

VISUAL ASYNCHRONY & TEMPORALLY EXTENDED CONTENTS

PHILIPPE CHUARD

Southern Methodist University

Temporal experiences, according to retentionalism, essentially have temporally extended contents: contents which represent distinct events at distinct temporal locations, and some of their temporal relations. This means, retentionalists insist, that temporal experiences themselves needn't be extended in time: only their contents are. The paper reviews an experiment by Moutoussis and Zeki, which demonstrates a colour-motion visual asynchrony (§2): information about motion seems to be processed more slowly than information about colour, so that the former is delayed relative to the latter. This, the paper argues, raises an important difficulty for retentionalism and its account of the temporal ontology of experiences: it suggests that a central background assumption about neural processing presupposed by retentionalism is false, at least in cases of visual asynchrony. The paper then explores the general significance of this result for retentionalism (§3).

THE ringtone blaring from the next table—"It's raining tacos"—seemed to last minutes, as did your excruciating auditory experience. It isn't just that the relative *durations* of these two events (the ringtone and your experience of it) appear to match: the ringtone consists in a progression of notes and your auditory experience seemingly unfolds through similar successive phases.

To suggest as much is a mistake, according to an influential account of *temporal experiences*—experiences, that is, of the *temporal features* of perceived events, such as their duration, order, succession, etc. Perceptual experiences have a representational content: just as the contents of thoughts or beliefs, whereby a belief that the French Revolution ended in 1814 needn't last twenty-five years, contentful psychological states more generally don't usually instantiate the temporal properties they represent, and the same goes for perceptual experiences:

Contact: Philippe Chuard <pchuard@smu.edu>

[T]he experience of *A* followed by *B* is backward-looking. That is, it occurs with the experience of *B*, all in one go, but it represents the temporally extended period of *A*'s preceding *B*. (Tye 2003: 88)

[. . .] experiences do not themselves have temporal structure that mirrors the temporal structure represented in their content. [. . .] for example, a temporal order experience does not involve first experiencing one event and *then* experiencing a second event [. . .]; it involves experiencing both events *at the same time*, even though they are experienced *as* happening at different times. (Lee 2014b: 1)¹

The view has received various labels: I'll use "retentionalism", as it seems to have become rather well entrenched by now.²

At the heart of retentionalism, then, lies the denial of any essential isomorphy or similarity between (a) the *temporal properties* of the events we perceive and (b) the *temporal properties* of our experiences of those events. That is, if perceived events—like a ringtone—are (i_a) extended in time with (ii_a) a given duration, composed of (iii_a) distinct temporal parts—the succession of notes in the ringtone—unfolding in (iv_a) a given order, retentionalists deny that an experience of the ringtone has to be (i_b) extended with (ii_b) roughly the same duration as the ringtone, or that (iii_b) the experience is to be segmented into successive temporal parts—corresponding to those of the ringtone—(iv_b) ordered in the same way.³

In contrast, some theories of temporal experience acknowledge or even exploit various temporal isomorphies of this sort. Some merely insist there are temporal similarities between (a) and (b) simply as a matter of phenomenological observation. Some go a little further, advancing different kinds of dependence relations between (a) events and (b) experiences thereof to explain why a given isomorphy holds.⁴ But there is a range of theories putting some relations of temporal isomorphy to work in order to account for the distinctive phenomenology of temporal experiences, say, or how temporal experiences present or represent events and their temporal properties, with different types of explanations

1. See also Broad (1923; 1925), Grush (2005; 2006; 2008), Lee (2014a), Pelczar (2010), and Strawson (2009).

2. Following Dainton (2017). Lee (2014a; 2014b) prefers "atomism"—unfortunately, that term can also be used for a view exclusively about the *mereology* of experience (see Lee 2014c; Chuard 2022) which needn't entail the distinctive commitments of retentionalism.

3. See Lee (2014a: 7–8) for different kinds of temporal isomorphy; Chuard (2020: §9.2.2) for some of the different theoretical roles they might play.

4. See, e.g., Locke (1689/1975: II, 14, §§6–8, 12), von Helmholtz (1910/1962: 22), James (1890/1952: 627–28), and Chuard (2017), for causal accounts.

put forward to that effect.⁵ In its negative strand, then, retentionalism rejects any such temporal isomorphy as required by an explanation of how we experience the temporal properties of perceived events, since retentionalists typically deny such isomorphies hold in the first place.

The central feature of perceptual experiences, the positive strand in retentionalism goes, is their *representational content*: it's possible, retentionalists insist, to represent an extended event with a certain duration and temporal structure *without* the experience itself being extended or structured in time (no successive temporal parts, no order)—indeed, it's in principle possible for the instantiation of such temporal contents in consciousness to be strictly instantaneous, on such a view.

The aim of this paper is to question one crucial assumption behind the retentionalist view. The focus won't be—not directly, at least—on the temporal properties of experiences *per se*, or any temporal isomorphy between experiences and the events they represent. Instead, the paper is concerned with *some* of the temporal properties of the *neural processes* responsible for the experiences in question and their content. Here's why: suppose the ringtone lasts about 3 seconds, starting with a *Re*, ending with a *Sol*. Retentionalism has it that both notes, separated by a little less than 3 seconds, can figure in the same *temporally extended content*, which becomes conscious “all at once”, as it were. This means that *both* notes are consciously heard at the same time (while being represented *as* successive). The implication, then, must be that neural processes behind your consciously hearing the first *Re* take longer than the processing underpinning your hearing the last *Sol*: about 3 seconds longer, that is. Hence, retentionalism must presuppose that the neural processes behind our conscious sensory experiences can be partially delayed in the following manner:

partially delayed processing (PDP): if some perceived event x occurring at $t-n$ and some later event y occurring at $t-n^*$ can both figure in the content of a conscious perceptual experience e occurring at t , then the neural processing NP_x prompted by x at $t-n$ and which results in experience e at t begins earlier—and lasts longer—than the underlying neural processing NP_y of y from $t-n^*$ to t .

That is, the neural processes behind the perceptual representation of an earlier stimulus in some temporally extended content are delayed *relative* to the pro-

5. Explanations of this sort are typically associated with forms of *extensionalism*, according to which temporal experiences must at the very least be *temporally extended* to present or represent extended events: see, e.g., Dainton (2000; 2008; 2014a; 2014b), Hoerl (2012; 2013), Phillips (2014a; 2014b), Rashbrook-Cooper (2017), Soteriou (2010; 2013).

cessing of a later stimulus figuring in the very same (token) content—conversely, the latter processing is “compressed” in time *relative* to the former.⁶

There are reasons to doubt such an assumption holds, however. The argument proceeds in two steps. First, I review certain empirically documented cases where the assumption of *partially delayed processing* (PDP) appears to be false: not only does the assumption lead to some implausible consequences, but it delivers the wrong results in the cases under consideration (§2).⁷ The second step aims to generalize the argument beyond the specific cases reviewed in §2: if *partially delayed processing* (PDP) is false, then a somewhat different temporal constraint must constrict how neural processing of perceptual stimuli unfolds in time, and we can ask what the most minimal but plausible deviation from *partially delayed processing* might amount to (§3). The obvious problem is that, unsurprisingly, such a temporal constraint undermines the central retentionalist claim about the *temporal ontology* of experiences. As a result, the retentionalist denial of temporal isomorphy between experiences and perceived events is jeopardized too, since the latter is grounded in the retentionalist proposal about the temporal ontology of experience. What’s more interesting, on the other hand, is that the resulting picture of temporal experiences remains perfectly compatible with the main considerations retentionalists advance to motivate their view.

1. Retentionalism

Three theses characterize the core of retentionalism—two concern the representational content of perceptual experiences, the third their temporal ontology:

Retentionalism

R1: *extended temporal contents:* the representational content of a perceptual experience can represent (a) a temporally extended event, (b) some of its distinct temporal parts, and (c) temporal features of, as well as temporal relations between, its temporal parts.

R2: *backward-looking contents:* such perceptual contents (d) do not represent all temporal parts of an extended event *as simultaneous*, but

6. Presumably, such partial delays admit of varying durations when it comes to temporally extended contents which represent more than just two non-simultaneous events.

7. I should note: while most of the evidence reviewed here pertains to vision, there seems to be no principled reason why the consequences drawn from such evidence cannot extend to neural processing in other sensory modalities too. And I’ve carried the discussion at that level of generality accordingly. Were it to turn out that such a generalization is unwarranted, this wouldn’t affect the main conclusions of the paper: owing to its commitment to PDP, retentionalism faces serious difficulties in its account of *visual* experiences with temporally extended contents.

represent (e) earlier temporal parts of the event represented as occurring before later temporal parts.

R₃: *unextended experiences*: a conscious perceptual experience with such an extended content needn't be temporally extended.

We *do* perceive the temporal features (duration, order, succession, etc.) of perceived events, the view assumes, and these R₁ to R₃ aim to offer an account of such experiences and their temporal content. The first thesis specifies the sort of temporal information conveyed in experience: it's not just that experiences represent events which are in fact temporally extended, but they are perceived *as such*, at least to the extent that distinct temporal parts of perceived events are represented at distinct temporal locations—that much is required for the perceptual representation of temporal relations between distinct events.⁸

The second thesis adds a claim about the *directionality* with which such temporal information is presented. In principle, R₁ is compatible with the suggestion that extended contents could, like expectations or predictions, be future-directed.⁹ One rationale behind R₂ is the idea that experiences and their representational contents causally depend upon the events they represent, and hence must succeed what they represent, if only by a short amount of time (Tye 2003: 88).

These first two claims, it's important to note, pertain to the contents of *perceptual* experiences *exclusively*: the retentionalist view distinguishes itself from a more traditional “memory account”, according to which memory plays a crucial role in explaining how we can become aware of the temporal features of perceived events (e.g., Reid 1855; LePoidevin 2007):

[. . .] the Atomist can reject Husserl's idea that temporal experience involves a special kind of “retentional” memory experience [. . .]. Atomic temporal experiences might involve just one kind of conscious perceptual experience, not differentiated between “retention” and “perception”. Moreover, an atomic temporal experience need not involve “retention” even in the weak sense that it involves retaining contents from immediately past experiences: a temporally extended content could include—perhaps exclusively—information about events that were not presented in any previous experiences. (Lee 2014a: 6)

8. How such locations are specified in the content of experience may vary between different versions—see, e.g., Lee (2014a; 2014b).

9. As some propose they are: see, e.g., Husserl (1991) and Grush (2005; 2006). Indeed, it's possible to add to retentionalism the claim that experiences have future-directed content as well, though most retentionalists seem dismissive of such a suggestion.

Lee isn't denying, of course, that we have various sorts of memories of very recent experiences—or that such memories accompany later perceptual experiences:¹⁰ only that they fulfil any explanatory function in the retentionalist account. This “no-dependence-on-memory” desideratum upon temporally extended *perceptual* contents can be phrased thus:

non-mnemonic temporal content: if a conscious experience e occurring at some time t represents an event y succeeding some event x , (i) e depends on sensory stimulation by x (at $t-n$) and y (at $t-n^*$), neurally processed so that (ii) *both* x at $t-n$ and y at $t-n^*$ can figure in the content of a conscious perceptual experience *for the first time* at t , and (iii) neither the representation of x nor that of y in e “retain the content” of an earlier conscious experience of x (at $t-n$) or y (at $t-n^*$).

Why insist on such a requirement? One reason simply goes back, it seems to me, to the central retentionalist idea that temporal experiences can be accounted for *solely in terms of their temporally extended contents*. For instance, Lee's argument for retentionalism—the “trace integration argument” (Lee 2014a: 13–16) to be discussed briefly in §3.4—rests on the undeniable observation that perceptual systems are equipped with mechanisms (such as visual motion detectors in MT) which process information from *successive* stages of a stimulus. This suggests there are mechanisms in the visual cortex or earlier which are perfectly apt to deliver the sort of temporally extended contents retentionalists posit: the retentionalist explanation, then, has no need for additional mnemonic processes to “retain [. . .] contents from immediately past experiences” (Lee 2014a: 6). This explains why, as Lee says above, retentionalism doesn't have to factor such contents into different components, one perceptual, the other mnemonic.¹¹

What conception of memory does the desideratum presuppose—what kinds of memories does it aim to rule out? Lee's formulation clearly states that what's being rejected as explanatorily irrelevant is any kind of memory “even in the weak sense that it involves retaining contents from immediately past experi-

10. Nor is he saying that an extended content “bears no relation to previous experiential contents” (Phillips 2018: 291). Retentionalists can allow that successive temporally extended contents are related in a variety of ways—e.g., by similarity, co-reference or semantic overlap, causal influence, etc. The only relation being denied here is that of conscious content-retainment through memory.

11. Another factor owes to the perceived fortunes of those more traditional “memory accounts” (see, e.g., Tye 2003: 87–88)—see also Dainton (2000: 125–27; 2017: §4.3), Hoerl (2009; 2013; 2017), Phillips (2010: 191–96; 2018: 288–89), for criticism of such accounts. That is, the difficulties facing traditional “memory accounts” might naturally drive retentionalists towards an alternative approach, free of any reliance on memory. Lee (2014a: 6) offers several other reasons—see Phillips (2018: 290–91) for discussion.

ences” (Lee 2014a: 6). This seemingly functionalist characterisation involves three features: the kind of memory at issue consists in (i) some *content*-involving mechanism, which operates (ii) on the content of conscious perceptual experiences that have just occurred (its inputs), and (iii) serves to “retain” them in consciousness (outputs). Presumably, “retaining” means that the content of a conscious experience remains conscious even once the experience has ended, or at least that it can make its way back into consciousness immediately after the initial experience.¹² Presumably, it’s also possible that the content thus retained *via* memory isn’t exactly identical to the content of the initial experience: it might largely overlap the original perceptual content, if the mnemonic content in question partially fades or degrades. Clearly, different kinds of memories might fall under the scope of this characterisation, including iconic memories, episodic memories of experiences just passed, as well as the sort of “retentions” Husserl (1991) posited.¹³ An advantage of such a characterisation, then, is that it promises to capture a shared trait of different kinds of memories, while ignoring their differences.¹⁴

The third thesis, R₃, concerns *experiences* themselves—*qua* psychological events or states—and their temporal ontology. Again, the underlying thought appears to be that, *insofar* as the phenomenology and content of our sensory awareness of temporally extended events can be fully accounted for *via* the sort of temporally extended perceptual contents specified in R₁ and R₂, there’s no reason to expect experiences with such contents to be extended in time, let alone mirror the temporal structure of the events they represent. It is this third thesis, then, which grounds the negative strand in retentionalism regarding temporal isomorphies—and thereby distinctively targets *extensionalist* approaches for whom temporal experiences must be temporally extended in order to present or represent the temporal features of perceived events.¹⁵

R₃’s formulation is often qualified along two dimensions: one temporal, the other modal. Temporally, first: perceptual experiences are in fact *strictly unextended in time*, one version of retentionalism has it:

12. Clause (iii), note, does not rule out suggestions that an earlier experience could causally impact a subsequent one—either by (a) speeding up the processing of the later experience, or by (b) somehow affecting its content (see Phillips 2018: 291)—since neither suggestion hangs on the *content* of the earlier experience *being retained in consciousness* afterwards.

13. For a brief but useful survey of different kinds of memories, see Palmer (1999: ch. 12).

14. See Lee (2014a: 6, n. 7; 2014b: 153). Phillips (2018) argues that any theory of temporal experiences relies in fact on memory in one sense or another. Phillips carries his discussion in terms, not of Lee’s characterization of memory *via* conditions (i) to (iii), but of his own, whereby memory involves “preserving past psychological success” (2018: 291).

15. See footnote 5 for references.

temporally restricted R₃: perceptual experiences of temporally extended events are *instantaneous*.

This is implied by another standard motivation for retentionalism:¹⁶

The Principle of Simultaneous Awareness (PSA): to be experienced as unified, contents must be presented simultaneously to a single momentary awareness (Dainton 2017: §3, p. 31).

Depending on what a “single momentary awareness” amounts to—specifically, whether such awareness is taken to occupy a single point, as it often seems to—PSA leads directly to a temporally restricted version of retentionalism.

Yet it’s not uncommon to grant that the relevant experiences do *in fact* have a little duration nonetheless:

[. . .] experiences, although they are not experiential processes, need not be instantaneous. An atomic experience might be a property instantiation enjoyed most fundamentally by a subject as they are over a short interval of time. Related to this, an atomic experience may be realized by an extended physical process (as I would argue all experiences are—see below). What makes this coherent is that the proper temporal parts of the realizing process need not themselves realize any experiences; in particular, they might be simply too short-lived. (Lee 2014a: 4)

Leading to a slightly weaker construal of R₃:

temporally relaxed R₃: perceptual experiences of temporally extended events have a relatively *brief duration*.

The other dimension along which R₃ can vary is whether distinct events (or temporal parts of events) figuring in temporally extended contents *must* be experienced at the same time. That is,

modally restricted R₃: perceptual experiences of temporally extended events have a content which *necessarily* represents successive events *simultaneously*—though not *as simultaneous*.

This version voices the main idea behind PSA, going back to Kant (1790/1980: 133), as well as Geach:

16. See, e.g., Dainton (2000: 133–42, 166; 2017: §3), James (1890/1952: 411), Miller (1984: 107–9), Phillips (2010: 179–80), and Tye (2003: 90).

Even if we accepted the view . . . that a judgement is a complex of Ideas, we could hardly suppose that in a thought the Ideas occur successively, as the words do in a sentence; it seems reasonable to say that unless the whole complex content is grasped all together—unless the Ideas . . . are all simultaneously present—the thought or judgement just does not exist at all. (Geach 1957: 104)¹⁷

Some retentionalists, on the other hand, opt for a slightly weaker proposal:

modally relaxed R₃: perceptual experiences of temporally extended events have a content which *can* represent successive events *simultaneously*—though not *as simultaneous*.¹⁸

This weaker version often serves to motivate the retentionalist denial that the delivery of temporal information in conscious experience *has to be* temporally extended and structured. It *needn't be*, the point goes:

Granted, I experience the red flash as being before the green one. But it need not be true that my experience or awareness of the red flash is before my experience or awareness of the green one. If I utter the sentence

The green flash is after the red flash,

I represent the red flash as being fore the green one; but my representation of the red flash is not before my representation of the green flash. In general, represented order has no obvious link with the order of representations. Why suppose that there is such a link for experiential representations? (Tye 2003: 90)¹⁹

Note that the *modally relaxed* version of R₃ might be construed differently, depending on how its modality (“can”) is interpreted. Indeed, it’s not entirely

17. See also Geach (1957: 104; 1969: 34), Anscombe and Geach (1961: 96), and for a useful discussion, Soteriou (2007; 2013). For the reference to Kant, see Dainton (2017: §2.4). For a contemporary echo of this point, see Lee (2014a: 14–16; 2014b: 154).

18. A word about the logical connections between these four readings of thesis R₃. The *temporally restricted* version entails the *modally restricted* one (as the *temporally relaxed* version is also meant to), but not *vice versa*, since the *modally restricted* version is compatible with the *temporally relaxed* version (simultaneity of presentation doesn’t imply such presentations are instantaneous). The *modally relaxed* claim is entailed by all three other construals, but is compatible with their respective negations.

19. See also Dennett and Kinsbourne (1992), Pelczar (2010), and Dainton (2014b) and Phillips (2010: 181–82) for discussion.

clear how the possibility in question should be understood. Perhaps, it's not even alethic, but epistemic: something like (M₁) "it is consistent with phenomenological data—or the nature of representation in general—that our temporal experiences are not in fact temporally extended or structured". On the other hand, if the modality is genuinely alethic, the modal claim at issue is meant to be stronger, I take it, than a *mere* logical or metaphysical possibility—in the sense, say, that even if *our* temporal experiences are *actually* extended and structured temporally, we find (M₂) some possible worlds containing experiences in some subjects (be they alien, or humans with a different psychological or neurophysiological make-up), which are not. Perhaps, the modality is (M₃) nomological, or at least (M₄) restricted to close worlds where human subjects are psychological and neurophysiological duplicates of actual humans.

These different interpretations, we'll see (§3.2), make a difference to whether the *modally relaxed* version of thesis R₃ is undermined by the considerations in the next sections. On the other hand, the *modally restricted* version of R₃ clearly appears to be falsified by such empirical considerations, as do the *temporally restricted* and *temporally relaxed* versions. This means that the negative strand in retentionalism is under threat: the retentionalist rejection of isomorphies between the temporal properties of experiences and those of the events experienced hinges on R₃, which relies on the assumption of *partially delayed processing* (PDP). If, however, the latter isn't empirically viable, this undercuts the retentionalist denial of temporal isomorphies.

2. Colour-Motion Visual Asynchrony

Now to the argument: I begin by reviewing one experiment which sheds some light on the timeframe in which stimuli and their different features are processed, and how this affects perceptual consciousness. I'll then explore its significance for the retentionalist view (§2.4).

2.1. Perceptual Latencies

The causal process whereby the occurrence of a stimulus in a subject's environment leads to a conscious experience of that stimulus takes time. Psychologists divide that interval into three rough periods:²⁰

20. See, e.g., Breitmeyer and Ögmen (2006), Holcombe (2009; 2015), Keetels and Vroomen (2011).

- *external latency*: the time it takes for a physical signal produced by the distal stimulus to reach the sensory receptors of a perceiving subject.
- *neural latency*: the time it takes for the physical signals produced by a subject's sensory receptors following proximal stimulation to "ascend the neural processing hierarchy", including different brain areas responsible for processing different features of the stimulus (Holcombe 2009; 2015).
- *online latency*: the time between completion of neural processing of a stimulus in central cortical areas and the occurrence of a conscious experience of that stimulus.²¹

There is evidence that different features of a stimulus can be processed at different speeds, and so incur differential *neural* latencies. For instance, differences in contrast intensity, or in chromatic brightness, can affect the speed at which dedicated neurons respond to such features: the more intense the stimulus, and the brighter the contrast, the faster the neuronal response. This can be observed early on in the lateral geniculate nucleus (LGN) of macaque monkeys, in their primary visual cortex, even in later stages of processing such as the superior temporal sulcus.²² There's also evidence that colour can be processed faster than motion.²³

2.2. *Differential Latencies in Vision*

In an influential experiment by Moutoussis and Zeki (1997a), the display consisted of squares moving up and down at an alternation rate of 1–2 Hz. As they moved, their colour also changed from green to red at the same alternation rate:

21. The distinction between neural and online latencies, note, doesn't precisely delineate which neural processes are to be included in each, and so where to draw the line precisely. A less rough characterization of the distinction would ultimately depend, I take it, on where the core neural correlates of consciousness (see, e.g., Chalmers 2000; Bayne & Howhy 2015; Koch 2004) are to be located, on what these core neural correlates exactly do (i.e., whether they function merely to (i) produce a representation of a stimulus feature, as opposed to (ii) making that representation accessible to consciousness, or both), and whether they should be included among the processes relevant to neural or online latencies (at least the latter, I presume).

22. See Allik and Kreegipuu (1998), Lee et al. (1990), Maunsell and Gibson (1992), Maunsell et al. (1999), Oram et al. (2002).

23. Note that the evidence from visual asynchrony to be presented below does not sit entirely well with some of the evidence regarding the greater speed of processing in the magnocellular pathway, see, e.g., Breitmeyer and Ogmen (2006: ch. 5), Maus et al. (2010: 487), which would appear to play a role in the processing of motion.

all squares simultaneously had the same colour and changed colour simultaneously, likewise with their direction of motion.²⁴

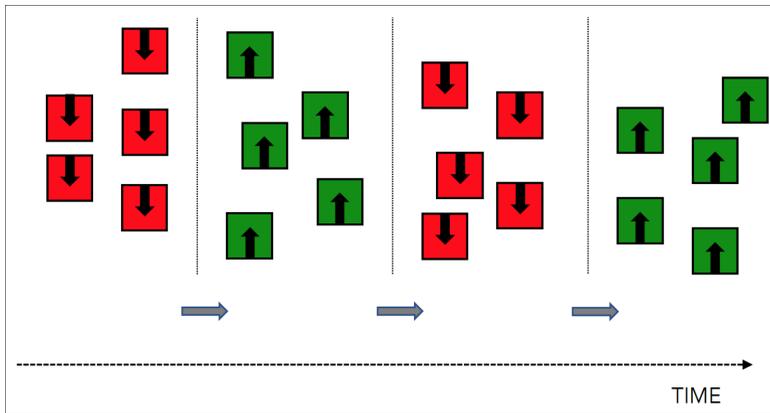


Figure 1: Illustration of two oscillations (in-phase trial) in Moutoussis and Zeki's (1997a) experiment.

The experiment included two series of trials: one where the period of oscillation for each feature took 565 ms, the other with oscillations at 706 ms (Moutoussis & Zeki 1997a; Bedell et al. 2003; 2006). In some trials, changes in motion direction and colour were *in phase* (i.e., the switch in both colour and motion was simultaneous), while in other trials they were not (the chromatic change occurred either midway through the motion, or so as to allow one colour to dominate through motion in one direction), and each combination was presented four times in mixed order (see Moutoussis & Zeki 1997a: 394, figure 2), with each trial consisting of ten repetitions. Subjects had to perform a *pairing (or matching) task*: that is, report the colour of the squares for a specified direction of motion, and *vice versa*.²⁵

24. FIGURE 1 represents just two oscillations and only serves to illustrate the general arrangement of the display. Moutoussis and Zeki (1997a) used between one and thirty coloured squares, presented against a black background on the computer screen. They note that the number of squares made no difference: the same results were obtained with just one or up to thirty squares—the point of using more squares was to “de-emphasize” shape perception.

25. Subsequent experiments varied in the following respects. First, while in the initial experiment, the *same stimulus* varied in colour and direction of motion (conjunctive stimulus), some experiments asked subjects to pair the colour and direction of motion of different stimuli (disjunctive stimulus), each on separate sides of a fixation point (Moutoussis & Zeki 1997b; Nishida & Johnston 2002). Second, some tested the asynchrony between colour and orientation (Arnold & Clifford 2002; Clifford et al. 2003). For more significant variations, see below.

Moutoussis and Zeki found that, for the squares to appear with a uniform colour when moving in one direction, the change in the direction of motion must in fact occur about 80 to 140 ms before the change in colour.²⁶ In other words, subjects systematically paired the direction of motion with the wrong colour:

Color-motion pairs perceived to occur simultaneously are not simultaneously present in the real world. By changing both the color and the direction of motion of objects rapidly and continuously and asking subjects to report which color-motion pairs were perceived as coexisting, it was found that motion was paired to the color present on the computer screen roughly 100 ms later (Moutoussis & Zeki, 1997a, 1997b). If there were no perception-time difference between color and motion, then the perceptual experience of the subjects should follow the reality occurring on the screen. These results thus suggest that the color of an object is perceived 100 ms before its direction of motion, a phenomenon we have called *perceptual asynchrony* and that we have attributed to the different processing-times necessary for the two functionally specialized systems to “finish” their corresponding jobs. These systems are thus not only processing but also perceptual systems, creating specific visual percepts in their own time and independently from one another. (Moutoussis 2014: 202–3)

The results vary somewhat as a function of the following factors:

- (i) *angle of direction*: if the direction of motion changes by less than 180 degrees (say the object’s trajectory shifts by 45 degrees only), motion needs to precede colour only by 80 ms.²⁷
- (ii) *relative luminance*: decreasing chromatic luminance reduced the effect by 25 ms; 40 ms when the luminance of the motion was decreased.²⁸
- (iii) *task*: the effect is observed when subjects are asked to report which colour they see when the object is moving, say, upwards (*pairing/matching task*). There are suggestions (Nishida & Johnston 2002; 2010; Bedell et al. 2003; Clifford et al. 2003) that the effect seems to disappear almost

26. Though this is the average typically cited in the literature, Moutoussis and Zeki (1997a: 396) report a delay between 78 and 84 ms for those trials with “best fit reversal” (i.e., when chromatic change and change in motion direction appear in phase to the subjects).

27. Amano et al. (2007), Arnold (2010), Arnold and Clifford (2002), Bedell et al. (2003).

28. Variations in intensity and luminance can affect other illusions such as the flash-lag effect: if the flashed target is dimmer, this can increase its delay relative to the moving target whereas, with the flashed target brighter, the delay can be shortened and even reversed—see Bedell et al. (2006), Purushothaman et al. (1998), Lappe and Krekelberg (1998), Ögmen et al. (2004), Patel et al. (2000).

entirely (reduced to about 6 ms) when subjects are tasked instead with reporting which *change* (colour or direction of motion) occurs first (*temporal order judgment*).²⁹

These results demonstrate that, owing to the differential neural latencies with which different features of a stimulus are processed, what are in fact non-simultaneous features (motion preceding colour) of non-simultaneous presentations of a stimulus (or of successive stimuli) appear to occur simultaneously in experience. This, we'll now see, has various implications for *online latencies*—the interval between the completion of neural processing for a representation of a given feature and when that feature becomes conscious.

2.3. Two Models

According to the *online model*,³⁰ a stimulus feature becomes conscious immediately after its processing by a dedicated neural mechanism is completed.³¹ Hence, differences in neural latencies between colour and motion result in the colour and motion of one stimulus (at a time) becoming conscious *at different times*, which serves to explain visual asynchrony. Because the direction of motion of an upward green stimulus takes longer to process (about 100 ms more, say), such a feature becomes conscious at the same time as the colour (red) of the *subsequent* downward stimulus, which is processed at greater speed. As a result, subjects report seeing a red square moving upward, followed by a green square moving downward.³²

29. Viviani and Aymoz (2001) reported an asynchrony of 50 ms with a temporal order judgment task—yet Nishida and Johnston (2010: 292) dismiss this as too small an effect. Nishida and Johnston (2002) *did* replicate the effect in several experiments, at an alternation rate of 250 ms, obtaining a delay between 100 ms and 150 ms when subjects were asked to report when colour and motion were “in phase” (Nishida & Johnston 2002: 360, figure 1; figure 3, p. 362). Oddly, it seems that, using temporal order judgments tasks, they found no effect *only* with an alternation rate of 2000 ms (figure 2, p. 361), or when only one dimension (colour or motion), but not the other, alternated at 250 ms.

30. Or “brain time” account: see Arnold (2005; 2010), Bartels and Zeki (1998; 2006), Holcombe (2015), Moutoussis (2012; 2014), Zeki (1993; 2003; 2015).

31. Or “almost immediately”: online latencies, according to the online model, could in principle amount to no latency at all (with 0 duration, that is). Still, it’s compatible with the online model that online latencies take a few milliseconds: what differentiates the online model from the “delayed perception” model below is (a) the relative size of the online latency, which (b) is the same for each feature on this model.

32. Nishida and Johnston (2002) dispute that visual asynchrony is explained by differential neural latencies between the processing of colour and motion. On their view, it owes instead to the fact that chromatic *changes* are first-order changes (dubbed “transitions”) while *changes* in direction of motion are second-order (“turning points”), which are associated with different time-markers

In contrast, Eagleman's "delayed perception" model has it that the last stages of processing leading to the generation of a conscious experience involve postdictive mechanisms, to the effect that the visual system "wait[s] for the slowest information to arrive" before "going live" — Eagleman suggests a delay of about 80 ms (2010: 223).³³ On this model, *online* latencies might outstrip *neural* latencies in the sense that the visual system generates a conscious representation only after all the synchronous features of a stimulus have been processed. But if the brain waits for the slower feature to be processed before generating a conscious experience of both, why expect any visual asynchrony at the conscious level, on this model? The timeframe is key: an 80 ms delay allows the most recent information about the direction of motion to be still missing *when* the processing of such information will be completed, say, 100 ms after chromatic information is processed. Hence the asynchrony: when neural processing of the colour (green) of the upward moving square is completed, the direction of its motion is still being processed, even after 80 ms, given a 100 ms neural latency. This is why, the suggestion must go, its colour (green) ends up being paired, inaccurately, with the direction of motion of the previous downward-moving red stimulus.³⁴

An additional possible difference between the two models concerns feature-binding in conscious experience. For the online model, different types of features (e.g., colour and motion) are processed separately and independently — as Moutoussis insists in the last sentence quoted earlier. And since, on this model, such features will become conscious as soon as their neural processing is completed, this means that the colour of a stimulus will "make it into" consciousness slightly before its direction of motion does, if the latter is processed more slowly. In other words, the online model suggests, cases of visual asynchrony provide some evidence that there is *no feature-binding* of the different features of a stimulus in conscious experience (following Zeki 1993). That is, there is strictly no single experience representing, say, a red square moving upward, on this version of the online model. Instead, an experience of redness and a distinct experience of upward motion occur more or less at the same time, which explains why subjects report them as simultaneous when engaged in a pairing task.

on the grounds that "turning points" take more time to be processed. But if temporal order judgment tasks produce little evidence of visual asynchrony, it's unclear why an explanation appealing to delays in tracking *changes* in colour and motion is offered to account for visual asynchrony (generated when subjects are *not* tasked with tracking such changes). For further discussion, see also Amano et al. (2007), Arnold (2010), Bedell et al. (2003; 2006), Clifford et al. (2003), Clifford (2010), Nishida and Johnston (2010).

33. The same duration as that posited by Eagleman and Sejnowski's (2000; see also Rao et al. 2001) postdictive account of the flash-lag effect — about which, see Maus and Nijhawan (2006; 2008), Maus et al. (2010), Nijhawan (2002), Ögmen et al. (2004), Patel et al. (2000).

34. For discussion, see Arnold (2010), Holcombe (2015), as well as Arstila (2019).

It might seem natural, on the other hand, to view the delay model as operating under the assumption that the different features of a stimulus *are* bound together in conscious experience: indeed, the point of the 80 ms online delay could be precisely to allow for feature-binding, whereby colour and motion, which are processed separately and at different speeds, are then bound together in a conscious representation of both features during the online delay. Subjects thus enjoy a *visual illusion* of a red square moving upward, which explains the pairing judgments subjects are tasked to report.³⁵

The little evidence there is to adjudicate between these models doesn't seem to support the delayed perception model, however.³⁶ Consider the illustration on the left of FIGURE 2:³⁷

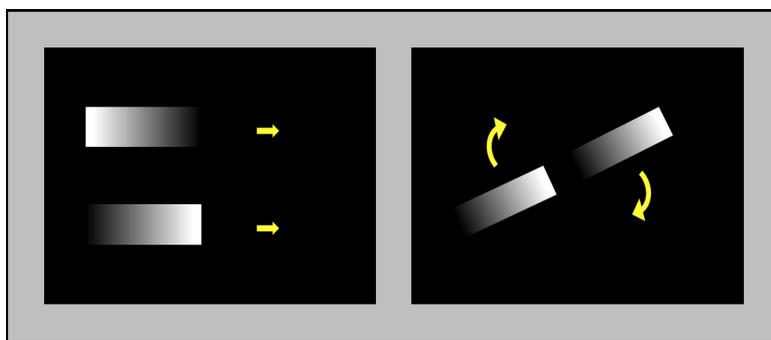


Figure 2: two versions of the Hess illusion.

35. For this “no binding” interpretation of the online model, see Bartels and Zeki (1998; 2006) and Zeki (2003; 2015). Such an interpretation isn't mandatory, however. The online model could allow, in principle, that feature-binding between colour and the direction of motion of an earlier presentation of the stimulus takes place once the colour is neurally processed and before a colour experience goes online. Conversely, on the delay model, with an online latency of 80 ms, the colour and direction of motion (of what are in fact successive stimuli) could become simultaneously, but *separately*, conscious, without being bound in a joint experience. Nevertheless, it's not clear, on such an interpretation of the delay model, what exactly an 80 ms delay allows the visual system to do (which the online model doesn't).

36. While there is some evidence of delayed online processing, Holcombe (2015: 829–30) argues it doesn't support the delayed perception model, since it concerns integration and compensation between vision and audition in light of the different “speeds” at which visual and auditory mechanisms process their dedicated stimuli (i.e., simultaneous visual and auditory stimuli can appear simultaneous despite the fact that auditory signals reach the auditory cortex faster than visual stimuli are processed in the visual cortex, by about 30 to 50 ms)—see Sugita and Suzuki (2003), Alais and Carlile (2005), Kopinska and Harris (2004). Note also that the evidence for such inter-sensory compensation has been questioned by Keetels and Vroomen (2011) and Van Eijk et al. (2008), while others have failed to replicate its effects: Arnold et al. (2005), Lewald and Guski (2004), Stone et al. (2001), Heron et al. (2007).

37. See Eagleman (2010: 222, figure 14.3). On the Hess illusion, see Wilson and Anstis (1969), and White et al. (2008)—compare Nijhawan (1992).

The rectangles are moving from left to right and back in synchrony—that is, in the same direction, at the same speed, occupying the same relative location at the same time. Whereas the top rectangle is brighter on its left side, the bottom rectangle has a brighter right side. If brighter stimuli are processed faster, one should expect the brighter side of each stimulus to be processed more rapidly than their darker side. So, were the online model to be correct, the reasoning goes, one should expect that, when both rectangles are moving to the right, the apparent length of the top rectangle should seem to contract somewhat while that of the bottom rectangle should seem to expand (Eagleman 2010: 221–23; Holcombe 2015: 825). The thought appears to be that, since the brighter side of each rectangle is processed faster than the other, it will become conscious earlier, if the online model is true. Thus, as it moves towards the right, the left-hand side of the top rectangle will become conscious before the right-hand side, and will retain this temporal advantage throughout its rightward motion. When the right-hand side becomes conscious a little later, it will appear at a location which, owing to its online delay, is closer to the location occupied by the faster processed left-hand side: closer, that is, than it in fact is, hence the apparent contraction. Similarly with the bottom rectangle: its brighter right-hand side will become conscious first, while the left-hand side is delayed and appearing at a location further apart from that of the faster processed right-hand side, thereby producing an apparent dilation of the stimulus when moving rightward.

Were such an illusion to be observed, advocates of both models seem to agree, it would rule in favour of the online model (Eagleman 2010: 221–26; Holcombe 2015: 825). Though the brighter side may be neurally processed faster than the other, the reasoning seems to go, the different neural latencies at play may still fall within an 80 ms online delay, for instance. Hence, according to the delayed perception model, both sides of the square could be fully processed before becoming jointly conscious following the 80 ms delay. That’s why there’s no reason to think, on the delay model, that one side of the square becomes conscious earlier than the other, and so no reason to expect any apparent contraction or expansion of the stimulus as a result.³⁸

Yet, it appears the sort of illusory experience the online model predicts does in fact occur. And while Eagleman (2010: 226) has conceded this much, he has tried to explain the illusion away as a function of the use of a neutral density filter to reduce contrast over parts of the screen. White et al. (2008) report having

38. Alternatively, if the neural latency between processing of the brighter side and that of the darker side goes over 80 ms, an 80 ms delay could nevertheless ensure that contraction and dilation effects are significantly minimized. For instance, with a neural latency of 100 ms, the delay model should predict that the darker side becomes conscious only 20 ms after the brighter side does (given the 80 ms online delay), compared with the 100 ms online latency between the two sides predicted by the online model.

produced the illusion *without* a density filter, however—in this case, a rotating Hess effect, similar to the illustration on the right of FIGURE 2, where the rectangles (or lines, in this case) look to curve—thereby suggesting that different parts of the figure with different degrees of illumination reach consciousness at different times, as the online model predicts.³⁹

2.4. Significance for Retentionalism

How does the evidence from visual asynchrony bear on the retentionalist view? At first sight, it might seem as though the online model is inimical to retentionalism: if stimuli become conscious very soon after being neurally processed, one might naturally expect that two successive events (e.g., two notes in a ringtone, or two successive presentations of the same red object) would become conscious successively, as soon as each is neurally processed. Conversely, the delayed perception model might look like a natural fit for retentionalism: an 80 ms delay could help explain how two successive events can nevertheless become simultaneously conscious in a temporally extended content, as retentionalism requires. If so, any evidence supporting the online model against the delay model (as in §2.3) ought to sound like bad news for retentionalism.

The situation is a little more complicated, though—and, in fact, worse for retentionalism. For even *if* the evidence had favoured the delay model rather than the online model, the former isn't really such a natural fit for retentionalism after all. In fact, evidence from visual asynchrony appears to undermine retentionalism in at least two respects, even when the “delayed perception” model is taken for granted.

To see this, it's important to keep two commitments of retentionalism clearly in sight. First, the “no-dependence-on-memory” desideratum for temporally extended perceptual contents:

non-mnemonic temporal content: if a conscious experience e occurring at some time t represents an event y succeeding some event x , (i) e depends on sensory stimulation by x (at $t-n$) and y (at $t-n^*$), neurally processed so that (ii) *both* x at $t-n$ and y at $t-n^*$ can figure in the content of a conscious perceptual experience *for the first time* at t , and (iii) neither the representation of x nor that of y in e “retain the content” of an earlier conscious experience of x (at $t-n$) or y (at $t-n^*$).

39. Similar illusions appear to support the online model, according to Holcombe (2015: 826): synchronous flashes that differ in luminance (Roufs 1963), or synchronous gratings that vary in contrast (Bernadete & Kaplan 1999; Stromeyer & Martini 2003), can generate apparent motion from the brighter/more brightly contrasted stimulus to the dimmer/less contrasted one.

Second, clauses (i) and (ii) in this desideratum have the following implication for the neural processing underpinning temporally extended perceptual contents:

partially delayed processing (PDP): if some perceived event x occurring at $t-n$ and some later event y occurring at $t-n^*$ can both figure in the content of a conscious perceptual experience e occurring at t , then the neural processing NP_x prompted by x at $t-n$ and which results in experience e at t begins earlier—and lasts longer—than the underlying neural processing NP_y of y from $t-n^*$ to t .

To repeat, retentionalism requires that non-simultaneous events represented at the same time in the same experience must have different neural or online latencies.

Yet, evidence from visual asynchrony suggests there are cases where this isn't so. For the evidence surveyed in §2.2 indicates, not just that (a) there doesn't seem to be any neural or online delays of the sort retentionalism needs, but that (b) for those neural latencies which do impact conscious experience, the resulting experiences lack a temporally extended content.

A brief summary of the evidence for (a) includes:

- (i) *no delay*: even with different *synchronous* features of a stimulus, when one feature (motion and its direction) has a longer processing latency than the other (colour), the brain doesn't appear to wait until the more slowly processed stimulus feature (motion) has been fully processed. Rather, each feature becomes conscious in its own time, with colour reaching consciousness before the direction of motion. This holds for the online model, as well as with the 80 ms delay posited by the delayed perception model—the differential online latency between colour and motion is simply much shorter (about 20 ms rather than 100 ms) on the latter.
- (ii) *constant processing*: if a colour of a given intensity is processed at a certain speed, it is processed roughly at that same speed through successive presentations of the stimulus (the same goes for the slower processing of direction of motion). This explains how the effect of Zeki and Moutoussis's experiment occurs *through successive presentations* of the display in a given trial (consisting of 10 repetitions): the upward direction of motion of an earlier presentation is wrongly paired with the red colour of a later presentation, followed by the downward motion of yet another earlier presentation being paired (wrongly again) with the green colour of a later presentation, *and so on*. There is no indication, that is, that processing

speeds or online delays somehow become shorter over successive presentations, as one would expect if *partially delayed processing* (PDP) were true.

As for (b):

- (iii) *no extended content*: in a pairing task, the direction of motion of an earlier stimulus (x at $t-n$) is reported as synchronous with the colour of a faster-processed later stimulus (y at $t-t-n^*$): though the motion of x at $t-n$ and the colour of y at $t-t-n^*$ seem to become conscious more or less at the same time (t), both features are judged (wrongly) to occur simultaneously.
- (iv) *delay not enough for extended content*: if the brain *did* wait an extra 80 ms before binding in consciousness the more slowly processed direction of motion of x at $t-n$ with the more quickly processed colour of y at $t-t-n^*$ (as the delay model predicts), such delay wouldn't suffice to guarantee that the resulting experience is representing two *successive* events or their features *as such*: the point of the delay model is to (1) explain how colour and motion are bound into a conscious representation of a simultaneous event (e.g., an upward-moving red square), where (2) the posited delay (80 ms) is short enough to explain how information about the direction of motion (e.g., the red square is in fact moving downward) arrives too late to be factored in such a conscious representation. Even on the delay model, that is, processing isn't *so delayed* as to allow that different *synchronous* features of the same stimulus (rather than successive ones) are represented in the same conscious experience—let alone that successive stimuli could be processed to become conscious simultaneously.

It isn't just that the evidence from visual asynchrony fails to line up with the main claims and commitments of retentionalism. The assumption of *partially delayed processing* (PDP), which the retentionalist postulation of temporally extended perceptual contents demands, seems inconsistent with what happens in cases of visual asynchrony.

Consider FIGURE 3, to begin with, which depicts the sort of experiences subjects appear to have in cases of visual asynchrony—with the distal stimuli at the bottom, the contents of the relevant experiences at the top (highlighted in grey), and the different processing latencies marked by the orientation and length of the different arrows in-between:⁴⁰

40. If the online and delay models do differ by their commitment to feature-binding, FIGURE 3 depicts the delay model only. On the online model, experience E_1 should be treated as two distinct roughly simultaneous experiences, one of red, the other of upward motion—likewise for E_2 and E_3 .

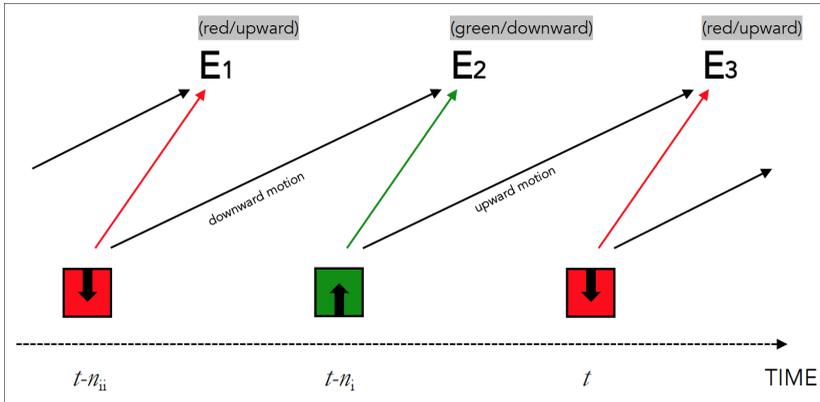


Figure 3: Standard model of experiential contents in visual asynchrony.

Note first that no model or explanation of visual asynchrony posits or needs to posit an experience at t (or thereafter) which represents *both* events x (stimulus presentation at $t-n_{ii}$) and y (stimulus presentation at $t-n_i$), let alone represents them as succeeding one another. The main difference between the online and delay models is how long after a stimulus is neurally processed do conscious experiences E_1, E_2, E_3 , occur: 80 ms or much less.

But why exactly couldn't the experiences depicted in FIGURE 3 have temporally extended contents instead? Can't a retentionalist simply re-describe the situation as in FIGURE 4: the subject's experience right after t (E_3) has a temporally extended content representing the stimulus' changes from $t-n_{ii}$ to t —that is, the downward motion of a green square, followed by the upward motion of a red square? FIGURE 4, that is, illustrates the kind of *partially delayed processing* (PDP) retentionalism would need to assume in this case:

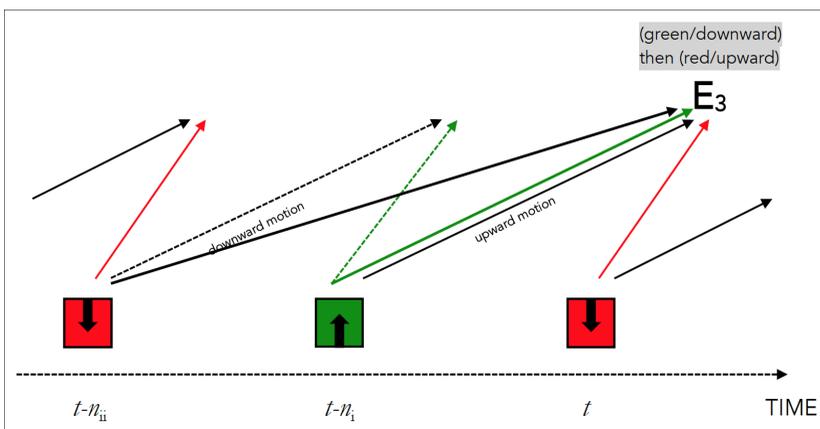


Figure 4: Experiential contents in visual asynchrony according to retentionalism.

There are three problems with this picture. First, for experience E_3 to represent both x (the stimulus at $t-n_{ii}$) and y (the stimulus at $t-n_i$) in a temporally extended content, online latencies would have to be significantly longer than what the evidence suggests. The online latency for the stimulus' downward motion at $t-n_{ii}$ would have to be significantly longer than the online latency for the stimulus' upward motion at $t-n_i$ in order for both to become conscious simultaneously after t (similarly with online latencies for the stimulus' green colour at $t-n_i$ and its red colour at t). The problem isn't just that there's no evidence that this is how online latencies unfold. The problem is that the evidence from visual asynchrony shows that such *partially delayed processing* (PDP) for the same type of features is in fact rather unlikely: processing of the direction of motion remains stable in such conditions, with a constant latency across successive presentations or repetitions. Otherwise, the asynchrony effect might not hold across the ten repetitions which constitute a given trial in Zeki and Moutoussis's experiment.

What's more, the online delays implied by such a suggestion should strike one as implausibly long. Suppose the delay in question encompasses processing of just two successive presentations of the stimulus (e.g., at $t-n_{ii}$ and $t-n_i$), even at the rapid alternation rates used in cases of visual asynchrony: between 250 ms between changes (in the version of Nishida & Johnston 2002) and up to 565 ms or even 700 ms (for Zeki & Moutoussis's 1997a version). Assume it takes only about 80 ms at the earliest for the second presentation of the stimulus at $t-n_i$ to become conscious (that is, assume a rather short neural latency with no added online latency, if only for the sake of argument): this version of retentionalism would still require processing delays of the first presentation of the stimulus at $t-n_{ii}$ to range between 330 ms and up to 780 ms, in order for an experience like E_3 in FIGURE 4 to have an extended content representing just two successive presentations of a stimulus at rapid alternations.⁴¹

Second, what about experience E_2 (right after $t-n_i$) which precedes experience E_3 (see FIGURE 3)? There's no reason why retentionalists should deny that an experience occurs at that time. But if the contents of E_2 and E_3 are both temporally extended (as in FIGURE 4) by virtue of *partially delayed processing* (PDP), the content of E_3 representing the succession [green downward moving square followed by red upward moving square] from $t-n_{ii}$ to t is likely to partially

41. For comparison's sake, standard estimations of how long neural processing altogether takes before a visual conscious experience occurs vary significantly between 80 ms and 200 ms: see, e.g., Fabre-Thorpe et al. (2001), Nowak and Bullier (1997), Thorpe et al. (1996), VanRullen and Thorpe (2001); as well as Koch (2004: ch. 15). In unusual circumstances, it could take up to 400 ms, as when processing is delayed *via* transcranial magnetic stimulation: see Herzog et al. (2016). For similar complaints regarding the unlikely length of the delay in delay models in general (*albeit* using different temporal estimates), including postdictive accounts, see, e.g., Arstila (2016: §4), Dennett and Kinsbourne (1992: 192), and Grush (2005: 215–16; 2008: 153).

overlap that of E_2 , since both have [green downward moving square] from the stimulus between $t-n_{ii}$ and $t-n_i$ as part of their content. The retentionalist view would then need to explain the status of that part of E_3 's content it shares with E_2 's. For if such content rests on neural processing already deployed for the same representation of the stimulus in an earlier experience (E_2), doesn't it violate the "no-dependence-on-memory" desideratum on temporally extended contents? That is, isn't the shared content [green downward moving square] from $t-n_{ii}$ to $t-n_i$ in fact mnemonically "retained" from E_2 to E_3 —and if not, why not?

Third, and more importantly, it becomes unclear why there should be any visual asynchrony in the first place. For if *partially delayed processing* (PDP) is true, it remains mysterious why, in experience E_3 (FIGURE 4), the upward motion of y at $t-n_i$ should end up being bound with the red colour of a *later* presentation of the stimulus at t , rather than being (accurately) bound with its own colour (green) at $t-n_i$ —*ditto* for the stimulus' features at $t-n_{ii}$. On both the online and delay models, what accounts for the asynchrony is the fact that there *isn't enough time* to complete neural processing of both synchronous features (i.e., colour and direction of motion at $t-n_i$) before one (colour) becomes conscious, owing to the longer neural latency of the other (motion direction). But given the longer neural or online latencies retentionalists need to posit in PDP, if the brain has all this additional time to produce a conscious experience of *successive* red and green moving squares, doesn't it thereby have ample time to fully process and accurately bind the synchronous colour and direction of motion of a given *synchronous presentation*? There's no reason to expect that cases of visual asynchrony arise, if *partially delayed processing* (PDP) is assumed, and no mechanism to explain why they nevertheless do.

Here, retentionalism faces the same difficulty the delayed perception model ran into: the latter, we saw, implies that illusions like the Hess illusion ought not to occur. The retentionalist has to posit even longer online delays, since such delays must encompass the completion of neural processing for several *successive presentations* or repetitions of the changing stimulus—whereas the 80 ms delay of the delayed perception model was meant to explain just *one* synchronous presentation of the stimulus at a time, not several. And if there's reason to rule out online delays of 80 ms, why think that longer online delays are somehow likelier?

The difficulties surveyed thus far arise in a specific set of conditions, where the alternations of the changing stimuli from green to red and moving up and down proceed at the rate of 1–2 Hz. It's in those conditions that evidence of visual asynchrony has surfaced. In the next section, I attempt to generalize the problem.

3. The Causal Argument

If *partially delayed processing* is false (at least in some cases), the question is what to replace it with—what might be the most minimal deviation from *partially delayed processing* (PDP) nevertheless compatible with evidence from visual asynchrony? To zero in on such a replacement, it may help to start with three related general assumptions about the mechanisms underpinning our perceptual systems: (a) perception of a stimulus essentially involves a *causal chain* from the occurrence of the distal stimulus in the subject's environment to a conscious experience of that stimulus, *via* proximal stimulation of the relevant sensory receptors and subsequent processing mechanisms, all the way to various dedicated cortical areas. Consequently, (b) such processes take time: there's a *temporal gap*, even if a very brief one, between the occurrence of a worldly event and its presentation or representation in conscious experience.⁴² This temporal gap is (c) relatively *constant, insofar as* the part of the causal chain beginning with the proximal stimulus (stimulation of a sensory organ) is concerned: even if not perfectly constant, there are constraints on the speed of processing throughout the perceptual system, to the effect that there's at least a lower and upper bound on how long it takes a stimulus to be processed and become conscious.

The constancy hypothesis, (c), is problematic for the retentionalist view, since it contradicts the assumption of *partially delayed processing* (PDP) retentionalism relies upon. The difficulty concerns what retentionalism says about the *temporal ontology* of experience—thesis R₃: namely, that experiences with temporally extended contents aren't, or needn't be, temporally extended or structured. Call this the "causal argument" against the retentionalist view. The argument should seem unsurprising: its interest lies mostly in exactly which retentionalist commitment(s) it undermines, which assumptions it leaves standing, and how retentionalists might try resisting it. Before outlining the argument, the constancy hypothesis needs some refinement.

3.1. Constant Processing

Is the temporal gap from proximal sensory stimulation to conscious experience really constant? After all, evidence from visual asynchrony reveals that the processing of different features of a stimulus (e.g., motion, orientation) is delayed by about 100 ms in comparison to others (e.g., brightness, colour).

42. With the obvious exclusion of familiar problematic cases, such as seeing the light from a faraway star. In any case, the main assumption about the causal and temporal structure of neural processing (c) concerns *neural* (as well as) *online* latencies, not *external* latencies.

Properly *contextualized*, the assumption that the temporal gap *is* constant remains consistent with such evidence. First, *relative* to those features (bright contrasts, more intense colours) that are neurally processed faster, there's no reason to think that the interval between proximal sensory stimulation and a resulting conscious experience doesn't stay constant *through successive presentations*. *Ditto* with those features processed somewhat less rapidly (motion and its direction, orientation, etc.). Second, though there is evidence that neural and online latencies vary between different sensory modalities,⁴³ the constancy hypothesis can be restricted to a specific modality. More importantly, the constancy hypothesis isn't at all a claim to the effect that different features, in different sensory modalities, can be neurally processed more or less at the same time. It's a claim, rather, about the relative speed and duration of neural processing *for successive presentations* of the same stimulus' features through the same sensory channels—it's in this sense that processing latencies are supposed to remain relatively constant *over time*.

Accordingly, different perceptible features within a sensory modality can be grouped in terms of the interval it takes for them to be neurally processed and reach consciousness. For each such group, one can obtain a different specific version of the following principle, the *causal constancy assumption*:

- (CCA) for a given sensory modality, and for some properties F, G, H, \dots of a stimulus x , perceived through that modality (and for some similar experimental settings w_1, w_2, w_3, \dots), the processing interval λ from proximal sensory stimulation caused by x to the onset of a conscious perceptual experience of F, G , or H, \dots is constant (across situations w_1, w_2, w_3, \dots).

Reference to “similar experimental settings w_1, w_2, w_3 ” merely serves to emphasise it is not enough to characterize such causal/temporal assumptions regarding neural processing solely in terms of the features being processed: the *conditions* in which they are perceived (e.g., whether adaptation can take place, what angular information about motion is presented, at what rate the features alternate, what stimuli they are preceded by, etc.) can make a difference too.⁴⁴

43. See note 36.

44. Including whether the features in question are attended or not, as attentional selection might make a difference to whether the processed features make it to consciousness at all, according to the view that attention serves as a gatekeeper for consciousness (i.e., no conscious perception without attention: see, e.g., Dehaene and Naccache (2001), and for a very helpful discussion of the evidence from inattentive blindness typically marshalled in support of such gatekeeping views, see Wu (2014: ch. 5; 2017). To circumvent this complication, I assume that, in CCA, the stimuli and their relevant features are fully attended—as seems to be the case anyway in experiments generating visual asynchrony.

Note also what talk of “constancy” amounts to in this context: it’s not just that the temporal gap λ after which a conscious experience of a given feature F occurs takes (i) *no less* than λ across successive proximal stimulations caused by F , but also that it takes (ii) *not much more* than λ in normal circumstances. Indeed, evidence from visual asynchrony doesn’t just support a conclusion about *how long* a conscious experience is delayed (or not) after proximal stimulation, owing to the *pace* of neural processing. It provides information about the *maximal relative length* of such a delay. For instance, in Moutoussis and Zeki’s (1997a) experiment, the evidence suggests it *always* takes longer for the visual system to process information about the direction of motion than about colour. But this delay (between 80 and 140 ms) isn’t just a lower bound: neural processing of the direction of motion is always behind the neural processing for colour *by more or less the same amount*. Significantly, this holds for *successive presentations of the same stimulus throughout a given trial* (in this case, ten repetitions, in the same conditions).

This means that λ needn’t be a strict or precise value, but a *limited range*: there is room for some variation between 80 and 140 ms, both across different trials carried out in different conditions, and through successive repetitions within a given trial. The latter, for instance, would allow that λ compresses somewhat over several successive presentations of the same stimulus with the same stimulus features presented in the same conditions. For example, processing of a green surface could well accelerate over the course of successive repetitions—thus, λ could be a little longer for neural processing at the onset of a stimulus compared to its offset. Such compression, however, still requires some minimum delay (λ ’s lower bound) in the processing of each successive presentation of the relevant stimulus feature.⁴⁵ As we’ll see, this is all the argument below really needs—*provided*, that is, that processing of a later presentation of the same stimulus feature isn’t *so* compressed as to become conscious *simultaneously* with an earlier one (in

45. See for instance Bachmann et al.’s (2003) *perceptual acceleration* model, posited to account for the flash-lag effect. In successive presentations of the same stimulus, the model assumes, processing of a later presentation in the succession can be “accelerated” by having a shorter online latency than earlier presentations (Bachmann 1989; Bachmann et al. 2003; Klotz & Wolf 1995). Bachmann et al. (2003: 285–88) estimate that acceleration starts within the first 100 ms of processing of a newly presented stimulus and lasts for 100–200 ms, after which online latency, though shorter than at stimulus onset, remains constant. Crucially, they do *not* assume that two non-simultaneous presentations of *the same stimulus* ever become conscious simultaneously (though non-simultaneous presentations—for up to 40 ms (Bachmann et al. 2003: 288)—of *different stimuli* seen in different conditions at different locations—one in a stream of similar stimuli, the other in isolation—might, to explain the flash-lag effect. In this case, however, the different stimuli appear as simultaneous, so there is no temporally extended content). For discussion, see Khurana and Nijhawan (1995), Nijhawan (2002).

which case, PDP holds, and processing wouldn't be constant at all, even in this moderate sense).

Finally, note that CCA doesn't presuppose the online or delay model. In particular, CCA can still be true if the delayed perception model is preferred: for even if the production of a conscious experience is delayed by about 80 ms to allow for the slower processing of other simultaneously presented features, 80 ms serves as a maximum too.

3.2. *The Argument*

Imagine seeing a red fire-truck cross your visual field at great speed—in a few seconds, it is gone. Bright red and seen in broad daylight, the fire-truck's colour is likely to be neurally processed relatively quickly. Its motion (not just the fact it is moving, but its trajectory through the scene) is the sort of temporally extended event one can perceive, according to retentionalism, *via* a temporally extended content. The example can be schematically described as follows:

- (i) at t_1 , the red fire-truck is located at L_1 .
- (ii) at t_2 , the red fire-truck is located at L_2 .
- (iii) at t_3 , the red fire-truck is located at L_3 .
- (iv) at t_4 , the red fire-truck is located at L_4 .
- (v) at t_5 , the red fire-truck is located at L_5 .

Suppose that each time— t_1 , t_2 , t_3 , t_4 , and t_5 —is separated from the next by about 200 ms, so that the interval from t_1 to t_5 lasts about 800 ms. Suppose also that the perceptual conditions from t_1 to t_5 (lighting, your visual attention focused on the truck, etc.) remain perfectly stable.

Next, apply CCA:

- (CCA) for a given sensory modality, and for some properties F , G , H , . . . of a stimulus x , perceived through that modality (and for some similar experimental settings w_1 , w_2 , w_3 , . . .), the processing interval λ from proximal sensory stimulation caused by x to the onset of a conscious perceptual experience of F , G , or H , . . . is constant (across situations w_1 , w_2 , w_3 , . . .).

It's quite possible that different instances of CCA, with different values for λ , need to apply: one for the colour of the fire-truck, another for its motion and its direction, or for its relative location, if these features are neurally processed at differ-

ent speeds. But since the colour of the fire-truck can be processed fairly rapidly, and is seen in the same perceptual conditions throughout, application of cca to the situation ((i)–(v)) delivers the following:⁴⁶

- (1) at $t_1 + \lambda$, S has a conscious visual experience of the fire-truck’s redness at L_1 .
- (2) at $t_2 + \lambda$, S has a conscious visual experience of the fire-truck’s redness at L_2 .
- (3) at $t_3 + \lambda$, S has a conscious visual experience of the fire-truck’s redness at L_3 .
- (4) at $t_4 + \lambda$, S has a conscious visual experience of the fire-truck’s redness at L_4 .
- (5) at $t_5 + \lambda$, S has a conscious visual experience of the fire-truck’s redness at L_5 .

This description of S ’s conscious experience through the relevant interval—(1) to (5)—can be summarized as follows:

- (6) from $t_1 + \lambda$ to $t_5 + \lambda$, S has a *succession* of different conscious visual experiences, each with a different representational content: each represents the fire-truck’s redness at a *different* location.

Obviously, this carries serious implications for the different interpretations of thesis R₃, the retentionalist claim about the temporal ontology of experience.

First, (6) clashes with the *temporally restricted* version of claim R₃—that perceptual experiences with temporally extended contents *are instantaneous*. Since neural processing takes time, and since the time it takes for a given feature F (colour) to be neurally processed in the same conditions is relatively constant, it takes no less and no more than λ for successive presentations of such a feature through t_1 to t_5 to be processed. In which case, successive presentations lead, *via* λ in cca, to a *succession* of *different* experiences. Even if neural processing of feature F *did shorten progressively* over the interval from t_1 to t_5 (becoming faster at $t_5 + \lambda$ than it was at $t_1 + \lambda$), the conclusion would still hold. For even a relatively inconstant and narrowing temporal gap λ^* would suffice to guarantee that an experience of the fire-truck’s colour from t_1 to t_5 occupies an interval from $t_1 + \lambda^*$ to $t_5 + \lambda^*$ and varies in content across that interval.

The *modally restricted* version of claim R₃—that the temporally extended content of a perceptual experience *must be* such that everything represented is represented *simultaneously*—is falsified too, for the same reason. Passage of a

46. The description in (1) to (5) contains a simplification: since λ represents the temporal interval between *proximal sensory stimulation* and conscious experience, $t_1 + \lambda$ omits the interval Δ for the external latency between the distal stimulus (the truck at location L_1) and proximal sensory stimulation (which I’m assuming is constant in this situation). Strictly speaking, “ $t_1 + \lambda$ ” should read “ $t_1 + \Delta + \lambda$ ” and so on.

bright red fire-truck in front of you from t_1 to t_5 entails, *via* CCA, that conscious experiences of the successive temporal parts of the passing truck—that is, the fire-truck at t_1 , and then at t_2 , and then at t_3 , etc.—are consciously experienced after roughly the same interval (λ). Hence, an experience of the truck’s redness at L_1 will become conscious at $t_1+\lambda$, followed at $t_2+\lambda$ by an experience of the truck’s redness at L_2 , and so on.

As a result, the *temporally relaxed* version of retentionalism—that a perceptual experience with a temporally extended content *can* have a brief extension but still lack temporal structure and *distinct* temporal parts—is also contradicted by (6). An experience of the fire truck from $t_1+\lambda$ to $t_5+\lambda$ will have different successive temporal parts with distinct contents (representing the truck’s redness at different locations).

Finally, whether the *modally relaxed* version of R3—that an experience *can* represent the successive temporal parts of a temporally extended event *simultaneously*—is incompatible with (6), depends on how the modality expressed by the relevant “can” is interpreted. If it’s meant to express a “mere possibility”—for example, that there *could* have been *some creatures* with unextended or unstructured experiences, as in (M_2)—then there’s no inconsistency between (6) and the modally relaxed version of R3. The fact that our experiences do not *actually* represent the successive temporal parts of extended events *simultaneously* doesn’t rule out that the experiences of other possible creatures—including possible humans with different kinds of brains—*could* have. Of course, such a possibility is rather uninformative in this context: so construed, the modally relaxed version of R3 may be true, yet tells us hardly anything about *our* actual experiences. If, on the other hand, the possibility in question is meant to be restricted to creatures *like us, with the kinds of brain we actually have*, in the conditions in which our actual brains usually operate—as in (M_3) or (M_4)—then it looks as though such a possibility *is* ruled out by (6): CCA is grounded in evidence about our actual neural processes, and implies that the perceived successive temporal parts of a moving fire-truck are *not* neurally processed so as to appear *simultaneously* in consciousness.⁴⁷

In short, every substantive interpretation of the third retentionalist thesis (R3) is undermined by CCA, the alternative to *partially delayed processing* (PDP).

47. As for (M_1) the purely epistemic interpretation of “can”—that simultaneous presentation is compatible with phenomenological data, say—it really depends on what evidence is brought to bear. Simultaneous presentation of successive temporal parts may well be compatible with phenomenal data—though see Dainton (2008; 2014a; 2014b) and Rashbrooke-Cooper (2017) for arguments that it isn’t, and Lee (2014b) for some responses—while being *incompatible* with physiological and behavioural evidence from visual asynchrony.

3.3. Response #1: Re-Description Strategies

As I said, the argument just articulated is hardly surprising. The interesting question is: is there room for retentionalists to resist it—without, that is, merely rejecting CCA by digging their heels in about PDP? A retentionalist could insist that most experiences in (1) to (5) above ought to be ascribed, at the very least, a temporally extended content. And indeed, nothing in the description from (1) to (5) above precludes the retentionalist from offering a more detailed description of the relevant experiences and their respective contents. For instance,

- (1) at $t_1+\lambda$, S has a conscious visual experience of the fire-truck's redness at L_1 .
- (2*) at $t_2+\lambda$, S has a conscious visual experience of the fire-truck's redness at L_1 & at L_2 .
- (3*) at $t_3+\lambda$, S has a conscious visual experience of the fire-truck's redness at L_2 & at L_3 .
- (4*) at $t_4+\lambda$, S has a conscious visual experience of the fire-truck's redness at L_3 & at L_4 .
- (5*) at $t_5+\lambda$, S has a conscious visual experience of the fire-truck's redness at L_4 & at L_5 .

This isn't the only way in which to sketch a proposal of this sort. Perhaps, some such contents are more extended than others: perhaps, the experience at $t_5+\lambda$ represents the whole sequence from t_1 to t_5 , say.

No matter how the details are worked out, the suggestion risks running afoul of the retentionalist's "no-dependence-on-memory" desideratum on temporally extended perceptual contents. For instance, the experience at $t_3+\lambda$ represents some of the same events (i.e., the truck being red at L_2) already represented in the experience at $t_2+\lambda$. Hence, the content of the later experience (3*) appears to "retain" part of the content of the preceding one (2*) in consciousness (in the sense specified by Lee—§1).⁴⁸

Worse, what the retentionalist ends up with, in any case, is still an extended experience temporally structured by way of its distinct temporal parts—*contra* R3—owing to the *distinct* extended contents in the succession from (1) to (5*). What this shows, interestingly, is that claims about the temporally extended contents of perceptual experiences—that is, theses R1 and R2 in the retentionalist view—can be severed from thesis R3 about the temporal ontology of experience,

48. In other words, retentionalism threatens to collapse into a traditional "memory account" simply by the fact that successive temporally extended contents which partially overlap thereby appear to meet the characterization of mnemonic processes ruled out by retentionalists (§1). This difficulty arises, note, when PDP is replaced by CCA, but also—as we saw (§2.4)—even when it isn't.

once *partially delayed processing* (PDP) is replaced by something like CCA. This serves to highlight the rather crucial role PDP plays, even if usually left unarticulated, in tying together the different strands of the retentionalist view.

3.4. Response #2: *Integrated Processing of Temporal Features*

Another facet of the causal argument worth noting: CCA does not clash in any way with the more compelling considerations advanced to motivate the retentionalist view. Hence, it seems such considerations are impotent in blocking the argument in §3.2.

For example, one of the most interesting and promising attempts in this regard is Geoffrey Lee's (2014a; 2014b) appeal to Reichardt detectors (2014a: 14): a simple model of a processing mechanism to explain motion perception, composed of one detector tracking edges at one location in the visual field and another detector tracking edges at another location. The first detector will be triggered by a moving stimulus first, followed by the second detector when the stimulus reaches the area it is sensitive to. But the information processed by the first detector must somehow be delayed until the second detector has been stimulated and its output processed, so that both pieces of information can be "integrated" in order to represent motion:

Suppose we have temporal information at the periphery that is contained in a temporally extended pattern of receptor stimulation, say on the retina. In order for the representation of this information to causally impact post-perceptual processing, each relevant temporal part of the initial extended stimulation has to *leave a trace* in the brain. If you consider the process leading up to, say, a verbal report of the information, you can see that each of these traces will have to be *simultaneously present* before the report is made: the alternative is that traces from certain parts of the stimulus no longer exist, and therefore can have no causal impact on the report. Furthermore, the traces will have to be *integrated* in the right way for simultaneous representation of each temporal stage of the input to be able to control later processes in an appropriate way: if each relevant stage of the input leaves a trace but the traces are in completely different neural populations that aren't functionally integrated, the information is not present in a useful form; it is not "explicit" in the relevant sense. For example, you can see in the motion detection example that if the representation of motion is going to have an appropriate later effect, the triggering of the first detector has to cause a trace that is integrated with the trace from the triggering of the second detector. *Models of temporal com-*

putation implicitly assume that setting up such traces and then simultaneously integrating them is the task that the brain has to perform. (Lee 2014a: 14–15)⁴⁹

Grant Lee’s very plausible rationale about the need for informational integration, as well as the rest of his argument. What does this show *à propos* CCA? Very little, I think.

True, if temporal features such as motion and duration *are* represented in conscious sensory experiences, the neural processing of such features must involve integration of locational information (e.g., that x is at L_1 at t_1 , that x is F at L_1 at t_1 , etc.) from successive stages of a moving or extended stimulus. And such integration will undoubtedly eventuate in some processing delay of the earlier stages. But notice: this is a point about *neural* latencies, that is, about the necessary processing requirements for the *representation* of *certain types* of features (indeed, it can serve to explain why motion and its direction incur longer neural latencies than, say, colour). It is *not, in and of itself*, a point about *online latencies*—about how and *when* our brains generate *conscious experiences*, which is what CCA is concerned with.

Thus, integration of locational information from successive presentations of a moving stimulus *for the purpose of representing its motion* will entail a processing delay. Yet, *during* the delay in question, the very same locational information (e.g., that x is at L_1 at t_1 , or that x is F at L_1 at t_1) can continue to be processed further in the neural hierarchy. This means that, in principle, such locational information could well become conscious *even before* its integration for motion representation has taken place. Motion representation, after all, is just one processing function: it needn’t signify the end of processing for a given stimulus feature involved in motion processing.

Processing models such as Reichardt detectors are thus perfectly at home with CCA—and with the evidence from visual asynchrony. The different latencies required by each apply to the production of *outputs* of different *types*, such as (i) a *neural representation* of feature F (e.g., motion) at some level of processing, and (ii) a *conscious experience* of feature G (e.g., location or colour). To represent x ’s motion, the relevant detectors must indeed integrate information from x ’s location L_1 at t_1 with information from x ’s location L_2 at t_2 , and so wait, by *neural* delay $\underline{\lambda}_M$, for the latter. As a result, the *neural* latency $\underline{\lambda}_M$ is likely longer than the *neural* latency $\underline{\lambda}_C$ which characterizes the processing of x ’s colour at t_1 . But if $\underline{\lambda}_C$ remains constant through processing of x ’s colour at t_1 , and then at t_2 , and t_3 , etc., so can λ_C in CCA, which sets the onset of conscious colour experience from $t_1 + \lambda_C$ to $t_3 + \lambda_C$, etc.—where interval λ_C in CCA, recall, covers both the neural latency $\underline{\lambda}_C$ and the online latency (if any) for conscious experience of colour.

49. NB: by “periphery”, Lee means, I take it, early stage of processing—e.g., in the retina.

In fact, CCA applies to λ_M just as well: if it takes $\underline{\lambda}_M$ to neurally process x 's motion from t_1 to t_2 , it also takes $\underline{\lambda}_M$ to process x 's motion from t_2 to t_3 , and then again from t_3 to t_4 , etc. Hence, λ_M , which determines the onset of the relevant motion experience in consciousness, can very well remain constant from $t_1 + \lambda_M$ to $t_2 + \lambda_M$, $t_3 + \lambda_M$, etc.—again, whether $\lambda_M > \underline{\lambda}_M$ or $\lambda_M = \underline{\lambda}_M$. In which case, successive presentations of a moving stimulus will give way to successive conscious experiences with different contents, by CCA—the kind of delayed processing at play in Reichardt detectors isn't enough to rescue the retentionalist thesis R₃.

4. Conclusion

Even if, *in principle*, unextended perceptual experiences *could* represent temporally extended events, there are reasons for thinking that, in fact, our conscious experiences *are* somehow “spread out” in time and structured accordingly. Such a conclusion derives from an eminently plausible assumption about the causal structure of our perceptual systems: it takes time for sensory information to be processed through a hierarchy of different neural areas. Such an interval is likely to vary somewhat, yet retain a lower and upper bound.

This explains why the retentionalist assumption of *partially delayed processing* (PDP), according to which successive stimuli could be processed at such different speeds as to enter consciousness simultaneously, is compromised by the evidence from visual asynchrony (§2.4). It's a problem for retentionalism, because the retentionalist thesis, R₃, about the temporal ontology of experience must presuppose PDP, we have seen.

On the other hand, a plausible alternative to *partially delayed processing*—CCA, which *is* compatible with visual asynchrony—can serve to mount a direct objection against R₃ and its different interpretations (§3.2). Given CCA, it's not true that experiences with temporally extended contents are (1) *instantaneous* (*temporally restricted* version), that (2) such experiences *must* represent the successive temporal parts of an extended event *simultaneously* (*modally restricted* version), or that (3) *slightly extended* but *very brief*—and temporally unstructured—experiences do so (*temporally relaxed* version). And since CCA is anchored in empirical considerations about our actual neurophysiological processes (including the evidence from visual asynchrony), it isn't in fact (4) *possible* in any relevant sense for the experiences of our neural duplicates to be unextended while carrying temporally extended contents (*modally relaxed* version, interpreted along the lines of M₄).

Accordingly, this paper aimed to suggest, the retentionalist thesis about the temporal ontology of experience (R₃) ought to be discarded on empirical grounds. Such a conclusion leaves the other retentionalist theses about temporally extended contents—R₁ and R₂—intact. Yet it implies that the mere avail-

ability of such extended contents isn't, in and of itself, enough to support a conclusion like R3—not without *partially delayed processing* (§3.3), that is. In this light, it's hardly surprising that considerations advanced to support the postulation of such extended contents—that is, the existence of Reichardt detectors, perfectly suited to deliver such contents—make little difference to the truth of the third retentionalist thesis (§3.4).

More importantly, if R3 and the assumption of *partially delayed processing* it rests upon are false, this leaves room for some isomorphy or similarity between the temporal properties of experiences and those of the events perceived, after all. What kind of temporal isomorphy? Does it mean that extensionalist views (for whom temporal experiences must be temporally extended) are thereby vindicated?

Not quite. Combined with cca, the toy example in §3.2 suggests at most that successive stimuli (or successive presentations of the same stimulus) and their different features reach consciousness successively. There is, in other words, some temporal *ordering* with which one consciously perceives the successive temporal parts of perceived events, at least. But then, synchronous features of the same stimulus may *not*, in fact, be consciously perceived simultaneously either—the evidence from visual asynchrony shows. Not much more, however, can be inferred from the evidence reviewed in this paper.

In which case, these results do not yet warrant the inference that our temporal experiences must be extended in time and temporally structured, as extensionalists argue. It could still turn out, for instance, that conscious perception largely consists in successions of very brief conscious snapshots, each representing only some feature of a given stimulus—the so-called “snapshot view”.⁵⁰ In fact, we have seen (§3.3), these results do not rule out a weakened version of retentionalism, according to which temporal experiences consist in successions of overlapping extended contents. Nevertheless, even if extensionalism isn't vindicated by the evidence discussed here, at least it retains the advantage (together with the snapshot view) of remaining entirely compatible with such evidence, unlike retentionalism.⁵¹

50. See, e.g., Locke (1689/1975), Prosser (2016), Chuard (2017), and Arstila (2018).

51. *Acknowledgments*: Thanks to the participants of the 2013 ANU Kioloa Consciousness Conference (especially David Chalmers, Colin Klein, David Pitt, and Jonathan Simons) for feedback on an ancient ancestor to this paper. Many thanks, as always, to my colleagues in the SMU work-in-progress gathering, and to Justin Fisher and Luke Robinson for their detailed comments in particular. Finally, all my gratitude to two referees for this journal for a host of genuinely helpful suggestions.

References

- Alais, David and Simon Carlile (2005). Synchronizing to Real Events: Subjective Audiovisual Alignment Scales with Perceived Auditory Depth and Speed of Sound. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6), 2244–47.
- Allik, Jüri and Kairi Kreegipuu (1998). Multiple Visual Latency. *Psychological Science*, 9(2), 135–38.
- Amano, K., A. Johnston, and S. Nishida (2007). Two Mechanisms Underlying the Effect of Angle of Motion Direction Change on Colour-Motion Asynchrony. *Vision Research*, 47, 687–705.
- Anscombe, G. E. M. and Peter Geach (1961). *Three Philosophers*. Oxford University Press.
- Arnold, Derek (2005). Perceptual Pairing of Colour and Motion. *Vision Research*, 45(24), 3015–26.
- Arnold, Derek (2010). Relative Timing and Perceptual Asynchrony. In R. Nijhawan and B. Khurana (Eds.), *Space & Time in Perception & Action* (254–77). Oxford University Press.
- Arnold, Derek and Colin Clifford (2002). Determinants of Asynchronous Processing in Vision. *Proceedings of The Royal Society: Biological Sciences*, 269(1491), 579–83.
- Arnold, Derek, Alan Johnston, and Shin'ya Nishida (2005). Timing Sight and Sound. *Vision Research*, 45, 1275–84.
- Arstila, Valtteri (2016). Theories of Apparent Motion. *Phenomenology and the Cognitive Sciences*, 15(3), 337–58.
- Arstila, Valtteri (2018). Temporal Experiences without the Specious Present. *Australasian Journal of Philosophy*, 96(2), 287–302.
- Arstila, Valtteri (2019). Time Markers and Temporal Illusions. In V. Arstila, A. Bardón, S. E. Power, and A. Vatakis (Eds.), *The Illusions of Time: Essays on Timing and Time Perception* (339–57). Palgrave Macmillan.
- Bachmann, Talis (1989). Microgenesis as Traced by the Transient Paired-Forms Paradigm. *Acta Psychologica*, 70, 3–17.
- Bachmann, Talis, Iris Luiga, Endel Pöder, and Kaupo Kalev (2003). Perceptual Acceleration of Objects in Stream: Evidence from Flash-Lag Displays. *Consciousness and Cognition*, 12, 279–97.
- Bartels, Andreas and Semir Zeki (1998). The Theory of Multistage Integration in the Visual Brain. *Proceedings of the Royal Society of London, B*, 265, 2327–32.
- Bartels, Andreas and Semir Zeki (2006). The Temporal Order of Binding Visual Attributes. *Vision Research*, 46, 2280–86.
- Bayne, Tim and Jakob Hohwy (2015). The Neural Correlates of Consciousness: Causes, Confounds, and Constituents. In S. M. Miller (Ed.), *The Constitution of Phenomenal Consciousness: Toward a Science and Theory* (155–76). John Benjamins.
- Bedell, Harold, Susana Chung, Haluk Ögmen, and Saumil Patel (2003). Color and Motion: Which Is the Tortoise and Which Is the Hare? *Vision Research*, 43, 2403–12.
- Bedell, Harold, Saumil Patel, Susana Chung, and Haluk Ögmen (2006). Perceptual Consequences of Timing Differences Within Parallel Feature-Processing Systems in Human Vision. In H. Ögmen and B. Breitmeyer (Eds.), *The First Half Second: The Microgenesis and Temporal Dynamics of Unconscious and Conscious Visual Processes* (245–58). MIT Press.

- Bernadete, Ethan and Ehud Kaplan (1999). The Dynamics of Primate M Retinal Ganglion Cells. *Visual Neuroscience*, 16, 355–68.
- Breitmeyer, Bruno and Haluk Ögmen (2006). *Visual Masking: Time Slices through Conscious and Unconscious Vision* (Oxford University Press).
- Broad, Charles (1923). *Scientific Thought*. Routledge.
- Broad, Charles (1925). *The Mind and Its Place in Nature*. Routledge.
- Chalmers, David (2000). What Is a Neural Correlate of Consciousness? In T. Metzinger (Ed.), *Neural Correlates of Consciousness: Empirical & Conceptual Issues* (17–40). MIT Press. (Reprinted in D. Chalmers, 2010, *The Character of Consciousness*, Oxford University Press)
- Chuard, Philippe (2017). The Snapshot Conception of Temporal Experience. In I. Phillips (Ed.), *The Routledge Handbook of Philosophy of Temporal Experience* (121–32). Routledge.
- Chuard, Philippe (2020). Temporal Consciousness. In U. Kriegel (Ed.), *The Oxford Handbook of the Philosophy of Consciousness* (188–207). Oxford University Press.
- Chuard, Philippe (2022). Experiential Holism in Time. *Mind & Language*, 37(4), 619–37.
- Clifford, Colin (2010). Dynamics of Visual Feature Binding. In R. Nijhawan and B. Khurana (Eds.), *Space & Time in Perception & Action* (199–215). Oxford University Press.
- Clifford, Colin, Derek Arnold, and Joel Pearson (2003). A Paradox of Temporal Perception Revealed by a Stimulus Oscillating in Colour and Orientation. *Vision Research*, 43, 2245–53.
- Dainton, Barry (2000). *The Stream of Consciousness*. Routledge.
- Dainton, Barry (2008). Sensing Change. *Philosophical Issues*, 18, 362–84.
- Dainton, Barry (2014a). Flow, Repetitions, and Symmetries: Replies to Lee and Pelczar. In N. Oaklander (ed.), *Debates in the Metaphysics of Time* (175–212). Continuum.
- Dainton, Barry (2014b). The Phenomenal Continuum. In D. Lloyd and V. Arstila (Eds.), *Subjective Time: The Philosophy, Psychology, and Neuroscience of Temporality* (101–37). MIT Press.
- Dainton, Barry (2017). Temporal Consciousness. *The Stanford Encyclopaedia of Philosophy*. Retrieved from <https://plato.stanford.edu/archives/fall2017/entries/consciousness-temporal/>
- Dehaene, Stanislas and Lionel Naccache (2001). Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework. *Cognition*, 79, 1–37.
- Dennett, Daniel and Marcel Kinsbourne (1992). Time and the Observer: The Where and When of Consciousness in the Brain. *Behavioral and Brain Sciences*, 15, 183–247. (Reprinted in N. Block, O. Flanagan, and G. Güzeldere [Eds.], 1997, *The Nature of Consciousness: Philosophical Debates*, MIT Press)
- Eagleman, David (2010). How Does the Timing of Neural Signals Map onto the Timing of Perception. In R. Nijhawan and B. Khurana (Eds.), *Space & Time in Perception & Action* (216–31). Oxford University Press.
- Eagleman, David and Terrence Sejnowski (2000). Motion Integration and Postdiction in Visual Awareness. *Science*, 287, 2036–38.
- Fabre-Thorpe, Michèle, Arnaud Delorme, Catherine Marlot, and Simon Thorpe (2001). A Limit to the Speed of Processing in Ultra-Rapid Visual Categorization of Novel Natural Scenes. *Journal of Cognitive Neuroscience*, 13(2), 171–80.
- Geach, Peter (1957). *Mental Acts*. Routledge.
- Geach, Peter (1969). *God and the Soul*. Routledge.

- Grush, Rick (2005). Internal Models and the Construction of Time: Generalizing from State Estimation to Trajectory Estimation to Address Temporal Features of Perception, Including Temporal Illusions. *Journal of Neural Engineering*, 2(3), 209–18.
- Grush, Rick (2006). How to, and How *Not to*, Bridge Computational Cognitive Neuroscience and Husserlian Phenomenology of Time. *Synthese*, 153(3), 417–50.
- Grush, Rick (2008). Temporal Representation and Dynamics. *New Ideas in Psychology*, 26, 146–57.
- Heron, James, David Whitaker, Paul McGraw, and Kirill Horoshenkov (2007). Adaptation Minimizes Distance-Related Audiovisual Delays. *Journal of Vision*, 7(13), 51–58.
- Herzog, Michael, Thomas Kammer, and Frank Scharnowski (2016). Time Slices: What Is the Duration of a Percept? *PLoS Biology*, 14(4), 1–12.
- Hoerl, Christoph (2009). Time and Tense in Perceptual Experience. *Philosophers' Imprint*, 9, 1–18.
- Hoerl, Christoph (2012). Husserl, the Absolute Flow, and Temporal Experience. *Philosophy and Phenomenological Research*, 86(2), 376–411.
- Hoerl, Christoph (2013). A Succession of Feelings, in and of Itself, Is Not a Feeling of Succession. *Mind*, 122(486), 373–417.
- Hoerl, Christoph (2017). On the View that We Cannot Perceive Movement and Change: Lessons from Locke and Reid. *Journal of Consciousness Studies*, 24(3–4), 88–102.
- Holcombe, Alex (2009). Seeing Slow and Seeing Fast: Two Limits on Perception. *Trends in Cognitive Science*, 13(5), 216–21.
- Holcombe, Alex (2015). The Temporal Organization of Perception. In Johan Wagemans (Ed.), *Oxford Handbook of Perceptual Organization* (820–39). Oxford University Press.
- Husserl, Edmund (1991). *On the Phenomenology of the Consciousness of Internal Time*. Kluwer.
- James, William (1952). *The Principles of Psychology* (Vol. 1). Dover. (Original work published 1890)
- Kant, Immanuel (1980). *Critique of Pure Reason* (N. Kemp Smith, Trans.). MacMillan. (Original work published 1790)
- Keetels, Mirjam and Jean Vroomen (2011). Perception of Synchrony between the Senses. In M. Wallace and M. Murray (Eds.), *The Neural Bases of Multisensory Processes* (147–77). CRC Press.
- Khurana, Beena and Romi Nijhawan (1995). Extrapolation or Attention Shift? *Nature*, 378, 566.
- Klotz, W. and P. Wolff (1995). The Effect of a Masked Stimulus on the Response to the Masking Stimulus. *Psychological Research/Psychologische Forschung*, 58, 92–101.
- Koch, Christoph (2004). *The Quest for Consciousness: A Neurobiological Approach*. Roberts & Company.
- Kopinska, Agnieszka and Laurence Harris (2004). Simultaneity Constancy. *Perception*, 33, 1049–60.
- Lappe, M. and B. Krekelberg (1998). The Position of Moving Objects. *Perception*, 27, 1437–49.
- Lee, B., J. Pokorny, V. Smith, P. Martin, and A. Valberg (1990). Luminance and Chromatic Modulation Sensitivity of Macaque Ganglion Cells and Human Observers. *Journal of the Optical Society of America, A*, 7, 2223–36.

- Lee, Geoffrey (2014a). Temporal Experience and the Temporal Structure of Experience. *Philosophers' Imprint*, 14(3), 1–21.
- Lee, Geoffrey (2014b). Extensionalism, Atomism, and Continuity. In N. Oaklander (Ed.), *Debates in the Metaphysics of Time*, 149–73. Continuum.
- Lee, Geoffrey (2014c). Experiences and Their Parts. In D. Bennett and C. Hill (Eds.), *Sensory Integration and the Unity of Consciousness* (287–321). MIT Press.
- LePoidevin, Robin (2007). *The Images of Time: An Essay on Temporal Representation*. Oxford University Press.
- Lewald, Jörg and Rainer Guski (2004). Auditory-Visual Temporal Integration as a Function of Distance: No Compensation for Sound-Transmission Time in Human Perception. *Neuroscience Letters*, 357(2), 119–22.
- Locke, John (1975). *An Essay Concerning Human Understanding*. P. Nidditch (Ed.). Oxford University Press. (Original work published 1689)
- Maunsell, John, Geoffrey Ghose, John Assad, Carrie McAdams, Christen Elizabeth Boudreau, and Brett Noerager (1999). Visual Response Latencies of Magnocellular and Parvocellular LGN Neurons in Macaque Monkeys. *Visual Neuroscience*, 16(1), 1–14.
- Maunsell, John and J. Gibson (1992). Visual Response Latencies in Striate Cortex of the Macaque Monkey. *Journal of Neurophysiology*, 68, 1332–44.
- Maus, Gerrit and Romi Nijhawan (2006). Forward Displacements of Fading Objects in Motion: The Role of Transient Signals in Perceiving Position. *Vision Research*, 46, 4375–81.
- Maus, Gerrit and Romi Nijhawan (2008). Motion Extrapolation into the Blind Spot. *Psychological Science*, 19, 1087–91.
- Maus, Gerrit, Beena Khurana, and Romi Nijhawan (2010). History and Theory of Flash-Lag: Past, Present, and Future. In R. Nijhawan and B. Khurana (Eds.), *Space & Time in Perception & Action* (477–99). Oxford University Press.
- Miller, Izchak (1984). *Husserl, Perception, and Temporal Awareness*. MIT Press.
- Moutoussis, Konstantinos (2012). Asynchrony in Visual Consciousness and the Possible Involvement of Attention. *Frontiers in Psychology*, 3, 1–9.
- Moutoussis, Konstantinos (2014). Perceptual Asynchrony in Vision. In D. Lloyd and V. Arstila (Eds.), *Subjective Time: The Philosophy, Psychology, and Neuroscience of Temporality* 201–16. MIT Press.
- Moutoussis, Konstantinos and Semir Zeki (1997a). A Direct Demonstration of Perceptual Asynchrony in Vision. *Proceedings of the Royal Society of London, B: Biological Sciences*, 264, 393–99.
- Moutoussis, Konstantinos and Semir Zeki (1997b). Functional Segregation and Temporal Hierarchy of the Visual Perceive Systems. *Proceedings of the Royal Society of London, B: Biological Sciences*, 264, 1407–14.
- Nijhawan, Romi (1992). Misalignment of Contours through the Interaction of Apparent and Real Motion Systems. *Investigative Ophthalmology & Visual Science*, 33(4), 974.
- Nijhawan, Romi (2002). Neural Delays, Visual Motion and the Flash-Lag Effect. *Trends in Cognitive Sciences*, 6(9), 387–93.
- Nishida, Shin'ya and Alan Johnston (2002). Marker Correspondence, Not Processing Latency, Determines Temporal Binding of Visual Attributes. *Current Biology*, 12, 359–68.

- Nishida, Shin'ya and Alan Johnston (2010). The Time Marker Account of Cross-Channel Temporal Judgments. In R. Nijhawan and B. Khurana (Eds.), *Space & Time in Perception & Action* (278–300). Oxford University Press.
- Nowak, Lionel and Jean Bullier (1997). The Timing of Information in the Visual System. In K. Rockland, J. Kaas, and A. Peters (Eds.), *Extrastriate Cortex in Primates* (205–41). Springer.
- Ögmen, Haluk, Saumil Patel, Harold Bedell, and Kaan Camuz (2004). Differential Latencies and the Dynamics of the Position-Computation Process for Moving Targets, Assessed with the Flash-Lag Effect. *Vision Research*, 44, 2109–28.
- Oram, M., D. Xiao, B. Drietschel, and K. Payne (2002). The Temporal Resolution of Neural Codes: Does Response Latency Have a Unique Role? *Philosophical Transactions of the Royal Society of London, B*, 357(1424), 987–1001.
- Palmer, Stephen (1999). *Vision Science: Photons to Phenomenology*. MIT Press.
- Patel, Saumil, Haluk Ögmen, Harold Bedell, and Vanitha Sampath (2000). Flash-Lag Effect: Differential Latency, Not Post-Diction. *Science*, 290, 1051a.
- Pelczar, Michael (2010). Must an Appearance of Succession Involve a Succession of Appearances? *Philosophy and Phenomenological Research*, 131(1), 49–63.
- Phillips, Ian (2010). Perceiving Temporal Properties. *The European Journal of Philosophy*, 18(2), 176–202.
- Phillips, Ian (2014a). Experience of and in Time. *Philosophy Compass*, 9(2), 131–44.
- Phillips, Ian (2014b). The Temporal Structure of Experience. In D. Lloyd and V. Arstila (Eds.), *Subjective Time: The Philosophy, Psychology, and Neuroscience of Temporality* (139–58). MIT Press.
- Phillips, Ian (2018). Consciousness, Time, and Memory. In R. Gennaro (Ed.), *The Routledge Handbook of Consciousness* (286–97). Routledge.
- Prosser, Simon (2016). *Experiencing Time*. Oxford University Press.
- Purushothaman, Gopathy, Saumil Patel, Harold Bedell, and Haluk Ögmen (1998). Moving Ahead through Differential Latency. *Nature*, 396, 424.
- Rao, Rajesh, David Eagleman, and Terrence Sejnowski (2001). Optimal Smoothing in Visual Motion Perception. *Neural Computation*, 13, 1243–53.
- Rashbrook-Cooper, Oliver (2017). Atomism, Extensional and Temporal Presence. In I. Phillips (Ed.), *The Routledge Handbook of Philosophy of Temporal Experience* (133–45). Routledge.
- Reid, Thomas (1855). *Essays on the Intellectual Powers of Man*. Derby.
- Roufs, Jacques (1963). Perception Lag as a Function of Stimulus Luminance. *Vision Research*, 3(1–2), 81–91.
- Soteriou, Matthew (2007). Content and the Stream of Consciousness. *Philosophical Perspectives*, 21, 543–68.
- Soteriou, Matthew (2010). Perceiving Events. *Philosophical Explorations*, 13(3), 223–41.
- Soteriou, Matthew (2013). *The Mind's Construction*. Oxford University Press.
- Stone, J. V., N. M. Hunkin, J. Porrill, R. Wood, V. Keeler, M. Beanland, M. Port, and N. R. Porter (2001). When Is Now? Perception of Simultaneity. *Proceedings of the Royal Society B*, 268, 31–38.
- Strawson, Galen (2009). *Selves*. Oxford University Press.
- Stromeyer, C. and P. Martini (2003). Human Temporal Impulse Response Speeds Up with Increased Stimulus Contrast. *Vision Research*, 43, 285–98.
- Sugita, Yoichi and Yôiti Suzuki (2003). Audiovisual Perception: Implicit Estimation of Sound-Arrival Time. *Nature*, 421, 911.

- Thorpe, Simon, Denis Fize, and Catherine Marlot (1996). Speed of Processing in the Human Visual System. *Nature*, 381, 520–22
- Tye, Michael (2003). *Consciousness and Persons: Unity and Identity*. MIT Press.
- Van Eijk, Rob, Armin Kohlrausch, James Juola, and Steven van de Par (2008). Audiovisual Synchrony and Temporal Order Judgments: Effects of Experimental Method and Stimulus Type. *Perception and Psychophysics*, 70(6), 955–68.
- VanRullen, Rufin and Simon Thorpe (2001). The Time Course of Visual Processing: From Early Perception to Decision-Making. *Journal of Cognitive Neuroscience*, 13(4), 454–61.
- Viviani, Paolo and Christelle Aymoz (2001). Colour, Form, and Movement Are Not Perceived Simultaneously. *Vision Research*, 41, 2909–18.
- von Helmholtz, Hermann (1962). *Treatise on Physiological Optics* (Vol. 3). J. P. C. Southall (Ed.). Dover. (Original work published 1910)
- Wilson, J. and S. Anstis (1969). Visual Delay as a Function of Luminance. *The American Journal of Psychology*, 82(3), 350–58.
- White, Alex, Daniel Linares, and Alex Holcombe (2008). Visuomotor Timing Compensates for Changes in Perceptual Latency. *Current Biology*, 18(20), 951–53.
- Wu, Wayne (2014). *Attention*. Routledge.
- Wu, Wayne (2017). Attention and Perception: A Necessary Connection? In B. Nanay (Ed.), *Current Controversies in Philosophy of Perception* (148–62). Routledge.
- Zeki, Semir (1993). *A Vision of the Brain*. Blackwell.
- Zeki, Semir (2003). The Disunity of Consciousness. *Trends in Cognitive Science*, 7(5), 214–18.
- Zeki, Semir (2015). A Massively Asynchronous, Parallel Brain. *Philosophical Transactions of the Royal Society, B*, 370, 1–14.