



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/172494/>

Version: Published Version

---

**Article:**

Viera, G. (2022) The perceived unity of time. *Mind & Language*, 37 (4). pp. 638-658. ISSN: 0268-1064

<https://doi.org/10.1111/mila.12331>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# The perceived unity of time

Gerardo Viera 

Department of Philosophy, University of Sheffield, Sheffield, UK

## Correspondence

Gerardo Viera, Department of Philosophy, University of Sheffield, 45 Victoria Street, Sheffield S3 7QB, UK.  
Email: g.viera@sheffield.ac.uk

## Funding information

ERC Consolidator Grant STYDS, Grant/Award Number: 726251

While we perceive events in our environment through multiple sensory systems, we nevertheless perceive all of these events as occupying a single unified timeline. Time, as we perceive it, is unified. I argue that existing accounts of the perceived unity of time fail. Instead, the perceived unity of time must be constructed by integrating our initially fragmented timekeeping capacities. However, existing accounts of multimodal integration do not tell us how this might occur. Something new is needed. I finish the paper by articulating the hurdles that must be overcome to provide an account of the perceived unity of time.

## KEYWORDS

format of representation, mental representation, multimodal perception, sensory integration, temporal perception

## 1 | INTRODUCTION

It is the shared duty of the various sensory systems to tell us how events in our environment are temporally structured. Through vision, audition, taste, touch, smell, and whatever other sensory systems we possess, we learn about when things occur around us. Yet, despite gathering this information through various sensory systems, we nevertheless perceive time as unified. The world appears to consist of a single timeline within which all of the events we perceive occur. It is in this way that there is a perceived unity of time.

This paper has two goals. First, to show that standard accounts of the perceived unity of time fail. Second, to articulate the explanatory hurdles that any adequate account of the perceived unity of time must overcome in explaining how temporal information is integrated across modalities and timescales. Acknowledging these hurdles reveals a host of questions about temporal perception that have not been directly addressed in the philosophical and scientific literature.

---

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Author. *Mind & Language* published by John Wiley & Sons Ltd.

The paper goes as follows. Section 2 provides a characterization of the perceived unity of time as a target for explanation. Section 3 lays out the standard accounts of the perceived unity of time. While they differ in their details, both accounts share a common strategy. They first provide a general explanation of how perception attributes temporal properties to perceived events (e.g., perceiving events as having durations or as standing in temporal relations to other events), and then, from this general attributive story, an explanation of the perceived unity of time supposedly emerges more or less straightforwardly. Section 4 argues that any account of the perceived unity of time adopting this strategy is bound to fail. Temporal perception is composed of initially fragmented timekeeping capacities and the explanation of the perceived unity of time must explain how these fragmented capacities are appropriately coordinated with one another. Section 5 draws out parallels that emerge between temporal and spatial perception and argues that the integration of temporal information across modalities must differ in significant ways from integration of spatial information. Finally, Section 6 articulates the explanatory challenges that any account of the perceived unity of time must overcome.

## 2 | DESCRIBING THE PERCEIVED UNITY OF TIME

When we, conceptually sophisticated adults, reflect on the world, the world appears to contain a single temporal dimension within which all of the events we perceive occur.<sup>1</sup> How do we explain this appearance? Why does the world seem to contain a single timeline and not multiple timelines associated with different modalities? This is the core of the perceived unity of time that will be the focus of this paper. In this section, we will do some unpacking to better grasp this target phenomenon.<sup>2</sup>

The first thing to notice is that the phenomenon at issue here is not a *purely* introspective or phenomenological one. The perceived unity of time concerns how *the world* appears to be temporally structured, and only in so far as our experiences seem to occur within the world, also speaks to how experiences seem to be temporally structured. When you leave your house, the thought that you forgot your keys might pop into mind after you hear the door click behind you. When we reflect on this episode, our thought coming to mind, that introspectively accessed event, seems to sit in the same timeline as the other events in the world that we perceive. It is the appearance of this unified timeline of events in the world that is our focus.

We can break down the target phenomenon into two separate perceptual capacities. First, there is what I will call *temporal localization* (or *localization*). We do not merely perceive events as having temporal structure, but we perceive events as being located in time relative to the present. Attending a musical performance, some things will seem to be occurring *now* (e.g., the sound of the sustained guitar), others will seem to have occurred at some point in the past (e.g., the initial sound of the note being struck), and often enough, we will have expectations that certain events will occur at some specific moment in the future (e.g., when the chorus will begin and the movements of the musicians will change). The perceived unity of time involves more than the appearance of a single timeline, but it involves the perception, across modalities

---

<sup>1</sup>Without access to how things appear to infants and non-linguistic creatures, the reflective evidence discussed here is restricted to language using adults.

<sup>2</sup>Aspects of the perceived unity of time, notably what I call *localization*, were described by Dennett and Kinsbourne (1992). While their primary aim was to debunk a particular theory of consciousness, *the multiple drafts model* (MDM) of consciousness they develop bears similarities to an account of the perceived unity of time suggested towards the end of Section 6.

and timescales, of events as being located at specific moments in that timeline relative to the present moment.

Second, there is what I will call *comparability*. We seem to have no introspectively available difficulties in comparing temporal properties across timescales and modalities. We can readily compare the duration of a seen flash of lightning and a heard crash of thunder. This is possible even though they are perceived through different modalities and at different timescales—milliseconds through seconds. Furthermore, when we compare these properties, we seem to understand them as being the same kinds of properties—*temporal properties*. In part, the taking of these properties as being of the same kind is a phenomenological datum. These experiences seem to have something phenomenally in common *qua their being temporal experiences*.<sup>3</sup> This contrasts with our abilities to make comparisons across magnitude types more generally. When asked subjects can compare the intensity of a flash of light with the intensity of a sound, and their responses show at least some intrapersonal stability (Spence, 2011). However, even though we can make these cross-magnitude-type comparisons, there is a lingering awkwardness to them. There is a sense that the compared intensities are not of the same type. There seems to be a phenomenological difference in the intensity of a sound and the intensity of a light. This lingering awkwardness is missing in the temporal case. We understand in the temporal case that the magnitudes being compared are all temporal properties regardless of which modality is used to detect them and over what timescale they occur.

While I have characterized localization and comparability in terms of first-person report, evidence for similar capacities can be found in the behaviors of human and non-human animals. Consider localization. To coordinate behaviors with events in the world, we need to gather information about *when* events are occurring and we need to be able to predict when events will occur. Think of what it takes to dance with a partner. You see and feel their movements and hear the music. Your movements have to be coordinated with the temporal information gathered through all of these different senses and your expectations of when the music and your partner's movements will change so that you know when and how to move. Many activities require similar localization. Consider what it takes for a predator to intercept prey. Some sharks, for instance, integrate information from their various senses to fix *when* to attack (Gardiner, Atema, Hueter, & Motta, 2014). Furthermore, for representations of when events occur to coordinate behaviors, they must locate events in *egocentric* temporal orderings, that is, relative to *now*. To anthropomorphize the situation, knowing that some event will occur at noon will not allow you to coordinate your behaviors with that event, unless you know how far from now noon is.<sup>4</sup> So, we have some reason for thinking that humans and non-humans locate events in time relative to the current moment to control behavior.

Let us consider the behavioral evidence for comparability by first noting an important contrast. A general finding on how animals, including humans, represent magnitudes is that there is a complex pattern of interactions between representations of different magnitude types. These patterns involve cases where the representation of a particular value for one magnitude-type distorts the representation of other magnitude-types (Pinel, Piazza, Le Bihan, & Dehaene, 2004; Walsh, 2003). To give just one example, displays with more objects in them are often perceived as having longer durations (Javadi & Aichelburg, 2012). While we can quantify the extent to which these different magnitude representations influence one another, we need not think these interactions increase the reliability of the representation of either magnitude type.

---

<sup>3</sup>Thank you to a referee for emphasizing this point.

<sup>4</sup>The point is similar to Perry's in *The problem of the essential indexical* (1979).

Instead, the interaction supposedly arises as a quirk of the machinery that underlies magnitude representation and behavior.<sup>5</sup> However, crossmodal temporal representations often do interact in ways that seem to facilitate reliably representing the world. To take one example, in a study by De Corte and Matell (2016), rats were trained to expect food at a particular location 10 seconds after a visual cue and 20 seconds after an auditory cue. The rats were then presented with a combination of the auditory and visual cues. What was observed was that the rats showed an expectation that the food would appear after 15 seconds. Unlike the pattern of cross modality type effects, these temporal representations interacted in ways that make sense if the perceptual system is aiming at integrating them to reliably represent the world—that is, the representations were treated as providing information about the same type of property in the world. This is what we would expect from comparability.

It may turn out that the reflective and behavioral evidence point to distinct phenomena. One's expectation of how that will turn out may trade on general assumptions about the division between perception for action and perception for awareness. However, we can remain neutral on this point since the important thing to notice is that humans and non-humans must have some way of integrating temporal information across modalities and timescales that allows for comparability and localization.

Two final points of clarification. First, what has been described so far is best understood as a competence or capacity. We are typically capable of locating events in time and making comparisons between instances of temporal properties. However, there are well known cases where we systematically fail to do so properly. For instance, we typically judge sounds as having longer durations than visual stimuli with the same objective durations (Wearden, Edwards, Fakhri, & Percival, 1998). Similarly, at very short interstimulus intervals, we can perceive a pair of events as being non-simultaneous, but fail to reliably perceive which event came first (Poppel, 1988).<sup>6</sup> In both cases, there is a performance failure—we fail to properly compare or locate temporal properties—that is revealed through empirical, not introspective, means. However, we often perform well in similar tasks, and the competence itself requires explanation. It is this general capacity that will be the focus of this paper.

Second, there are similarities between this paper's target and Molyneux's question. Is time represented via an amodal format, that is shared by the different modalities and to which no single modality has a unique claim, or is time represented multimodally, in that each modality represents time in a modality specific way (Richardson, 2014)? Nothing in the characterization of the perceived unity of time demands that we make a decision about this at this moment. Similarly, nothing in the characterization of the perceived unity of time, as studied in animals or adults, provides us with an answer to Molyneux's question with regards to time. Will a newly sighted person, for instance, be able to compare a seen duration with a felt one? In laying out the target phenomenon, we should remain open that this question may be answered either way. However, as we shall see in Section 4, I will argue that some temporal properties are represented amodally while others are represented multimodally. Therefore, we have some reasons to doubt that Molyneux's question can be given a common answer for all aspects of temporal experience.<sup>7</sup>

---

<sup>5</sup>For different explanations of this “quirk” see (Pinel et al., 2004; Walsh, 2003). In some cases the quirk might be beneficial, in others not.

<sup>6</sup>For another case involving localization failures, see Holcombe (2015).

<sup>7</sup>This is in contrast to Richardson (2014). Thank you to a referee and Jacob Berger for raising the connection between the target of this paper and Molyneux's Problem.

### 3 | THE STANDARD APPROACHES

Most accounts of the perceived unity of time adopt one of two general approaches—internal clock approaches and mirroring approaches. Despite their differences, these approaches share a common strategy. First, they provide a general explanation for how perception attributes temporal properties to perceived events. Then, from that attributive story, an explanation of the perceived unity of time emerges straightforwardly. Nothing further needs to be posited to account for the perceived unity of time over and above what is posited in the explanation of how temporal properties are attributed to events. In this section, we will look at these two approaches. Since the aim of this paper is to show that any account of the perceived unity of time that adopts this general strategy is bound to fail, I will not be considering whether these accounts succeed by their own lights.

#### 3.1 | Internal clock approaches

Our initial sensory responses to the world are largely controlled by external events impacting our sensory transducers. When you see a flash of light, photons impact the retina and a flurry of visual processes begin to unfold. When the light disappears those visual processes soon end. According to internal clock approaches, these initial sensory responses to events in the world do not themselves represent the temporal properties of those events. Instead, the temporal contents of perception are contributed by an internal clock that monitors the timing of these proximal sensory processes, and on the basis of the temporal measurements of those processes, perception as a whole attributes temporal properties to perceived events. While different versions of the internal clock model have been developed over the years, we will focus on *scalar expectancy theory* (SET) (Gibbon, Church, & Meck, 1984) since it is arguably the most influential internal clock model in the literature.<sup>8</sup>

The original variant of SET accounted for the perception of time through a three-component system involving a supramodal pacemaker-accumulator clock, a memory store, and a decision/comparator mechanism. The pacemaker produces *pulses* at a regular rate that are tallied by an accumulator system. Since the pulses are produced at a regular rate the total number of pulses tallied during some interval provides a measurement of the duration of that interval. It is this supramodal pacemaker-accumulator mechanism that is used to measure the temporal properties of the various modality specific sensory processes. These initial measurements, represented via total number of accumulated pulses, are then compared to stored pulse-based measurements for either particular events or averaged measurements for event types (Jones & Wearden, 2003). The result of the comparison is a relative duration judgment in which the current measured event is determined to be longer than, shorter than, or of equal duration to the measurement in memory. On the basis of this entire process, temporal properties are attributed to the events in the world responsible for producing the proximal sensory responses measured by the internal clock. This system, that primarily represents interval durations, can attribute a range of temporal properties to events in the world. Temporal order is given in terms of the interval separating event boundaries. Properties like rhythm, rate, and so forth, are attributed by applying simple mathematical operations on initial interval representations (Gallistel, 1990).

---

<sup>8</sup>Similar arguments could be raised against other internal clock models like the striatal beat frequency model (Matell & Meck, 2004).

In this way, we have a general story for how perception attributes various temporal properties to events across modalities and timescales.

The transition from the attributive story to an explanation of the perceived unity of time can be summed up in slogan form—the unity of the timekeeping mechanism accounts for the perceived unity of time. Comparability is easily accounted for. Since durations, for instance, are all encoded via that same pulse-based supramodal timekeeping system, which already includes a mechanism for comparing pulse-based measurements, there is nothing to distinguish crossmodal from intramodal comparisons of temporal properties. Only the inputs to the centralized timekeeper differ between the crossmodal and intramodal cases. Localization is explained through the single supramodal clock mechanism that provides a common temporal ordering for all of the events that we perceive. Events are then located within this temporal ordering by noting the length of the interval that separates the particular event from the current moment. Nothing is needed beyond the attributive machinery to explain the perceived unity of time.

Other variants of SET replace the single supramodal pacemaker-accumulator mechanism with modality specific clock mechanisms as part of the overall timekeeping system (Chen & Yeh, 2009; Wearden et al., 1998). This move is largely motivated by observed variation in the precision of temporal perception across modalities. For instance, each sensory modality has a different minimum ISI needed to reliably perceive two stimuli as non-simultaneous (Poppel, 1988). Similarly, discrimination thresholds for interval lengths differ depending on the modality of the stimuli marking interval boundaries (Grondin, 2003). To account for these differences, the move is to posit modality specific clocks that pulse at difference rates (see discussion in Chen & Yeh, 2009). From here, the attributive story remains much the same. Initial measurements produced by pacemaker-accumulator mechanisms are compared to stored measurements, and these comparisons are the basis for the attribution of temporal properties to perceived events.

Once we introduce modality specific clocks with differing pulse rates, we get an explanation for the variation in timekeeping precision across modalities. More precise modalities, like audition, have clocks with faster pulse rates. However, this causes a problem for comparability, since  $N$ -pulses from the auditory clock will represent a different duration than  $N$ -pulses from a slower clock, such as vision's. In fact, there is evidence of increased variability in temporal judgements when subjects are asked to make crossmodal comparisons (Penney, Gibbon, & Meck, 2000; Zhang & Zhou, 2017). Nevertheless, a simple multiplication operation can allow for a normalized means of encoding temporal information via a common code. The result is that despite the differences in the initial pulse-based codes, a common pulse-based code is easily obtained (we can remain neutral as to whether this is an amodal code). Therefore, the very same explanation of comparability is given in the modal specific clock version of SET as was given in the single clock variant.

Localization, however, requires more than just modality specific clocks. Consider what is required for the attribution of crossmodal temporal relations—such as the temporal interval separating a flash of lightning and a crash of thunder. Within an internal clock framework, one would have to introduce a supramodal clock mechanism to make these attributions (for instance, this is done by Klink, Montijn, & van Wezel, 2011).<sup>9</sup>

---

<sup>9</sup>This section concerns pure internal clock accounts. Pure clock-based accounts of temporal perception must introduce supramodal clock mechanisms. However, nothing stops an internal clock theorist from introducing the sort of crossmodal mechanisms described in Section 4.2. Hybrid versions like this are closer to what I endorse, however, the evidence presented in Section 4 suggests much more variation in timekeeping mechanisms than mere modal specificity.

So, on this internal clock approach to temporal perception, we attribute temporal properties to events in the world through a combination of modality specific and supramodal clock mechanisms.<sup>10</sup> Localization is accounted for by this supramodal clock mechanism in the same way as it was in the single clock variant of SET. The supramodal clock not only provides a common temporal order for all of the events that we perceive but also provides a means of locating events within that ordering. Once again, nothing is needed over and above the attributive machinery to account for the perceived unity of time.

### 3.2 | Mirroring approaches

The other approach to explaining the perceived unity of time appeals directly to the temporal properties of sensory processes themselves. According to mirroring approaches, the temporal contents of perception (or experience) mirror the temporal properties of perceptual (or experiential) processes themselves.<sup>11</sup> To see understand the approach, consider the following example.<sup>12</sup>

Imagine an approaching thunderstorm. When the storm is far away, the flashes of lightning will appear to occur prior to the crashes of thunder. As the storm approaches, the apparent gap between the lightning and thunder shrinks, until the thunder and lightning appear simultaneous. According to the mirroring approach, perception is able to represent these changing temporal relations in virtue of the changes in the temporal relations between experiences of the lightning and experiences of the thunder. As the perceived gap shrinks there will be a corresponding (i.e., mirroring) shrink in the actual temporal interval between experiences (or the corresponding perceptual states). The account can then be generalized beyond temporal order to all of the temporal contents in perception. We might perceive the thunder as lingering longer than the short strike of lightning, and this too will mirror the relative durations of the experiences of thunder and the experiences of lightning. Through this mirroring, perception latches onto and attributes temporal properties to perceived events.

If we adopt a mirroring approach, then, once again, we have a simple explanation of the perceived unity of time.<sup>13</sup> All that is needed to get this explanation of the perceived unity of

---

<sup>10</sup>Maniadakis and Trahanias (2016) develop a similar account inspired by striatal beat frequency models.

<sup>11</sup>For examples see Arstila (2015), Dainton (2000), Foster (1991), Mellor (1981), Phillips (2014), Rashbrook (2013). These authors do not maintain that experience having a particular temporal property is sufficient for that property appearing as part of the experience's content. Rather, they claim that it is necessary that if experience has a certain temporal content, then the experience's temporal structure will mirror this content. Something else makes a certain temporal property of experience part of its content. Nevertheless, mirroring plays a content *enabling* role. Without the appropriate mirroring, experience could not have its temporal contents. Some (Arstila, 2015; Foster, 1991; Mellor, 1981), go further, suggesting that mirroring plays a content *determining* role, in that the temporal contents of experience are determined by some subset of the temporal properties of experience itself (for instance, the duration of an experience may determine duration content, while the date on which an experience occurs would not be reflected in experience). On either interpretation, similar accounts of temporal unity are possible, and the same objections raised in the next section would apply. See Lee (2014) and Watzl (2012) for general criticisms of mirroring views.

<sup>12</sup>This section's goal is to assess the mirroring view's account of the perceived unity of time. I am granting that consumer systems can utilize information carried by the timing of sensory processes. A referee pointed out that granting them this may be too generous since a challenge raised in this paper is to provide an account of how temporal information is transformed into a format usable by consumer systems. Lee (2014) has argued this story ultimately undermines mirroring accounts.

<sup>13</sup>While the phenomena might ultimately be related, the perceived unity of time should be distinguished from the phenomenal unity of consciousness (Bayne & Chalmers, 2003). While some theorists discussed in this section, for example, Dainton and Rashbrook, argue that mirroring cannot account for the phenomenal unity of consciousness, they nevertheless appeal to mirroring to account for the perceived unity of time.

time off the ground is a simple assumption about the metaphysics of time. Since our modality specific sensory processes themselves occur within the single timeline of events in the world—that is our metaphysical assumption—then we can explain why the world appears to be temporally unified by appealing to the unity of time itself. Modality specific perceptual processes occurring at particular moments in a single worldly timeline. This accounts for localization since the appearance of events as occupying a common timeline is inherited from our experiences occupying a single timeline. Comparability is also explained by the general mirroring principle. Any consumer system capable of making use of the temporal contents encoded in one modality must be able to exploit the temporal properties of those sensory processes, since these are the proximal physical properties of the perceptual system in virtue of which perception attributes temporal properties to perceived events. However, it is the very same type of temporal property across perceptual systems that carries this content. Therefore, any consumer system capable of making use of the temporal information encoded in one modality can, in principle, utilize the temporal information found in the other modalities (provided it has access to the relevant sensory processes). Comparing temporal properties of events detected through different modalities simply involves comparing the temporal properties of more proximal sensory processes. Once again, nothing over and above the machinery needed to account for how temporal properties are attributed to events in the world is needed to account for the perceived unity of time.<sup>14</sup>

At this point, it is useful to notice a common feature to the internal clock and mirroring accounts of the perceived unity of time. In both cases, temporal perception is conceived as a single psychological phenomenon (or at least, a sufficiently homogenous assortment of phenomena). As a result, a general story for how perception attributes temporal properties to perceived events seems plausible. Once we have this general story, then the perceived unity of time is easily explained since there will be something like a common code in which perception encodes temporal information. In some cases, researchers are explicit about why they think this. For instance, a recent paper (Hartcher-O'Brien, Brighthouse, & Levitan, 2016) argues that a good reason for pursuing a unified mechanism/explanation that underpins temporal perception is that the various temporal properties that we perceive appear so intimately related to one another. That is, from an observation of the perceived unity of time, it is assumed that temporal perception is a singular psychological phenomenon, which gives rise to these sorts of approaches for understanding the perceived unity of time. In the next section, this assumption that temporal perception is a singular psychological phenomenon will be at issue.

#### 4 | THE FRAGMENTATION OF TEMPORAL PERCEPTION

In the first part of this section, I will argue that “temporal perception” does not pick out a single psychological capacity. It instead acts as an umbrella term picking out various timekeeping capacities that are specialized for specific aspects of the temporal structure of the world. Then, by looking at two specific timekeeping capacities, I will argue that temporal perception employs mechanisms that represent time in radically different ways. Therefore, no general story about how perceptual systems attribute temporal properties to perceived events is possible. Therefore,

---

<sup>14</sup>How mirroring theorists account for comparability is often unclear. Some, like Phillips (2012), introduce a type of internal clock to exploit the timing of perceptual processes. Others, like (Arstila, 2015), appeal to a comparator mechanism.

the standard accounts of the perceived unity of time fail. Instead, an account of the perceived unity of time must explain how the unity of time is constructed from initially fragmented time-keeping capacities.

The situation regarding temporal perception parallels what occurred in the literature on memory. Memory was initially understood as a single psychological capacity to retain information for later use. However, as research progressed, memory was no longer seen as a single capacity. Instead it was seen as various different psychological phenomena to be studied on their own terms (Craver, 2007). A general theory of memory was abandoned. Instead, theorists attempted to understand how specific forms of memory operate and how forms of memory might interact. The same goes for temporal perception. We cannot generalize from one time-keeping capacity to another. Instead, we must theorize about each capacity on its own terms and then uncover how these capacities interact.

This section describes ways in which specific timekeeping capacities can be selectively intervened upon while leaving other timekeeping capacities unaffected. The conclusion is that timekeeping capacities come apart along at least three different dimensions—timescales, modalities, and temporal-property-types.<sup>15</sup>

Pharmacological and mechanical interventions provide evidence for timescale specific divisions among timekeeping capacities. For instance, haloperidol and midazolam both impair temporal discriminations at around the one-second timescale, however, of the two, only haloperidol also impairs discriminations around 50 ms (Rammsayer, 1999). Similar dissociations are found through the use of rTMS. Applied to dorsal frontal areas, rTMS selectively impairs discriminations around one second (Jones, Rosenkranz, Rothwell, & Jahanshahi, 2004), while rTMS applied to the cerebellum selectively impairs discriminations in the millisecond range (Koch et al., 2007).

Psychophysics experiments show that timekeeping capacities can be selectively intervened upon along modality and temporal property-type dimensions. Consider first modality specific cases. It is well known that saccades distort the perception of the temporal properties of visual stimuli presented at the target location of the saccade during a short temporal window centered on saccade execution (Burr, Tozzi, & Morrone, 2007). When a single visual target is presented during this window at the appropriate location, subjects perceive the target as having a compressed duration. When a sequence-pair of stimuli is similarly presented, subjects often perceive a reversal of their objective temporal order. Importantly, saccades only influence the perception of visual stimuli. Auditory stimuli, for instance, presented alongside the visual ones do not undergo corresponding distortions.

Psychophysics also shows the selective manipulability of temporal property-type specific capacities, for example, capacities to perceive duration versus sequences, rates, and so forth. Consider *the oddball illusion* (Tse, Intriligator, Rivest, & Cavanagh, 2004). Subjects are initially presented with a series of *standard* stimuli that are identical with regards to their temporal properties (e.g., duration, ISI, etc.) and are of the same non-temporal type (e.g., if they are flashes of light, then they will be of the same color, intensity, etc.). After the presentation of the standard sequence, subjects are shown an *oddball*, which is identical to the standards with regards to its temporal properties, but differs in some salient non-temporal way (e.g., it might be a different colored light). Subjects reliably perceive the oddball as having a significantly longer duration than the standards (up to 50% longer).

An internal clock theorist could try and account for this dilation through an increase in the clock pulse-rate due to the novel oddball (this is what Tse et al., 2004 propose). However, if that

---

<sup>15</sup>For further evidence see Paton and Buonomano (2018).

were the case, then the other temporal properties of the oddball should be equally distorted, since the clock distortion would distort measurements of these other properties. To test this, Eagleman and colleagues conducted a version of the oddball study, reported in Eagleman (2008), in which the standards and the oddball flickered at a fixed rate. If the dilation resulted from an increase in the pulse-rate of a general-purpose clock, then the oddball should seem to flicker more slowly than the standards. However, the study showed that there was no effect of this sort. Only the perceived duration of the oddball was influenced, not its flicker rate. In this way, timekeeping capacities specialized for specific types of temporal properties can be selectively intervened upon.<sup>16</sup>

At this point, an important dialectical point needs emphasis. Any of these selective interventions could in principle be explained as resulting from some change in the operation of a single timekeeping mechanism or the inputs to those mechanisms. Perhaps this is most clear in the cases of selective distortions along modality specific lines. However, that explanatory strategy loses plausibility when we consider the full range of cases. Unless our goal was to salvage a centralized clock model, it is unclear why we should think that the inputs to a centralized clock would differ along timescale, modality, and temporal property type lines in the way these selective distortions would require. Furthermore, if there were a centralized clock (or clock network), then we would expect to find cases where subjects undergo a general disruption to their timekeeping capacities (in the same way the subjects may lose the ability to perceive faces, surface color, etc.). However, no cases like that exist.<sup>17</sup> Of course, nothing here demands that one abandon the idea that there is a centralized clock, however, this is not due to anything specific about temporal perception but rather, concerns the general underdetermination of theory by data. The resulting theory would become increasingly ad hoc to accommodate this evidence. Our best, least ad hoc, explanation, then, is one in which we take temporal perception to be fragmented.

This alone, however, does not show that the standard approaches to the perceived unity of time fail. These capacities could all employ a common code underpinned by SET-like timekeeping mechanisms. In what follows, we will focus on specific timekeeping capacities and show that the explanatory demands that are placed on models of those capacities give us good reasons for thinking that these capacities employ distinct types of representational mechanisms. As a result, no general account of how perceptual systems attribute temporal properties to events is possible, and therefore, no account of the perceived unity of time that relies on one can be succeed.

## 4.1 | Specific timekeeping capacities: Case #1 duration

Let us begin by considering duration perception at very short timescales. Classic approaches to this capacity have appealed to dedicated clock mechanisms, with SET being a prime example. Yet, emerging models are doing without dedicated clock mechanisms. Instead, they account for many rudimentary timekeeping capacities as arising from *intrinsic properties* of neural systems throughout the brain (Ivry & Schlerf, 2008; Paton & Buonomano, 2018). Intrinsic models have

---

<sup>16</sup>Johnston, Arnold, and Nishida (2006) also show selective distortions of modality and temporal property specific perception.

<sup>17</sup>Surveying the literature reveals no cases like this. One explanation for their absence, raised by a referee, is that impairing a centralized clock may result in an elimination of consciousness altogether. However, this would make temporal content unique amongst perceptual contents, since most, if not all, perceptual contents seem to be capable of being impaired while preserving consciousness. Without reasons for thinking temporal perception is unique in these ways, the absence of general “time blindness”, along with the evidence raised in this section, count against centralized clock approaches.

the advantage of providing a ready explanation for highly localized distortions in temporal perception and for why all sensory experiences seem to have some temporal content (i.e., because this content is provided by the same mechanisms that provides non-temporal content to experience) all while not having to posit any additional machinery over and above what is already in place for non-temporal capacities.

For the purposes of this paper, I will focus on a single type of intrinsic model, *state-dependent network models* (SDN models).<sup>18</sup> The choice of SDN models is meant to be illustrative of the general trend within the timekeeping literature as the same conclusions could be drawn from other intrinsic timekeeping models (e.g., Eagleman & Pariyadath, 2009; Lebedev, O'Doherty, & Nicolelis, 2008) or variants of the internal clock models (although, evidence in favor of the SDN models will be discussed). However time is encoded at short timescales will contrast sharply with how time is encoded for other timekeeping capacities (e.g., crossmodal temporal order perception).<sup>19</sup> The particular focus on the SDN models is largely due to their being particularly well-developed variants of these emerging intrinsic models.

According to SDN models, recurrent neural networks (RNNs) underpinning a wide-range of non-temporal capacities can also underpin rudimentary timekeeping capacities at very short timescales in the following way: Within each RNN we can distinguish between a system's *active states*, which are the different spatial distributions of spiking activity within the network, and the *hidden states*, which are the modulatory states of the system that control how the active states develop over time.<sup>20</sup> As a RNN receives input, a particular pattern of active states will unfold as a function of the incoming signal and the initial hidden states of the network. As time passes, the system's active states change as a function of the modulatory influence of the hidden states. As a result, at any given moment there will be a particular subset of neurons within the RNN that are most active which provides a spatial code for duration without the need for a dedicated clock.

Evidence for SDN models comes from a variety of sources. First, as a proof of concept, the time-dependent activity patterns described by SDN models have been found in artificial RNNs and with in vitro neural populations (Finnerty, Shadlen, Jazayeri, Nobre, & Buonomano, 2015; Goel & Buonomano, 2014).

Second, SDN models have accurately predicted a novel pattern of variability in temporal perception at very short timescales that are only accommodated by other models through the inclusion of post-hoc assumptions. This variability arises when subjects are presented with a pair of stimuli and are tested to see how they perceived the duration of the second stimuli (Buonomano & Karmarkar, 2002; Spencer, Karmarkar, & Ivry, 2009). In one condition, the ISI between the stimulus pairs was held constant, in the second condition, the ISI varied. Subjects showed an increased variability in the perceived duration of the second stimuli in the varied-ISI condition when the stimuli had durations of less than 150 ms. Internal clock accounts did not predict that variability of this sort would be restricted to these timescales. SDN models did. Since the particular dynamics of a RNN is not only a function of the incoming sensory signal but also of the RNN's state when the stimulus arrives, by varying the ISI between the two

---

<sup>18</sup>For details of SDN models see Buonomano (2000), Buonomano and Karmarkar (2002), Buonomano and Maass (2009), Paton and Buonomano (2018).

<sup>19</sup>Different coding schemes are found in ramping models (Lebedev et al., 2008), oscillation models (Kosem et al., 2014), and efficiency coding models (Eagleman & Pariyadath, 2009).

<sup>20</sup>See Buonomano and Maass (2009), Goel and Buonomano (2014) for details of the interaction between active and hidden states.

stimuli, the experiment varied this initial state in unpredictable ways resulting in varied perceptual performance.

No dedicated clock mechanism is needed. Instead, local networks have the intrinsic ability to keep track of some temporal properties of perceived events. Furthermore, since each RNN will have a different internal structure, each RNN will employ a different spatial coding for temporal information. Finally, since the ability of any RNN to carry information about time is due to its local stimulus history, it follows that there is no common interval that all RNNs encode information about. Rather, RNNs look backwards different distances depending on their local stimulation histories.

One criticism that has often been raised against SDN models, and intrinsic timekeeping models more generally, is that it is unclear how to scale SDN models for longer timescales and crossmodal timekeeping (Ivry & Schlerf, 2008). However, once we accept that temporal perception is fragmented, this objection loses its force. All accounts of temporal perception need to provide a story of cross-timescale and cross-modal integration. This is not unique to intrinsic models.

## 4.2 | Specific timekeeping capacities: Case #2 temporal order

Here we will shift focus from duration perception to crossmodal order perception. Models of this capacity must meet significantly different demands than ones for duration perception. In particular, any account of temporal order perception must accommodate the perceptual system's ability to rapidly recalibrate the perceived temporal order of events. To see what this recalibration is like consider the following two cases.

First, consider a study by Stetson, Cui, Montague, and Eagleman (2006).<sup>21</sup> Subjects were asked to press a button and then after a variable delay, with an average length of 35 ms, a flash of light would appear on the screen in front of them. Subjects had to respond whether the button press occurred before or after the flash of light. If the light followed the button press by approximately 20 ms subjects would be equally likely to report the button press as occurring earlier than or later than the flash of light. This provided the baseline *point of subjective simultaneity* (PSS) at which the two stimuli appeared simultaneous. A delay was then inserted between the button press and the flash of light such that the light appeared on average 135 ms after the button press. After several trials the PSS shifted to where the visual stimulus had to follow the button press by 44 ms to be perceived as simultaneous. The shifting PSS already showed that there was some recalibration in perceived temporal order, but the interesting finding came when the extended delay was abruptly replaced with the original 35 ms delay. Stimulus pairs, with an ISI of 35 ms, that were originally perceived as involving a flash of light after the button press were now reliably perceived as though the flash of light occurred before the button press! Perceived temporal order was reversed despite there being no change in stimulus timing.

For the second case consider a study by (Köse, Gramfort, & van Wassenhove, 2014). Subjects were presented a pair of stimuli—a pulsing sound and light. Both stimuli pulsed at the same frequency—1 Hz—but were slightly out of phase with one another. Subjects initially perceived the stimuli as being out of sync, however, subjects quickly began to perceive the two stimuli as being in phase, despite there being no change in the incoming stimulation.

---

<sup>21</sup>For similar findings using other sensory-sensory pairings, see Chen and Vroomen (2013), Navarra, García-Morera, and Spence (2012), Vroomen and Keetels (2010).

Two general accounts were proposed to explain these types of effects. In the first, which is readily accommodated by the internal clock and mirroring approaches, recalibration arises through a shift in the timing of our sensory processes. The auditory and visual pulses seemed to synchronize with one another through a synchronizing of the auditory and visual processing of these stimuli. The second proposal explains these recalibration effects without any shift in the timing of sensory processes. Instead, a representational mechanism for attributing temporal relations to the perceived events is recalibrated.

To decide between these accounts, a critical test was performed combining the recalibration studies with imaging methods to determine whether shifts in perceived temporal order corresponded with shifts in the timing of sensory processes. Interestingly, the two cases brought about conflicting results. In the Stetson et al. study, there was no shift in the latencies of tactile or visual processing. Instead, there was increased activity in the anterior cingulate cortex and medial frontal cortex (regions the authors suggest are involved in conflict monitoring).<sup>22</sup> In the Kosem et al. study, however, there was a corresponding shift in the timing of auditory and visual processes. Prior to calibration, ERP showed that the auditory and visual processes were oscillating in step with the oscillations of their respective stimuli. However, after calibration, the oscillations in the visual and auditory processes fell into phase with one another. While the explanation for the different effects is not clear, it may have to do with the types of sequences used in both studies. The Stetson et al. study used non-rhythmic sequences, while the Kosem et al. study used rhythmic sequences. Since there are reasons for thinking that rhythmic and non-rhythmic sequence perception engage distinct networks (Grahn & Brett, 2009), the different imaging results might be the result of the rhythmic/non-rhythmic difference.

Given the results of the Stetson et al. study, there must be an explanation of perceived temporal order, and its recalibration, that does not directly appeal to the timing of sensory stimulation or initial sensory processing. To explain these effects, researchers often posit a decision/comparator mechanism based on known perceptual opponency mechanisms (Cai, Stetson, & Eagleman, 2012; McDonald, Teder-Sälejärvi, Russo, & Hillyard, 2005; Roach, Heron, Whitaker, & McGraw, 2011).

The model from Cai et al. (2012) will serve as a nice example of this sort of mechanism.<sup>23</sup> A series of delay tuned neurons respond to particular temporal asynchronies between motor and visual processing. Think of these as a series of neurons tuned with Gaussian response profiles centered on these specific asynchronies, for example, motor-leading-visual by 50 ms, 30 ms, 20 ms, 0 ms, -20 ms, and so forth. These delay tuned neurons feed excitatory and inhibitory signals to a pair of summation nodes. One node receives primarily excitatory signals from motor-leading neurons and vice versa. This differential input produces opponent behavior in these summation nodes. In a calibration neutral state, the differential activity in these nodes simply reflects asynchronies in motor-visual processing. If the activity in the two nodes is identical, then the motor and visual events are represented as simultaneous. If there is a difference in activation, then a sequential order is represented, and the relative activity of the nodes encodes the length of the separating interval.

In models like this, recalibration can occur in several ways. Adaptation can directly influence the summation nodes, directly changing the encoding of temporal order, or adaptation

---

<sup>22</sup>Similar results were found using MEG and EEG (Simon, Noel, & Wallace, 2017; Stekelenburg, Sugano, & Vroomen, 2011).

<sup>23</sup>There are different approaches to explaining these recalibration effects, however, they all distinguish the timing of perceptual processes with represented temporal content which is all we need (Chen & Vroomen, 2013).

can influence the delay-tuned neurons, influencing the inputs to the summation nodes. Alternatively, recalibration can also result from a change in the timing of initial sensory processes, which will result in differences in the activation of the summation nodes. The model provides the resources needed to account for both the Stetson et al. and Kosem et al. studies (although, we need not insist that a single model accounts for all temporal order perception and recalibration).

The important thing to notice, however, for our purposes is that the mechanisms proposed to account for these aspects of temporal order perception latch onto their content in ways that differ from how SDN or SET models latch onto their content. Furthermore, not only do they exploit different representational strategies to get at their respective properties in the world, they exploit different properties of neural systems to encode this information. In both cases temporal information is given a spatial code—the distribution of activity across a population encodes the relevant information—however, the mappings from spatial patterns to temporal contents differ. A consumer system capable of using the information concerning an interval length encoded in one of these timekeeping systems need not be in any position to use the information encoded in the other. This is even the case if the consumer system is causally sensitive to the activity in both networks. Causal sensitivity is not enough to make use of the information. Consumer system must be able to decode these causal influences.

At this point we can discharge the argument. Temporal perception is fragmented. It is not a single capacity, but is instead composed of various specialized timekeeping capacities. When we try and account for these distinct capacities, not only do the capacities themselves seem to demand different things of the models that would account for them, but the models of these capacities that are currently being developed describe mechanisms that latch onto and encode temporal information in a variety of ways. As a result, the standard approaches to the perceived unity of time are bound to fail. No general story is forthcoming for how perception attributes temporal properties to perceived events and existing empirical evidence suggests that no such story is possible. Another strategy is needed to explain how the perceived unity of time emerges from our initially fragmented timekeeping capacities.

## 5 | PARALLELS BETWEEN TIME AND SPACE

One plausible means of making progress on understanding how temporal information is integrated across and within modalities is to look at spatial perception. In both cases, we seem to perceive the world as consisting in unified or seamless dimensions within which perceived events and objects are located. Furthermore, these aspects of perception depend on the integration of information initially encoded in multiple representational mechanisms. In this section, I will argue that despite their superficial functional similarities, explanations of how spatial information is integrated within and across modalities cannot be applied to the temporal case. To show this, we will begin by focusing on the perception of visual space before turning to the multimodal case.

The visual system parses the incoming retinal signal through a series of specialized filtering mechanisms that preferentially respond to specific stimulus features such as orientation, direction of motion, depth, color, and so forth. Many of the cortical systems responsible for processing these features have a map-like retinotopic structure where adjacent locations in the cortical maps encode information corresponding to adjacent retinal locations (Gardner, Merriam, Movshon, & Heeger, 2008; Wandell, Dumoulin, & Brewer, 2007). Despite this

initial feature segregation, we nevertheless perceive these features as being located within a common space. Visual space appears unified. While there is disagreement over the details, the general story for how spatial information is integrated in early vision is largely accepted. The story appeals to two aspects of *how* spatial information is encoded in retinopic maps (Robertson, 2003).

First, since each map shares a common retinopic structure, and therefore represents the same visual space, the process of integrating spatial information in vision is one of coordinating the different retinopic maps. An analogy for thinking about this integration is to think of coordinating map layers in a computer program. As long as the map layers share a common format, then integrating them can be accomplished by functionally superimposing one on the other. All that's needed are common landmarks, or anchor points, across the maps to line up the layers.

Second, simultaneous activity across these maps guides their coordination. The visual system exploits the assumption that simultaneous activity patterns across maps results from co-instantiation of features by objects. For instance, in a simplified case, if there are simultaneous spikes of activity in the color map and the motion map, then the visual system will behave as though there is a common colored and moving object rather than independent instantiations of color and motion. It is these simultaneity activity patterns, that at least in development, act as anchor points for coordinating the distinct visual feature maps. In many accounts, this coordination produces mappings between feature maps and a retinopic master map in which bound feature groups (or objects) are constructed (Koch & Ullman, 1985; Robertson, 2003).

Comparability, for visual space, is given by the common retinopic structure of the visual maps, and localization is given by the coordination of these maps. Notice, that an account of how visual features are represented does not provide an account of the unity of visual space. These representations must be integrated.

However, neither of the structural features of spatial representations exploited by models of spatial integration have analogs in the temporal case. First, there is no analog to retinotopy across the different timekeeping mechanisms in perception.<sup>24</sup> Second, simultaneity of sensory processes cannot be used as anchor points to coordinate the different timekeeping mechanisms. As we saw, in some cases the temporal contents of perception come apart from the temporal structure of perceptual processes. Sensory processes with identical temporal structure can represent events as standing in radically different temporal relations. Therefore, the story given for the integration of spatial information in vision cannot be applied to temporal perception.

The multimodal case is similar. There is a unity to our multimodal perception of space—the various senses locate objects within a common space around the individual. It cannot be the case that this integration is entirely accounted for by exploiting a common map-like structure (let alone retinopic structure), since different sensory modalities utilize different structured representations of space. A further complication is that there are multiple systems for integrating spatial information across modalities.<sup>25</sup>

Of particular interest to us is the role of the parietal cortex in the representation of personal space (i.e., the space immediately surrounding the body) (Sereno & Huang, 2014).

---

<sup>24</sup>Holcombe (2015) makes similar points. Retinopic maps carry information about a shared spatial visual space and retinopic structure facilitates the utilization of this information. That initial sensory areas may carry temporal information through resemblance is not enough to establish a parallel to retinotopy.

<sup>25</sup>Not discussed here are the superior colliculus (King, 2004) and entorhinal cortex (Soman, Muralidharan, & Chakravarthy, 2018). Both cases exploit map-like spatial representations that have no temporal analogs.

Despite this system integrating information across modalities, the explanation of this integration parallels the explanation in the visual case. Neurons in the postcentral gyrus (along with other areas) integrate multisensory signals where their responses to multisensory signals from a given location exceed the sum of the activation elicited by single sensory signals. This integration is thought to occur through a two-stage process. First, individual sensory modalities involve modality specific map-like representations of space. Then, through an exploitation of simultaneity across these maps, correspondences are formed between the modality specific maps and the multimodal map in parietal cortex (Bernasconi et al., 2018).<sup>26</sup> We have essentially the same story as we did in the visual case, the only difference is that the initial modality specific spatial maps are differently structured. However, simultaneity across cortical maps is essential to the process.

The point of this section is not to deny that simultaneity will play a role in the integration of temporal information. Neural integration is largely a story of the temporal and spatial convergence of neural signals. However, in the temporal case, simultaneity plays a semantical role. The various maps being integrated represent *how the world is now*. As a result, simultaneous activity across these maps can produce a complex representation of the world *right now*. However, in the temporal case, there is no clear-cut relationship between the timing of sensory processes and their temporal content. Yet, that is what would be needed to exploit simultaneity in the same way as it is exploited in spatial integration. Some other, content sensitive, explanation is needed for the integration of temporal information.<sup>27</sup>

## 6 | TOWARDS THE PERCEIVED UNITY OF TIME

Standard explanations of the perceived unity of time fail. We also cannot simply import models of spatial integration to the temporal domain. The empirical evidence described so far also shows that we cannot account for the perceived unity of time by simply introducing further clock mechanisms. Internal clocks track the temporal properties of internal mental processes. Yet, the temporal structure of perceptual processes comes apart from their temporal contents. Whatever account of the perceived unity of time we give must be sensitive to the *temporal contents* of perception and not merely the timing of perceptual processes. We need something new.

While it is not clear what allows for the construction of the perceived unity of time, this is for future interdisciplinary research, what we can do is articulate the hurdles that must be overcome in accounting for the perceived unity of time from our initially fragmented timekeeping capacities.

### 6.1 | Temporal localization

Standard approaches attempted to explain localization by appealing to the timing of sensory processes. The central reason why these approaches failed is that the perceiving temporal location of events comes apart from the timing of sensory processes. Simply adding more attributive machinery will not help either. We may come to know that some event has a certain duration

---

<sup>26</sup>The development of the superior colliculus may give insight as to how modality specific maps are manipulated to establish this correspondence. In that case, retinotopic maps in early development influence the structure of auditory and haptic maps (Doubell, Skaliora, Baron, & King, 2003).

<sup>27</sup>Simultaneity might play this role integrating *some* timekeeping mechanisms, but it cannot do it for the full range needed to account of the perceived unity of time.

and that some event stands in certain relations to other events, but none of this specifies *when* the event occurs relative to the current moment.<sup>28</sup> Something must anchor events, and the properties attributed to them, to particular moments in time.

Carlos Montemayor (2013) argues that an indexical component contributed by the perceptual mechanisms for detecting simultaneity plays this anchoring role.<sup>29</sup> Our perceptual system is capable of representing events, separated by as much as 240 ms, as occurring simultaneously. Montemayor argues that this *temporal window of integration* not only represents events as being simultaneous but indexes them as occurring *now*. In this way, perceptual timekeeping mechanisms possess a referential/indexical function in addition their attributive functions, that locates events in time relative to *now*.

Montemayor's proposal is clearly on the right track. Its success does not require a universal mapping between the temporal contents of perception and the timing of perceptual processes—there need only be one system that contributes this indexical component. However, it leaves unexplained why the world appears to have a single temporal dimension as opposed to multiple. We indexically locate ourselves as located *here*, but the world appears to have three spatial dimensions. Something beyond a mere indexical is needed to account for the dimensionality of localization. Providing an explanation of this referential component and the dimensionality of temporal perception is the first hurdle that must be overcome.

## 6.2 | Comparability and translation

Comparability requires something other than what is required by localization. An account of comparability must provide translation procedures by which individual consumer systems can use the temporal information carried in various formats by different timekeeping mechanisms.

One possibility is that there is a single code, either an amodal code or the coding scheme for one modality takes priority, into which the distinct temporal representations are translated, which consumer systems then utilize. Another possibility is that consumer systems might have their own propriety codes, suited for their particular needs (e.g., motor control vs. lexicalization), into which they translate various temporal representations. These are open empirical possibilities for future research required to explain the perceived unity of time. Furthermore, how we account for this translation might answer the temporal version of Molyneux's question (e.g., whether or not there is a common code for temporal information and whether or not the capacity to integrate this information is innate or acquired).

Once again, we find an aspect of the perceived unity of time that cannot be solved by introducing further clock mechanisms. Whatever account of translation we provide must be one that while operating on the local non-semantic properties of neural systems, must nonetheless respect their semantic content.

A final point is needed. There are three broad-stroke options we can adopt to explain the perceived unity of time. Unity might depend on a single unified representation of time that integrates the information encoded in peripheral timekeeping mechanisms. This unitary representation of time may then play the causal/function role of a unified experience of time that is utilized by consumer systems, including those involved in introspection. Therefore, despite perceptual processes that underpin temporal perception being fragmented, temporal experience

---

<sup>28</sup>The similarities with Perry (1979) are intended.

<sup>29</sup>Maniadakis and Trahanias (2016) provides a similar account. The same worry applies to their view as well.

might be unified (under a certain reading of unity). Another option is that fragmentation extends beyond mere perceptual processing and applies to experience as well. Again, there may be a single representational system that integrates information in peripheral timekeeping mechanisms, but this representation only provides experience with multimodal and cross-timescale temporal content. Other temporal content may be contributed through the operation of the peripheral timekeepers. Knocking out this central integrator would result in selectively knocking out experiences of integrated temporal properties while leaving non-integrated temporal experiences intact.<sup>30</sup> Alternatively, there may simply be *no single* integrated representation of time and we account for the perceived unity of time without a single place where it all comes together. The consumer systems that drive time sensitive behaviors may employ various integrated representations of time. As long as they are coordinated, then reflective and behavioral responses will appear coherent. On this construal, while time may strike us as unified in the ways that I described, there may not be a single unified representation of time on the basis of which time strikes us so.<sup>31</sup>

## 7 | CONCLUSION

The world appears to have a unified temporal structure. The observation of the world appearing this way may tempt us to believe that temporal perception, or experience, is unified or unitary. However, this is not so. Temporal perception is initially fragmented. The perceived unity of time is constructed from these fragmented capacities. However, at present there is no theory that explains this unity. Something new is needed. Specifying what this is requires new work.

## ACKNOWLEDGMENTS

This paper benefitted from feedback from many people. I would like to thank members of Bence Nanay's group at the University of Antwerp for incredible feedback. I would also like to thank audiences at the 2018 Central APA, University of Milan, University of Tuebingen, York University, and Idaho State University. I also want to give special thanks to Todd Ganson, Zachary Kofi, Emma Esmaili, and two anonymous referees for their comments.

## ORCID

Gerardo Viera  <https://orcid.org/0000-0002-3183-2294>

---

<sup>30</sup>Deficits of integrated temporal processing in autism may be clinical cases of this sort (Stevenson et al., 2014).

<sup>31</sup>This last option is inspired by Dennett and Kinsbourne (1992) and their multiple drafts model (MDM) of consciousness, however, the positions are not identical. They developed their MDM in contrast with what they called *the Cartesian theater* according to which there is a place in the brain where processing or information becomes conscious. They use crossmodal temporal experience to argue that there is *no moment or place* where information becomes conscious. Rather, at any moment there are multiple representations of the world, and depending on context, some of these representations will gain *fame in the brain* and become conscious. Similarly, on this construal of the perceived unity of time, there may be multiple representations of how events are temporally structured, and some of these may gain *fame* by being utilized by downstream processes. Yet, this claim about integrated temporal information is neutral about what makes any representation conscious. It may be that for any one of these integrated representations to become conscious it must “enter” the *Cartesian theater*. See (Arstila, 2016) for arguments that issues about temporal perception do not obligate us to adopt any specific theory of consciousness.

## REFERENCES

- Arstila, V. (2015). Defense of the brain time view. *Frontiers in Psychology*, 6, 1350.
- Arstila, V. (2016). The time of experience and the experience of time. In B. Mölder, V. Arstila & P. Øhrstrøm (Eds.), *Philosophy and psychology of time* (pp. 163–186). Cham: Springer International Publishing.
- Bayne, T. J. & Chalmers, D. J. (2003). What is the unity of consciousness? In A. Cleeremans (Ed.), *The unity of consciousness*. Oxford: Oxford University Press.
- Bernasconi, F., Noel, J.-P., Park, H. D., Faivre, N., Seeck, M., Spinelli, L., ... Serino, A. (2018). Audio-tactile and peripersonal space processing around the trunk in human parietal and temporal cortex: An intracranial EEG study. *Cerebral Cortex*, 28(9), 3385–3397.
- Buonomano, D. V. (2000). Decoding temporal information: A model based on short-term synaptic plasticity. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 20(3), 1129–1141.
- Buonomano, D. V. & Karmarkar, U. R. (2002). How do we tell time? *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, 8(1), 42–51.
- Buonomano, D. V. & Maass, W. (2009). State-dependent computations: Spatiotemporal processing in cortical networks. *Nature Reviews Neuroscience*, 10(2), 113–125.
- Burr, D., Tozzi, A. & Morrone, M. C. (2007). Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nature Neuroscience*, 10(4), 423–425.
- Cai, M., Stetson, C. & Eagleman, D. M. (2012). A neural model for temporal order judgments and their active recalibration: A common mechanism for space and time? *Frontiers in Psychology*, 3, 470.
- Chen, K.-M. & Yeh, S.-L. (2009). Asymmetric cross-modal effects in time perception. *Acta Psychologica*, 130(3), 225–234.
- Chen, L. & Vroomen, J. (2013). Intersensory binding across space and time: A tutorial review. *Attention, Perception, & Psychophysics*, 75(5), 790–811.
- Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford: Oxford University Press, Clarendon Press.
- Dainton, B. (2000). *Stream of consciousness: Unity and continuity in conscious experience*. Abingdon: Routledge.
- De Corte, B. J. & Matell, M. S. (2016). Interval timing, temporal averaging, and cue integration. *Current Opinion in Behavioral Sciences*, 8, 60–66.
- Dennett, D. C. & Kinsbourne, M. (1992). Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences*, 15(2), 183–201.
- Doubell, T. P., Skalióra, I., Baron, J. & King, A. J. (2003). Functional connectivity between the superficial and deeper layers of the superior colliculus: An anatomical substrate for sensorimotor integration. *Journal of Neuroscience*, 23(16), 6596–6607.
- Eagleman, D. M. (2008). Human time perception and its illusions. *Current Opinion in Neurobiology*, 18(2), 131–136.
- Eagleman, D. M. & Pariyadath, V. (2009). Is subjective duration a signature of coding efficiency? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1525), 1841–1851.
- Finnerty, G. T., Shadlen, M. N., Jazayeri, M., Nobre, A. C. & Buonomano, D. V. (2015). Time in cortical circuits. *The Journal of Neuroscience*, 35(41), 13912–13916.
- Foster, J. (1991). *The immaterial self: A defence of the cartesian dualist conception of the mind*. Abingdon: Routledge.
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- Gardiner, J. M., Atema, J., Hueter, R. E. & Motta, P. J. (2014). Multisensory integration and behavioral plasticity in sharks from different ecological niches. *PLoS One*, 9(4), e93036.
- Gardner, J. L., Merriam, E. P., Movshon, J. A. & Heeger, D. J. (2008). Maps of visual space in human occipital cortex are retinotopic, not spatiotopic. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 28(15), 3988–3999.
- Gibbon, J., Church, R. M. & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences*, 423, 52–77.
- Goel, A. & Buonomano, D. V. (2014). Timing as an intrinsic property of neural networks: Evidence from in vivo and in vitro experiments. *Philosophical Transaction of the Royal Society B*, 369(1637), 20120460.
- Grahn, J. A. & Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex*, 45(1), 54–61.

- Grondin, S. (2003). Sensory modalities and temporal processing. In H. Helfrich (Ed.), *Time and mind II: Information processing perspectives* (pp. 61–77). Cambridge, MA: Hogrefe & Huber Publishers.
- Hartcher-O'Brien, J., Brighouse, C. & Levitan, C. A. (2016). A single mechanism account of duration and rate processing via the pacemaker–accumulator and beat frequency models. *Current Opinion in Behavioral Sciences*, 8, 268–275.
- Holcombe, A. (2015). The temporal organisation of perception. In J. Wagemans (Ed.), *The Oxford handbook of perceptual organisation*. Oxford: OUP Oxford.
- Ivry, R. B. & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences*, 12(7), 273–280.
- Javadi, A. H. & Aichelburg, C. (2012). When time and numerosity interfere: The longer the more, and the more the longer. *PLoS One*, 7(7), e41496.
- Johnston, A., Arnold, D. H. & Nishida, S. (2006). Spatially localized distortions of event time. *Current Biology*, 16(5), 472–479.
- Jones, C. R. G., Rosenkranz, K., Rothwell, J. C. & Jahanshahi, M. (2004). The right dorsolateral prefrontal cortex is essential in time reproduction: An investigation with repetitive transcranial magnetic stimulation. *Experimental Brain Research*, 158(3), 366–372.
- Jones, L. A. & Wearden, J. H. (2003). More is not necessarily better: Examining the nature of the temporal reference memory component in timing. *The Quarterly Journal of Experimental Psychology B, Comparative and Physiological Psychology*, 56(4), 321–343.
- King, A. J. (2004). The superior colliculus. *Current Biology*, 14(9), R335–R338.
- Klink, P. C., Montijn, J. S. & van Wezel, R. J. A. (2011). Crossmodal duration perception involves perceptual grouping, temporal ventriloquism, and variable internal clock rates. *Attention, Perception & Psychophysics*, 73(1), 219–236.
- Koch, C. & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, 4(4), 219–227.
- Koch, G., Oliveri, M., Torriero, S., Salerno, S., Lo Gerfo, E. & Caltagirone, C. (2007). Repetitive TMS of cerebellum interferes with millisecond time processing. *Experimental Brain Research*, 179(2), 291–299.
- Kösem, A., Gramfort, A. & van Wassenhove, V. (2014). Encoding of event timing in the phase of neural oscillations. *NeuroImage*, 92, 274–284.
- Lebedev, M. A., O'Doherty, J. E. & Nicolelis, M. A. L. (2008). Decoding of temporal intervals from cortical ensemble activity. *Journal of Neurophysiology*, 99(1), 166–186.
- Lee, G. (2014). Temporal experience and the temporal structure of experience. *Philosophers' Imprint*, 14(3), 1–21.
- Maniadakis, M. & Trahanias, P. (2016). When and how-long: A unified approach for time perception. *Frontiers in Psychology*, 7, 466.
- Matell, M. S. & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Cognitive Brain Research*, 21(2), 139–170.
- McDonald, J. J., Teder-Sälejärvi, W. A., Russo, F. D. & Hillyard, S. A. (2005). Neural basis of auditory-induced shifts in visual time-order perception. *Nature Neuroscience*, 8(9), 1197–1202.
- Mellor, D. H. (1981). *Real time*. Cambridge: Cambridge University Press.
- Montemayor, C. (2013). *Minding time: A philosophical and theoretical approach to the psychology of time*. Leiden: Brill.
- Navarra, J., García-Morera, J. & Spence, C. (2012). Temporal adaptation to audiovisual asynchrony generalizes across different sound frequencies. *Perception Science*, 3, 152.
- Paton, J. J. & Buonomano, D. V. (2018). The neural basis of timing: Distributed mechanisms for diverse functions. *Neuron*, 98(4), 687–705.
- Penney, T. B., Gibbon, J. & Meck, W. H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *Journal of Experimental Psychology. Human Perception and Performance*, 26(6), 1770–1787.
- Perry, J. (1979). The problem of the essential indexical. *Noûs*, 13(December), 3–21.
- Phillips, I. (2012). Attention to the passage of time. *Philosophical Perspectives*, 26(1), 277–308.
- Phillips, I. (2014). Experience of and in time. *Philosophy Compass*, 9(2), 131–144.
- Pinel, P., Piazza, M., Le Bihan, D. & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, 41(6), 983–993.

- Poppel, E. (1988). *Mindworks: Time and conscious experience*. San Diego, CA: Harcourt Brace Jovanovich.
- Rammesayer, T. H. (1999). Neuropharmacological evidence for different timing mechanisms in humans. *The Quarterly Journal of Experimental Psychology B, Comparative and Physiological Psychology*, 52(3), 273–286.
- Rashbrook, O. (2013). An appearance of succession requires a succession of appearances. *Philosophy and Phenomenological Research*, 87(3), 584–610.
- Richardson, L. (2014). Space, time, and Molyneux's question. *Ratio*, 27(4), 483–505.
- Roach, N. W., Heron, J., Whitaker, D. & McGraw, P. V. (2011). Asynchrony adaptation reveals neural population code for audio-visual timing. *Proceedings. Biological Sciences*, 278(1710), 1314–1322.
- Robertson, L. C. (2003). Binding, spatial attention and perceptual awareness. *Nature Reviews Neuroscience*, 4(2), 93–102.
- Sereno, M. I. & Huang, R.-S. (2014). Multisensory maps in parietal cortex. *Current Opinion in Neurobiology*, 24, 39–46.
- Simon, D. M., Noel, J.-P. & Wallace, M. T. (2017). Event related potentials index rapid recalibration to audiovisual temporal asynchrony. *Frontiers in Integrative Neuroscience*, 11, 8.
- Soman, K., Muralidharan, V. & Chakravarthy, V. S. (2018). A model of multisensory integration and its influence on hippocampal spatial cell responses. *IEEE Transactions on Cognitive and Developmental Systems*, 10, 637–646.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73(4), 971–995.
- Spencer, R. M. C., Karmarkar, U. & Ivry, R. B. (2009). Evaluating dedicated and intrinsic models of temporal encoding by varying context. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1525), 1853–1863.
- Stekelenburg, J. J., Sugano, Y. & Vroomen, J. (2011). Neural correlates of motor–sensory temporal recalibration. *Brain Research*, 1397, 46–54.
- Stetson, C., Cui, X., Montague, P. R. & Eagleman, D. M. (2006). Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron*, 51(5), 651–659.
- Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M. & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *The Journal of Neuroscience*, 34(3), 691–697.
- Tse, P. U., Intriligator, J., Rivest, J. & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics*, 66(7), 1171–1189.
- Vroomen, J. & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics*, 72(4), 871–884.
- Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, 7(11), 483–488.
- Wandell, B. A., Dumoulin, S. O. & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, 56(2), 366–383.
- Watzl, S. (2012). Silencing the experience of change. *Philosophical Studies*, 165(3), 1009–1032.
- Wearden, J. H., Edwards, H., Fakhri, M. & Percival, A. (1998). Why “sounds are judged longer than lights”: Application of a model of the internal clock in humans. *The Quarterly Journal of Experimental Psychology Section B*, 51(2b), 97–120.
- Zhang, H. & Zhou, X. (2017). Supramodal representation of temporal priors calibrates interval timing. *Journal of Neurophysiology*, 118(2), 1244–1256.

**How to cite this article:** Viera G. The perceived unity of time. *Mind & Language*. 2020; 1–21. <https://doi.org/10.1111/mila.12331>