventions such as transcranial magnetic stimulation (TMS), point to a role of the parietal cortex in relative timing. Relative timing has traditionally been modelled as a ‘race’ between competing neural signals. We propose an updated race process based on the integration of sensory evidence towards a decision threshold rather than simple signal propagation. The model suggests a general approach for identifying brain regions involved in relative timing, based on looking for trial-by-trial correlations between neural activity and temporal order judgements (TOJs). Finally, we show how the paradigm can be used to reveal signals related to TOJs in parietal cortex of monkeys trained in a TOJ task.

1. Timing and the brain

Timing is central for myriad aspects of behaviour. All forms of life generate behaviours that play out according to precise sequences in time, spanning multiple temporal scales. Complex behaviours such as locomotion or grasping and manipulation require coordinated temporal cascades of muscular activations over hundreds of milliseconds to seconds; repetitive movements such as sawing a plank or chopping an onion follow rhythms on the scale of seconds; and activities such as foraging [1] can be timed to anticipate food availability on the scale of hours to days.

Time is equally important for interpreting the world around us. ‘What’ and ‘where’ do not suffice; we also need to understand ‘when’, to intuit temporal relationships among sensory stimuli, and to take actions within optimal time windows to effect desired outcomes.

One can consider different temporal relations between events. Metrical judgements regarding the amount of time elapsed between events are called interval timing. Computing time intervals is inherent to behavioural conditioning [2,3] and provides the basis for rhythmic movements and behavioural cycles, ranging from subseconds to hours [4]. Interval timing has been intensively studied, and is the focus of many of the contributions in this volume.

The focus of this paper is instead on relative timing. Relative timing refers to the ordinal relationship between events, i.e. comparisons between instants of time. Relative timing is an essential, although often overlooked aspect of perception. For example, relative timing bears significance for perceptual grouping. In the visual domain, elements in the visual scene that change together over time, as well as space, are grouped together, and become part of the ‘same object’ [5].
The same grouping principles of Gestalt psychology apply to auditory scene analysis: sounds that start and stop together as well as those that co-modulate in amplitude or frequency can be grouped together and identified as distinct auditory objects [6]. Moreover, the sequential arrangement of speech syllables and the pauses between them, as well as the temporal structure within syllables and phonemes is fundamental for speech recognition [7].

In addition to interpreting our sensory environment, encoding relative timing is critical for higher cognitive functions. An important example is our ability to draw causal inferences. In assessing whether two events might be ‘cause’ and ‘effect’, the philosopher David Hume proposed eight criteria, among them contiguity in time and space and temporal precedence of cause over effect [8]. That is, if event B follows (or is perceived as following) event A, then it is plausible that A causes B; however, the reverse cannot be true. Thus, perceiving time and estimating the order of temporal events, as well as the errors in that estimation, influence how we establish causal inferences between events.

The purpose of this paper is to provide a brief, critical review of relative timing mechanisms. We authors are mainly neurophysiologists, and our primary motivation is to stimulate our fellow neurophysiologists to think about the topic of relative timing. For this reason (and for the sake of brevity), our review covers only selected aspects of relative timing. We focus on studies that examine relative timing within single sensory modalities, particularly vision, across different levels of investigation. For a more comprehensive review on the temporal organization of perception, see [9].

2. Implicit computations of relative timing

Relative timing computations in the brain could be further categorized as implicit, unconscious computations or as explicit judgements. In the case of implicit relative timing, the ordinal relationship between events is critical, but the events are not experienced as discrete, and the perceptual quality is typically distinct from that of the individual events. Typically, these phenomena require very brief temporal intervals between the underlying events, on the order of tens of milliseconds or less. An example is visual direction selectivity, a sensitive visual ability in vertebrate and invertebrate species, and a common attribute in single neurons in the visual system, such as the H1 neurons of the fly visual system [10], some ganglion cells in the mammalian retina [11] and many neurons in the middle temporal (MT) area of the primate visual cortex [12,13]. Two spatially displaced visual stimuli that are flashed in rapid succession are perceived as a unitary stimulus that moves from the location of the first stimulus towards the location of the second stimulus. The relative timing of the two stimuli is paramount, because if their timing is reversed, then the perceived direction is likewise reversed. At the level of single direction-selective neurons, if the two stimuli are spatially displaced along the preferred-null directional axis of the neuron, then the cell will respond much more strongly when one stimulus is presented before the other stimulus (preferred direction) than when the temporal order between the two stimuli is reversed (null direction) [11,12].

Implicit relative timing is also a prominent feature of auditory coding. Relative order is critical for recognition of complex sounds such as speech, and sensitivity to stimulus order is common at the level of single neurons in the auditory system. For example, some bat species use frequency-modulated (FM) ultrasonic vocalizations to determine spatial parameters of targets [14]. Neurons in the inferior colliculus of little brown bats are selective for FM sweep direction (rising or falling frequencies) [15,16]. Direction selectivity for FM tones is also present in the cat’s inferior colliculus [17] and auditory cortex [18]. Sound localization based on intra-aural time disparities is another prominent example of implicit relative timing. The relative time of arrival of sound at the two ears is translated into a horizontal location of the sound source: sounds that reach (or are presented to) the right ear before the left ear are perceived to originate to the right of the midline (and vice versa), with a temporal resolution on the order of microseconds [19]. With sufficiently brief inter-aural disparities, subjects detect a unitary sound source at a precise horizontal location; they do not detect distinct signals at the two ears, much less make an explicit judgement regarding the relative arrival times. At the cellular level, neurons in the mammalian superior olivary complex and avian nucleus laminaris detect coincident arrival of neuronal signals from the cochlear nuclei on either side of the brainstem, corresponding to precise inter-aural delays [19,20].

These examples of implicit timing typically depend on information that converges on single neurons and is integrated on the temporal scale of the membrane time constant (tens of milliseconds). Relative temporal encoding also extends to interactions between neurons. For example, spike-timing-dependent plasticity is a ubiquitous modulation of synaptic strength between pre- and postsynaptic neurons that depends critically on the precise order between pre- and postsynaptic electrical activity [21]. Long-term potentiation occurs when presynaptic spikes lead excitatory postsynaptic potentials and postsynaptic spikes by up to approximately 20 ms, whereas long-term depression results when postsynaptic spikes lead presynaptic spikes by approximately 20–100 ms [22,23]. Thus, relative order of neuronal activity among synaptic partners is a key determinant of whether synaptic efficacy increases or decreases, with no explicit representation of relative time.

3. Explicit computations of relative timing: temporal order judgements and simultaneity judgements

In contrast to implicit timing computations, in which events presented in close succession are rapidly integrated to yield a distinct percept, animals are also capable of perceiving multiple discrete events and judging their temporal order. Explicit temporal judgements are crucial to organizing the sensory scene and inferring causal relationships among events. The rest of this paper will focus on explicit relative timing.

An example of an explicit judgement of temporal order can be found in soccer. When a penalty kick is awarded, the referee must judge the timing of the goalkeeper’s forward movement relative to when the shooter’s foot strikes the ball (if the goalkeeper moves first, then the kick should be retaken). In experimental settings, judgements regarding the temporal relationship of two events can be categorized as which event occurred before the other, referred to as temporal order judgement (TOJ), or whether the two events occurred simultaneously or not, referred to as simultaneity judgement.
The double-headed arrow (blue) indicates the just notable difference (JND), subjective simultaneity (PSS), the SOA that corresponds to 50% on the ordinate. SOA is positive when stimulus A is presented first and negative when stimulus B appearing first is plotted as a function of SOA, defined as 'stimulus A lead time'.

**Figure 1.** Hypothetical psychometric function from a temporal order judgement (TOJ) experiment. Two stimuli (A and B) are presented with a range of stimulus onset asynchronies (SOA). Subject's probability of reporting stimulus A appearing first is plotted as a function of SOA, defined as 'stimulus A lead time'. SOA is positive when stimulus A is presented first and negative when stimulus B is presented first. The single-headed arrow (green) corresponds to the point of subjective simultaneity (PSS), the SOA that corresponds to 50% on the ordinate. The double-headed arrow (blue) indicates the just notable difference (JND), defined as half of the interquartile range. (Online version in colour.)

Historically, relative timing judgements were brought to attention from the field of astronomy in the early-nineteenth century, when astronomers estimated the transit time of stars using the ‘ear-and-eye method’. Astronomers measured distance using graticule lines superimposed on an image plane of the telescope, and measured time by listening to beats of a clock; they then matched which beat of the clock corresponded to the instant the star passed a particular graticule line, i.e. a TOJ between a visual and an auditory signal. Variations in these judgements among astronomers stimulated some of the first thinking about the perception of relative time [24], and were among the seminal observations that led Gustav Fechner and Wilhelm Wundt to launch the fields of psychophysics and experimental psychology [25–27].

TOJ and SJ are the means to study the instants in time that observers attribute to external events (‘when’). Comparing the ‘when’ of two events in relation with each other is more approachable than estimating the absolute time of occurrence of events, because relative timing does not require precise indexing of time. Accordingly, relative timing tasks have been used to study diverse phenomena, such as attention [28,29], perceptual latency [30,31], dependence of sensory latency on stimulus intensity [32] and speed–accuracy trade-off [33].

The general TOJ paradigm is as follows. Two events are presented with a time interval between them, called the stimulus onset asynchrony (SOA). The SOA is varied from trial to trial. The observer’s task is to report the stimulus that (perceptually) appeared first. The percentage of ‘reporting one of the stimuli appearing first’ is plotted as a function of SOA (figure 1). To quantify the observer’s performance in a TOJ task, a cumulative Gaussian function is typically fit to the data, which provides two parameters. One parameter is the point of subjective simultaneity (PSS), often taken as the SOA that corresponds to the 50% value on the ordinate. At the PSS, either choice is equally likely, and observers are maximally uncertain about the order [34]. The PSS does not necessarily correspond to the point of objective simultaneity, where SOA is zero; neither does an SOA = 0 necessarily yield a percept of simultaneity in observers. Thus, the PSS is an index of the bias (horizontal offset) of the psychometric function. The second parameter is the just notable difference (JND) or threshold, corresponding to half of the interquartile range [35]. This is an index of the slope of the function: the finer the precision of the judgement, the smaller the JND [34] (figure 1). In SJ tasks, the percentage of trials reported as simultaneous is plotted as a function of SOA, with the standard deviation of the underlying Gaussian distribution (the derivative of the sigmoidal psychometric function) taken as the observer’s sensitivity to asynchrony (equivalent to JND), and the centre of the distribution taken as an estimate of PSS [35].

Although TOJ and SJ are related, data from an SJ experiment do not necessarily transfer to a TOJ experiment in the same subject. That is, when we judge that two events are not simultaneous, we do not necessarily know their order. For example, observers might apply different criterion levels when performing the tasks. In fact, it has been shown that the maximum of subjective simultaneity inferred from SJ does not coincide with that inferred from TOJ [28,36].

In addition to TOJ and SJ, ternary response tasks have also been used, in which subjects make one of three choices: report which of two stimuli appeared first, or report whether the two stimuli appeared simultaneously. While the ternary response task would seem to address both simultaneity and temporal order, subjects’ TOJs typically change depending on whether they have the option to report simultaneity (ternary task) or not [37]. Thus, researchers have generally studied either TOJ or SJ in isolation [35].

### 4. Human psychophysical studies

In visual TOJ experiments with human subjects, various studies have reported thresholds (corresponding to approx. 80% probability of reporting one stimulus first) in the range of 50–80 ms [38–40]. Interestingly, thresholds do not seem to vary substantially for different sensory modalities. In a tactile TOJ task in which subjects had to report which hand (left versus right) was first stimulated mechanically, Yamamoto & Kitazawa [40] reported an average threshold of 74 ms. Kanabus et al. [41] tested TOJ for auditory as well as visual stimuli. The auditory stimuli were tones of different frequency, whereas the visual stimuli were different coloured pulses of light flashed at the same location by a light-emitting diode. The authors found that threshold SOA (75% per cent correct responses) was similar between the modalities, approximately 40 ms [41]. These findings are intriguing, because sensory latencies in the sensory periphery are much briefer for auditory and tactile stimuli than for visual stimuli, and thus presumably more reliable from trial to trial. The comparable TOJ thresholds between visual and non-visual modalities suggest that the reliability of TOJs is likely limited by central processes, or even that TOJs for different sensory modalities could be adjudicated by a common central mechanism, or a common brain area [34,42] (but see also [43] for a discussion about modulation of TOJ thresholds by other factors). As discussed in §3, an important aspect of TOJ experiments is that objective (SOA = 0) and subjective simultaneity (the PSS) of a pair of stimuli are typically not the same. In some
models of TOJ, a non-zero PSS is taken to reflect differences between the ‘arrival latencies’ of the signals corresponding to each stimulus at the brain area(s) that judges the temporal order [34]. These delays must represent the sum of latencies occurring from initial transduction of the stimulus in the sensory periphery, through integration and propagation of signals between synaptic stages to the ultimate decision stage(s) in the brain. Delays may vary with different attributes of the stimulus, such as contrast and eccentricity of visual stimuli [34], and have been taken as evidence about the relative strength or latency of these sensory signals. However, considerable evidence has accumulated to challenge the idea that TOJ is determined only by differences in arrival latencies, requiring an updated model for TOJ (see §7 of this review).

The PSS is also affected by changes in internal cognitive state. A prominent example is attention. Attention has always been intertwined with temporal order studies. More than a century ago, British psychologist Titchener [44, p. 251] proposed his law of prior-entry, in which ‘the object of attention comes to consciousness more quickly than the objects which we are not attending to’. Indeed, modern experimental psychologists have confirmed that exogenous and endogenous attention bias PSS towards the attended stimulus (i.e. the attended stimulus is reported first), both for intramodal TOJs (vision, auditory and tactile) [29,45–49] as well as cross-modal TOJs [43]. One concern is that the effects of attention could be due to response bias [50]. For example, if subjects are asked to attend to the right side of a display, then they are more inclined to report the first stimulus as the one appearing on the same right side. Using an orthogonal design, in which the dimension of response (e.g. report the colour of the first stimulus to appear) is orthogonal to the dimension for which attention is cued (e.g. left versus right), helps to mitigate response bias [45]. For example, to rule out response bias as an explanation of the attention effects on TOJ, Shore et al. [45] asked subjects to judge the order of presentation of two visual stimuli (a horizontal and a vertical line) and found that, although response bias has a large influence, exogenous (and to a lesser extent endogenous) attention cues could still affect the perceived order of visual stimuli. Similar results have been observed in the somatosensory domain when response bias was reckoned and an orthogonal design was used [36,46].

Schneider & Bavelier [30] proposed a number of possibilities for how attention could affect TOJ: (i) through sensory interaction between cue and stimulus (exogenous attention), (ii) by reducing the transmission time of the attended stimulus (attending to a feature or spatial location) and (iii) by affecting the decision mechanism. Regarding (i), exogenous (visual) attention cues are by definition presented near the item to be attended; thus, sensory interactions are inevitable. It is debatable whether these sensory interactions should be considered ‘attention’ in the context of TOJ. One way to address this confound might be to provide the exogenous cue using a different sensory modality than that used for the TOJ; for example, a spatially localized auditory or tactile cue could be used to exogenously attract spatial attention in a visual TOJ task [45]. With regard to transmission times (ii), little evidence supports the notion that attention decreases latency of neural signals, at least in animal models. In single unit neurophysiological studies in monkeys, Lee et al. [51] found that directing attention towards (or away from) a stimulus has little effect on the latency of neuronal signals in area V4 (approx. 1 ms decrease in latency for either 100% or 25% contrast stimuli). In another study, Sundberg et al. [52] found less than 1.5 ms of latency change owing to attention at the highest contrast level that they tested, again in area V4. Similarly, Bisley et al. [53] found that allocation of attention had no effect on response latencies in lateral intraparietal (LIP) cortex. Thus, latency differences owing to attention, if any, are slight. In regard to (iii), within a signal-detection framework, it is well established that attention can affect decision criteria as well as enhance sensory processing [54–57]. In a TOJ context, attention could affect TOJs by selectively affecting the decision criterion, such as changing the evidence threshold for a competitive ‘race process’ between the competing stimuli (see §7).

Importantly, changes in relative delays, as indexed by the PSS, have also been used to provide clues for the location of brain areas involved in TOJs, as described in the following sections.

5. Neuropsychology of temporal order judgements

Historically, it was the British neurologist Critchley [58] who reported timing deficits in patients with damage to the parietal cortex. However, the parietal lobe, in particular the right parietal lobe, gained much more attention in the late 1990s. Husain et al. [59] reported that patients with right parietal damage had deficits in an attentional blink task, indicating difficulties in temporal processing. Harrington et al. [60] found deficits in duration perception (interval timing) in individuals with right hemisphere damage; moreover, the subjects’ temporal performance was correlated with their ability to direct non-spatial attention. Based on patient studies, Husain & Rorden [61] suggested that the temporal–parietal junction (TPJ) encodes temporal information in the visual modality. Batelli et al. [62,63] carried out a series of experiments on parietal neglect patients, and based on their findings, proposed that the right parietal lobe constitutes a ‘when’ cortical visual pathway. Although these experiments did not directly involve TOJs, the results are strongly suggestive of timing deficits in neglect patients, i.e. patients with damage to the right parietal cortex.

Other neuropsychological studies provide more direct evidence for parietal involvement in TOJs. Rorden et al. [64] presented one stimulus in each visual field of patients with damage to right parietal cortex and found that the patients reported the ipsilesional stimulus first unless the contralesional stimulus was presented at least 200 ms earlier. Sinnett et al. [65] presented one shape in each hemifield of right parietal patients and found that the contralesional stimulus had to be presented approximately 200 ms before the ipsilesional stimulus in order for patients to report them with equal frequency. Baylis et al. [66] asked patients with either left or right parietal damage to report which of two stimuli was the second to appear, and found that a lead of approximately 200 ms was necessary for the contralesional stimulus to be reported as frequently as the ipsilesional stimulus, regardless of the side of the lesion.

A limitation of some of the earlier patient studies is that they were confounded by response bias [64,67]. Because parietal patients show strong biases towards stimuli on their ipsilesional side [68], in TOJ experiments they might report the stimulus on their ‘good’ side as appearing first whenever they are uncertain of the temporal order. This issue is most prominent when instructions are given to subjects...
(e.g. attend left). Orthogonal experimental designs reduce this confound (see §4). When such a design was used to control for response bias [65], shifts in PSS were still observed in patients [66], indicating both left and right parietal lesions can still affect TOJ judgements, with contralesional visual stimuli appearing later than ipsilesional stimuli.

While neuropsychological studies with patients have suggested changes in PSS resulting from brain lesions, the studies have generally not reported changes in sensitivity in TOJ tasks. This is unfortunate, because changes in sensitivity are less prone to response bias and could be more directly related to TOJ itself rather than related to secondary processes such as attention. Future neuropsychological studies on TOJ could be enhanced by a more quantitative psychophysical approach that includes measurements of sensitivity.

6. Localizing brain areas involved in temporal order judgement in normal subjects

A variety of behavioural studies have provided clues about areas involved in TOJ, at least with respect to vision. One approach has been to compare order judgements when both stimuli are presented to one eye (monoptic) versus when the stimuli are presented separately to the two eyes (dichoptic). The threshold for order judgements is lower when stimuli are presented monoptically [69,70], suggesting involvement of brain areas with monocular neurons in TOJ, such as V1 or visual thalamus. Observers also have a bias to perceive foveal stimuli before peripheral stimuli [69]. However, in the monoptic versus dichoptic and foveal versus peripheral comparisons, the two stimuli were presented spatially very close to each other, with only arc minutes of separation. TOJs at this scale may be best subserved by neurons with small receptive fields (RFs) in early visual areas. In addition, two successive stimuli in close spatial proximity could produce a dominant low-level motion signal that could allow observers to perform the task based on motion cues rather than order judgement per se [71].

With larger spatial separations of stimuli, localization studies in healthy human subjects again point to a role of parietal cortex in TOJs. Woo et al. [38] used temporary disruptive techniques to study TOJ in humans. While subjects made order judgements between two visual stimuli, one in each hemifield, the authors applied a single TMS pulse to either the left or right posterior parietal cortex. They found that the processing of the contralateral stimulus was delayed for 20–30 ms, but only when TMS was applied on the right, but not on the left side. The disruptive effect was evident only when the TMS pulse was given 50–100 ms after the onset of the first stimulus, which corresponds to when visual signals should reach parietal visual areas. In an fMRI study, Davis et al. [39] asked subjects to perform a TOJ, and compared it with an equally difficult shape judgement with one stimulus in each visual hemifield. Using fMRI, the authors looked at differential brain activity between the two tasks, and found bilateral activation in the TPJ specific to the TOJ task. However, a potential confound of the experiment is that only the TOJ task (and not the shape task) required selection in time and processing of the onset of events. Thus, the authors performed a follow-up control experiment in which both tasks required discriminating brief events concurrent with the onset of the visual stimuli, and only left TPJ activity was found in the TOJ task. This is in contrast to the study of Woo et al. [38,39] which showed a right parietal dominance in TOJ tasks. These apparently contrasting results point to the fact that fMRI results are correlational and do not necessarily address the causal involvement of a specific brain area in TOJ tasks, when compared with patient studies [63] and TMS studies [38]. Currently, TMS is the only available technique that can interfere directly and acutely with cognitive functions in humans, and TMS studies implicate the right parietal cortex in TOJ.

7. Models of temporal order judgement

Localization studies point to particular areas in the brain that could be involved in TOJs, but they do not address the underlying neural mechanisms. How are TOJs determined at the level of neural circuits? Forty years ago, Sternberg & Knoll [34] posited a straightforward model for TOJ, in which (neural) signals elicited by two competing stimuli are transmitted along independent channels in the brain, ultimately converging on a decision stage that compares the relative arrival latencies. In this view, stochastic variations in arrival times of the two signals give rise to the trial-by-trial variations in perceived temporal order, even for the same SOA. Arrival-time-based models of this sort have generally not specified a neural instantiation, but to our reading they allude mainly to signal propagation—the time needed for action potentials to move along axons. This view is reminiscent of the mechanism of horizontal auditory localization, which is based on the relative propagation times of neural signals from the two ears to a site of binaural comparison in the brainstem [72].

A great deal of evidence has accumulated that challenges arrival-time-based models for relative timing judgements. First, as mentioned in §3, many studies have shown differences between temporal order judgements and simultaneity judgements, which differ only in the instructions given to the subjects. Perceptual differences between TOJ and SJ paradigms imply differences in decision criteria, something that is not easily captured by arrival-time-based models. Second, temporal judgements between different sensory modalities with different peripheral processing speeds (e.g. vision and audition) imply some kind of central recalibration, which is more readily explained by different decision criteria than by normalization of propagation times [73]. Third, prior entry (attention) effects on relative timing judgements could also be accounted for by changes in decision criteria, but it is difficult to imagine how attention could affect propagation speeds, which would seem more ‘hard wired’. Indeed, there is little evidence that attention significantly affects neuronal response latencies in animal experiments (see §4). These and other observations argue that arrival-time-based models do not provide a general explanation for relative timing judgements.

The implication of malleable decision criteria in relative timing judgements suggests an alternative model grounded in signal-detection theory (SDT). As applied to the brain, SDT examines how decisions are adjudicated based on sensory evidence and internal factors. Particularly relevant to relative timing judgements are decision models in which sensory data are evaluated as they are collected over time. These models include sequential analysis and diffusion/drift or race models (for a review, see [74]). Race models constitute a contest to accumulate or integrate sensory evidence towards
competing alternative decision outcomes. Evidence for possible alternative outcomes is integrated and compared with a decision criterion or threshold; the evidence that first surpasses criterion ‘wins’ and leads to the corresponding percept or outcome. In this view, different decision criteria could result from different evidence thresholds for competing decision outcomes, or different (unequal) starting points or integration rates for the competing integration processes. We propose that TOJ could be a race-like accumulation of evidence, compressed over short time intervals on the order of the SOAs.

In race models for TOJ, what could cause the stochastic variations in perceptual report from trial to trial? As mentioned above, transmission times, in the sense of action potentials propagating down axons, are unlikely to contribute substantial variation in TOJ experiments, because axonal propagation times are very brief and reliable, and there is little evidence that propagation times can be modulated. However, stochastic variation in TOJ from trial to trial could arise in the strength of the sensory responses to the competing stimuli (e.g. the number of spikes evoked), or in the synaptic integration involved in propagating neural signals from one stage to the next. Implicit in this ‘bottom-up’ view is that there is a causal relationship between neuronal firing at relatively early stages of sensory processing and the animal’s subsequent perceptual report; that is, trial-to-trial variation in neuronal firing constitutes sensory noise that influences the subsequent decision stage. This might seem obvious to the point of triviality, but it implies an important experimental test: by looking for trial-by-trial correlations in neuronal firing with perceptual report, one could identify brain areas that are particularly important for perceptual judgements, such as TOJ. For example, while there are trial-to-trial variations in visual responses in the retina, these variations might not be well correlated with the animal’s perceptual report near threshold, whereas variations in downstream visual areas could be more strongly correlated. This would imply that the noise in the subsequent stages was more dominant to the ultimate decision stage, and that the sources of noise were fairly independent from stage to stage.

Much of the previous work relating firing of cortical neurons to perception has been based on the paradigm developed by Newsome et al. [75,76] in which monkeys discriminate the net direction of motion of noisy random-dot kinematograms in a two-alternative forced choice manner, and signal their perceived direction by a saccadic eye movement. In the race model proposed by Shadlen et al. [77,78], the random-dot stimulus would affect two ensembles of neurons with overlapping RFs, but with opposite direction preferences. Spikes in the competing neuronal pools would be integrated and thresholded, and the neuronal pool that reaches threshold earlier would determine the perceived direction [77,79].

In principle, a similar competitive race process could underlie TOJs in the brain. Distinct populations of neurons could encode the two stimuli, with the neurons’ RFs corresponding to the two spatially distinct stimulus locations in the TOJ paradigm. Spike counts would be pooled within each neuronal ensemble, integrated and thresholded, as described by a race process. The neuronal pool that reaches threshold earlier would lead to the percept of the corresponding stimulus appearing first.

In the random-dot kinematogram experiments, the animals typically view the stimulus for hundreds of milliseconds; thus, in weighing the perceptual decision, the brain could integrate sensory information for an extended period of time [80]. However, in TOJ experiments, subjects are able to discriminate temporal order with SOAs as short as a few tens of milliseconds. The informative neuronal spikes for the integration and race process are presumably confined to a time window on the scale of the narrow SOA, because if spikes were integrated for a long time after both stimuli have been turned on, then the neuronal signals would presumably be no longer informative of which stimulus appeared first. Thus, for TOJs, a bottom-up mechanism might be subsumed by an integration/race process compressed to the timescale of the SOA.

An implication of a race process compressed to the timescale of the SOA is that if one were to look for neuronal responses that covary with a subject’s TOJs, presumably the most informative correlations would be with respect to the earliest part of the neuronal response following the onset of the stimulus. In this view, if the response to the stimulus in a neuron’s RF happens to be slightly larger on a given trial, then the neuronal pool corresponding to that RF location should reach threshold a little sooner, and the subject would tend to perceive the stimulus in that RF as appearing first. If instead the RF response happens to be slightly smaller, then the subject would tend to perceive the stimulus in the RF as appearing second.

At first glance, it might seem odd to suggest that neurons representing only one of two stimuli could contribute to a TOJ between the two stimuli. It might seem that a neuron subserving TOJs should have large RFs encompassing both stimulus events, to allow for a direct comparison. However, this does not necessarily follow. For one, random variations in the firing of the two competing neuronal pools representing one or the other stimulus (but not both) could affect the decision at a later stage of the process. If so, neurons could play a causal role in the percept without being ‘directly’ involved in the decision process. For example, in the classic models of perceptual decisions in the visual motion pathway, evidence for one motion direction or the other is posited to be processed in distinct MT neuronal pools, with the integration taking place downstream in the visual hierarchy, e.g. in area LIP [74,81]. Moreover, the integration-to-threshold underlying the race process could also occur in independent pools of decision-related neurons. With a two-alternative forced choice design to ensure an unambiguous behavioural outcome, it seems plausible that at some stage the competing neuronal pools should interact in a winner-take-all manner (perhaps by mutual inhibition), but this could occur outside the sensory system (for example in the oculomotor system, in the classic perceptual decision experiments). Thus, there is no _a priori_ reason to expect that neurons involved in the decision mechanism must have a sensory representation of both stimuli in the TOJ task.

In interpreting neuronal data in TOJ experiments, another important consideration is that the subject’s report of temporal order is invariably separated in time from the narrow time window of the SOA when the crucial neuronal spikes are presumably accruing. For example, psychophysical thresholds for SOA in TOJ tasks are on the order of tens of milliseconds, yet behavioural response times are typically on the order of hundreds of milliseconds. During the time between the onset of the stimuli and the behavioural response, the neuronal data might be ‘buffered’ in some manner, or there could be secondary modulations in neuronal firing because of selective attention to one stimulus or the other (e.g. to the first stimulus to occur) or because of selective motor preparation. For
example, secondary attentional modulation can affect detection of transient sensory events that occurred even hundreds of milliseconds earlier [82]. Thus, neuronal activity that correlates with the animal’s judgement of temporal order might not be causal (bottom-up) to the animal’s judgement, and should be interpreted cautiously. For example, modulations in neuronal activity occurring hundreds of milliseconds after both stimuli have been presented in a TOJ experiment could be related to selective attention after the decision has been made (i.e. top-down rather than bottom-up).

As a final word on models for relative timing judgements, we point out that simple race models cannot explain all aspects of relative timing perception. In some cases, ‘high level’ Gestalt constraints can override simple temporal relationships among the incoming stimuli (for a review, see [9]). For example, when two auditory tones are presented with a temporarily intervening noise burst, the two tones are perceived as grouped together, followed by the noise burst (for an example, see [83]). Observations such as these are difficult to reconcile with SDT-based race models, but do not negate the ideas of race processes. For example, grouping constraints may occur at a higher processing level in the brain, with the ability to reverse the outcome of race processes occurring at lower processing levels.

8. Neurophysiological mechanisms of relative timing

(a) Human electrophysiology

With the model for TOJs in mind, it is useful to examine experiments that have measured neuronal activity during TOJ experiments. We start with non-invasive recording methods in human subjects. As described above, for the same SOA, ideally one would like to identify neuronal activity that varies on a trial-by-trial basis with the subject’s perceived temporal order. This approach is difficult in humans, because most non-invasive means for recording from humans lack single-trial resolution. Therefore, in the few studies on the topic, rather than looking for trial-by-trial correlations, the authors instead tried to bias perceived temporal order in some manner, and then looked how the bias affected neuronal activity averaged over many trials. For example, McDonald et al. [47] recorded event-related potentials (ERPs) over the right and left occipital cortex while subjects judged which of two visual stimuli, presented to the left and right of the fixation point, appeared first. The authors biased attention to one side or the other by presenting an auditory stimulus that was spatially offset to the left or right of the midline, and that occurred 100–300 ms before the onset of the corresponding visual stimulus. The auditory cue produced a PSS shift of approximately 70 ms towards the side of the auditory cue. For the SOA = 0 condition, the authors found that, for a given occipital electrode, the magnitude of the ERP signal was enhanced when the auditory cue was contralateral to the electrode compared with when the cue was ipsilateral. However, there was no change in the latency of the ERP signal [47]. For the SOA = 70 ms condition (close to PSS), the authors found the early ERP components were offset in time by approximately 70 ms between the left and right occipital electrodes, much like the visual stimuli themselves. But when the SOA is close to the PSS, it might be expected that neuronal signals would be more closely aligned in time—if those signals indeed reflect the subjective perception of near simultaneity. Instead, the 70 ms separation of the occipital ERP signals better reflected the physical visual stimulus than the subjective perception, perhaps indicating that the occipital cortex (i.e. V1) is not strongly involved in the TOJ. In this view, stronger correlations with subjective TOJ might be found in higher brain areas. This issue could also be addressed if, for a fixed SOA, the authors had sorted their data according to the observers’ perceptual report before averaging, to reveal neuronal signals related to the subject’s subjective percept of temporal order rather than to the visual stimulus.

In another study, Vibell et al. [48] instructed subjects to attend to one sensory modality (touch or vision) in a bimodal TOJ task, and found a difference of approximately 40 ms between the PSS for the attend-to-touch condition and the attend-to-vision condition. In the ERP data collected over scalp occipital leads, the authors observed latency shifts in P1, N1 and N2 components as well as P300 potential when attention was directed to the visual modality compared with when attention was directed to the tactile modality. The authors concluded that attention decreases the latency of the visual signal, consistent with prior entry theory. However, the ERP components peaked at approximately 150 ms for P1 and approximately 440 ms for P300 [48]. The late time course could indicate that the modulation of these signals is secondary to the TOJ, perhaps reflecting post-decisional attention towards the first stimulus detected. In addition, the authors were not able to identify or reliably measure C1, the earliest component originating primarily from V1, which could have provided a more direct read-out of early stages of visual processing. One should also bear in mind that even a change in the latency of an ERP component does not necessarily indicate a change in the latency of individual neurons [51]. Each component of the ERP reflects the pooled activity of large ensembles of cells that could have different individual latencies. Attention could preferentially modulate the magnitude of responses of neurons with different response latencies, and the summation of the magnitude changes could cause an apparent change in the latency of the ERP component [51]. Moreover, latency changes, if they exist at all, are unlikely to arise from changes in the actual propagation of neural signals, but rather from the way that those signals are integrated or read-out by postsynaptic neurons. For example, attention could better synchronize presynaptic inputs to cause a more concerted—and thus earlier—postsynaptic response. In this sense, a modulation that affects response amplitude may not be so different from a modulation that affects latency.

An important challenge in human neurophysiology (such as multi-channel EEG) is to provide sufficient spatial resolution to distinguish activation by the two stimuli in a TOJ task, as in the study by McDonald et al. [45]. Under these conditions, comparisons could be made between the neural signals for opposite perceptual reports in response to the same SOA. Properly designed experiments [84] could then reveal neuronal responses that are related to the subject’s perceived temporal order rather than related to the physical properties of the stimuli.

(b) Single-unit electrophysiology in monkeys

Single-unit physiology in behaving animals has the advantage of providing the most precise information about the
timing of neuronal responses. Monkeys provide an attractive model species for this purpose because the animals can be trained in complex timing tasks. A number of studies have examined neuronal correlates of interval timing in trained macaques. For example, Leon & Shadlen [85] found that LIP neurons represent elapsed time when the animals had to make eye movements to indicate whether the duration of a test stimulus was shorter or longer than a reference duration. Mayo & Sommer [86] trained animals to compare a time interval demarcated by two successive flashes to a reference interval and produce different eye movements, and found that the dynamics of visual adaptation in frontal eye field neurons matched the animals’ temporal discrimination. Additional studies have revealed potential neuronal correlates of interval timing in tasks in which animals made movements at the end of proscribed intervals, without explicit prompting [87–90]. However, to the best of our knowledge, there have been no published studies on neuronal correlates of TOJs in the monkey.

Two of the authors of this review recently trained two macaque monkeys to perform a TOJ task in which the animals reported the order of appearance of two visual stimuli, presented with a range of SOAs from 0 (simultaneous) to 126 ms [91]. The two stimuli were spatially configured so as to prevent the animals from using apparent motion as a cue to judge relative timing. In addition, rather than having the animals directly indicate the first stimulus that they perceived to appear (e.g. by making an eye movement to one stimulus or the other), we instead imposed a delay period followed by one or the other stimulus (randomly chosen) turning green. The animals were trained to release a lever if the green stimulus was the first stimulus that had appeared, or to continue holding the lever if the green stimulus was the second stimulus that had appeared. This design ensured that the animal’s percept of temporal order was dissociated from how the animals signalled that percept. This approach is important, because many parietal neurons also have activity related to movements of the eyes or arms, which complicates interpretation of neuronal signals related to perceptual judgements per se [84].

After training for months, the animals reached a high level of performance. The threshold SOA (82% maximum performance) was 34 ms for one animal and 41 ms for the other animal, and both animals’ performance was significantly better than chance even for SOA = 9 ms, the shortest non-zero SOA, we could present given the frame rate of the stimulus monitor. These threshold SOAs are faster, or at least comparable to, thresholds reported in human TOJ experiments [38–40], even though that the animals could not use motion cues to perform the TOJ task.

After the animals’ behavioural thresholds stabilized, we recorded from single neurons in the LIP. After mapping the RF of each unit, one stimulus was positioned to fall within the RF, whereas the other stimulus fell outside the RF. The animals performed the TOJ task, whereas neuronal data were collected. We analysed the neuronal responses as a function of the animal’s judgement of temporal order on each trial—whether he reported the stimulus in the RF as appearing first or second. LIP units tended to respond more strongly when the animal reported the stimulus in the RF as appearing first (figure 2). This difference in neural response was absent in the baseline period before the stimuli appeared, yet was observed both in the early visual response (phasic period, 40–100 ms after stimulus appeared in RF) and more reliably in the late visual response (tonic period, 100–250 ms after the stimulus appeared in the RF).

To quantify this effect, we used a metric called choice probability (CP) to analyse the responses of each neuron [92]. CP is related to the difference between the distributions of responses when the animal reports one or the other perceptual outcome [93]. CP of 1.0 would indicate that the neuronal firing was a perfectly reliable predictor of the animal’s TOJ, whereas a CP of 0.5 would indicate that the neuron’s firing was completely uninformative of the TOJ.

A most interesting case is when the stimuli were presented simultaneously, i.e. SOA = 0. In this case, there was no meaningful signal in the visual stimulus (at least with respect to temporal order), and the animal’s judgements accordingly tended to be nearly evenly divided between the two possible outcomes. Among the population of LIP neurons that we studied for the SOA = 0 condition, approximately 10% of the neurons had CP that was statistically significantly different than 0.5 during the brief phasic visual response period, and approximately 25% had CP that was statistically significant during the tonic period. Similar results were obtained for other brief SOAs for which the monkeys’ behavioural judgements were divided between the two possible outcomes. Note that because the CP analysis was only conducted within a set of trials with the same SOA, the physical visual stimulus (i.e. the location of one stimulus and the

Figure 2. Response of a single LIP neuron for SOA = 0 ms. (a) Shows spike-rate function aligned to the appearance of the stimulus in RF, for all trials. (b) Depicts the same data, but with the trials divided by whether the animal reported the stimulus in the RF appearing first (solid blue) or second (dotted red). Spike-rate functions for individual units were generated by convolving 1-ms-binned histograms with a Gaussian function (SD = 15 ms). (Online version in colour.)
other stimulus, as well as the temporal offset between them) was identical between the trials; thus, differences in neuronal activity could not be attributed to differences in low-level visual input, but rather were related to the animals’ TOJ.

In summary, the variations in the firing of LIP neurons reflected the judgement of temporal order on a trial-by-trial basis. These data are at least consistent with parietal neurons playing a role in the perception of visual temporal order. Across the neuronal population, there was the relationship between the earliest period of visual responses and the animal’s TOJ, although this relationship was weaker or more difficult to detect than that emerging later in the trial. However, these recordings were from single neurons, and the relevant time window so brief that only a few spikes were fired during that time, making the detection of correlations challenging. It is possible that stronger correlations would emerge if simultaneous recordings were made from many neurons as the animals performed the TOJ task. It is also likely that other brain areas, in parietal cortex and elsewhere, are involved in the TOJ, perhaps even more directly than LIP. Nonetheless, the relationship between neuronal signals in the early visual transient responses and the animal’s decision near threshold is consistent with a bottom-up model for TOJs, in which trial-by-trial variation in sensory responses to the competing stimuli determine the perceived order. Stronger correlations between neuronal activity and TOJ emerged later in the trial, over hundreds of milliseconds. However, these relatively late signals could be more related to processes secondary to the TOJ; for example, attention could be drawn to the location of the stimulus that the animal judged to appear first (although the animals were required to maintain fixation during that period). More experiments are needed to tease apart these components of TOJ, but we believe the basic experimental paradigm is valuable for addressing underlying neuronal mechanisms, both in animal studies and in non-invasive neurophysiological experiments in humans.

9. Concluding remarks

Understanding the neuronal underpinnings of TOJ will shed light more generally on decision-making in the brain, and could provide a more mechanistic framework for interpreting the effects of attention. Neurophysiological experiments on TOJs would benefit from the framework of SDT, which has been applied successfully to the study of other perceptual decisions. In particular, looking for trial-by-trial correlations between neuronal activity and the subjective perception of temporal order can provide important clues about the brain circuits involved in TOJ, and will help to address whether TOJs are a distributed or centralized brain function. Properly designed electrophysiological experiments should work near perceptual threshold to identify neuronal signals that are correlated with temporal judgements per se, and should factor out sensory input and motor output as sources of neuronal response variation [84]. This approach is particularly applicable to animal studies in which trial-by-trial neuronal responses can be recorded with high fidelity, but similar general approaches could also be applied to human electrophysiological or imaging studies.

Acknowledgements. The authors thank Richard Born, Patrick Cavanagh, Alex Holcombe, John Maunsell and Thomas Luo for discussions on parts of the material covered in this paper.

References


