15 Temporal Mental Imagery

Gerardo Viera and Bence Nanay

Mental Imagery

The definition of mental imagery as perceptual processing that is not triggered by corresponding sensory stimulation in the relevant sense modality is a fair summary of the way the concept of mental imagery is used in psychology and neuroscience (Nanay, 2018a; Nanay, forthcoming; Pearson et al., 2015). In a recent review article on mental imagery, for example, the authors say: "We use the term 'mental imagery' to refer to representations . . . of sensory information without a direct external stimulus" (Pearson et al., 2015: 590). And in a much older paper written by Kosslyn, Behrmann, and Jeannerod, the authors characterize the concept of visual mental imagery as "'seeing' in the absence of the appropriate immediate sensory input" (Kosslyn, Behrmann, and Jeannerod, 1995a: 1335).

But it is also easy to see that many everyday examples of mental imagery will also fit this definition. When you close your eyes and visualize an apple, there is no sensory input at all – your eyes are closed. But there is early cortical perceptual processing (sometimes as early as in the primary visual cortex – see Kosslyn et al., 1995b; Page, Duhamel, and Crognale, 2011; Slotnick, Thompson, and Kosslyn, 2005). You have early cortical perceptual processing that is not triggered by corresponding sensory stimulation in the relevant sense modality.

A lot more needs to be said about how to cash out this definition of mental imagery in general and temporal mental imagery in particular. Take vision as an example (see Bullier, 2004; Grill-Spector and Mallach, 2004). The light hits your retina. Then this sensory stimulation is processed via the primary visual pathway that connects neural networks in the retina to the primary visual cortex (V1) via the lateral geniculate nucleus (LGN) in the thalamus. Outputs from V1 activate other parts of the visual cortex and are also fed forward to a range of extrastriate areas (V2, V3, V4/V8, V3a, V5/MT).

Given that the early visual cortices are retinotopic, we can identify direction-sensitive neurons in, say, the primary visual cortex, that are sensitive to the retinal activation of a certain part of the retina. If there is retinal activation of a certain shape, then these direction-sensitive neurons will fire reliably. If, for example, the sensory stimulation is the visual input of a straight horizontal line in the middle of the visual field, then those direction-sensitive neurons in the middle of the field of the retinotopic primary visual cortex that are sensitive to horizontal input will fire.

This counts as corresponding sensory input. But a part of the primary visual cortex can be active even though there is no corresponding sensory input. When this happens, we talk of mental imagery.

In the visual case, it is easy enough to check whether there is (spatial) correspondence between the early cortices and the retina, given the retinotopy of the early visual cortices. Simply put, if there is a triangle in the middle of the retina and there is also an isomorphic triangle in the primary visual cortex, we have a match: The perceptual processing in V1 is triggered by corresponding sensory stimulation in the relevant sense modality.

But if there is no triangle in the middle of the retina but there is still a triangle in the V1, then we have no match: The perceptual processing in the primary visual cortex is not triggered by corresponding sensory stimulation in the relevant sense modality. In other words, we have an instance of mental imagery.

Such correspondences are relatively easy and straightforward enough to check in the case of V1. But it is neither easy nor straightforward if we go a bit further up in the visual processing hierarchy or if we focus on the non-visual sense modalities. Or if we focus on temporal and not spatial correspondence.

So instead of only relying on retinotopy, which is a convenient but not usually feasible way of assessing correspondence between sensory input and perceptual processing, we need a more general method of telling cases of correspondence from cases of non-correspondence (that is, mental imagery).

Suppose that a certain sensory stimulation-type, S1, reliably causes the perceptual processing of a type, P1, in a specific agent, A1. Each time A1 gets S1 as sensory input, A1's perceptual system engages in P1 processing. However, if we get P1 in this agent, but P1 is not triggered by S1 but rather by S2, which does not reliably cause P1 in A1, then we have a case of mental imagery. Indexing to an agent is required here, since for any given stimulation-type, different people may exhibit different subsequent sensory processing as a result of a range of individual differences.

Some clarifications are needed. S1 is a sensory stimulation-type, so it is a type of event that happens to our sense organs. P1 is perceptual processing, by which we mean processing in early cortical areas, from V1 to MT. There is an ongoing debate about how to delineate perception from cognition (or from post-perceptual processing, see Beck, 2018; Nanay, 2012; Phillips, 2019), but we do not need to take sides in this debate here (on tricky issues such as whether face perception is genuine perception). Even those who would want to restrict perceptual processing to the bare minimum would count the areas V1 to MT as perceptual areas.

Mental imagery may or may not be voluntary or conscious. Remember that it is defined in terms of early cortical activation with a certain kind of etiology (namely early cortical activation that is not triggered by corresponding sensory stimulation). So no necessary reference to consciousness or volition is built into this definition.

Amodal completion will count as mental imagery in the sense just described (Nanay, 2010, 2018b). When you see a cat behind a picket fence, you amodally complete the occluded parts of this perceived cat. There is early cortical perceptual processing of the occluded outlines of the hidden parts of the cat, but this processing does not correspond to any outline on the retina (see Bakin, Nakayama, and Gilbert, 2000; Ban, Yamamoto, and Hanakawa, 2013; Bushnell et al., 2011; Emmanuoil and Ro, 2014; Hazenberg et al., 2014; Hedgé et al., 2008; Komatsu, 2006; Kovacs et al., 1995; Lee and Nguyen, 2001; Lee et al., 2012; Pan et al., 2012; Scherzer and Ekroll, 2015; Shibata et al., 2011; Smith and Muckli, 2010; Sugita, 1999). The part of the retina that would correspond to the amodally completed outline is the homogenous monochrome white of the picket.

The amodal completion example is important for a number of reasons. First, it highlights that the concept of correspondence should be understood as local correspondence. In some (most) cases of amodal completion, the completion can be fully explained in a bottom-up manner: The entire retinal stimulation does reliably cause the amodally completed outlines in the entire V1. But the local input of the missing (because occluded) outline does not reliably cause the activation of this outline in the V1. So while there is global correspondence between the entire state of retinal activation and the subsequent spatiotopic sensory processes, there is no local correspondence and hence the amodal completion of this occluded outline (in V1) is not triggered by (locally) corresponding sensory stimulation in the relevant sense modality. Amodal completion amounts to mental imagery.

Amodal completion is also important because it highlights that mental imagery can be bottom-up but it may also be top-down influenced. The amodal completion of the hidden parts of the cat often depends on our prior knowledge of the anatomy of cats. Mental imagery can be bottom-up or top-down as can amodal completion. And the same is true of temporal mental imagery.

The Case for Temporal Mental Imagery

A helpful aspect of this way of thinking about mental imagery is that this could be applied to the temporal case very easily. Visual sensory stimulation reliably leads to V1 activation in 30 milliseconds (see Rauschenberger et al., 2006; Rolls and Tovee, 1994; Thorpe et al., 1996, for summaries). If we have V1 activation but no visual sensory stimulation that would have preceded this V1 activation by 30 milliseconds, then there is no temporal correspondence. The perceptual processing (in V1) is not triggered by temporally corresponding sensory stimulation. We have an instance of temporal mental imagery. As with mental imagery in general, temporal correspondence is also relative to the individual – some of us are slower than others. Temporal mental imagery is perceptual processing that is triggered by spatially corresponding sensory stimulation in the appropriate sense modality, but where this perceptual processing does not temporally correspond with the incoming stimulation.

There is a certain ambiguity in the definition of temporal mental imagery as early perceptual processing that is not triggered by temporally corresponding sensory stimulation. In cases in which there is no local correspondence of any sort between perceptual processes and sensory stimulation, there will by default also be no temporal correspondence (this would include cases in which you close your eyes and imagine an apple). Mental imagery of this sort would automatically count as temporal mental imagery. However, for our purposes in this chapter, we mean something narrower by temporal mental imagery. The cases on which we will focus are ones in which there is local sensory stimulation that corresponds, along some dimension with the perceptual processes, yet fails to temporally correspond with early perceptual processes. For instance, in the visual case, we have in mind cases such as predictive and postdictive perception in which there are perceptual processes that retinotopically correspond with the sensory stimulation but fail to correspond with the timing of the sensory stimulation.

Temporal correspondence can fail in two directions: The perceptual processing may come earlier than it should – this is a case of “predictive temporal mental imagery.” Alternatively, it may come later than it should – this would amount to “postdictive temporal mental imagery.”

One important advantage of this way of thinking about temporal mental imagery is that it can help us to explain a recurring theme in thinking about the experience of time. This was summarized memorably by William James, who writes that:

> . . . the practically cognized present is no knife-edge, but a saddle-back, with a certain breadth of its own. (James, 1890: 609)

In other words, our experiences have a certain temporal thickness. But what does this mean exactly? Here is a more contemporary philosophical spin on what James had in mind:

> The dynamic content of our experience at short timescales is metaphysically dependent on the content of experience over longer timescales (Phillips, 2011: 3).

So when we have an experience of, say, watching a football fly through the air and bounce off the goalpost, our experience should not be characterized as the sequence of dimensionless point-like experiences. Rather, my experience of the ball right now somehow represents the ball a split second ago and also represents where the ball would be in a split second. This phenomenon is often described, following James, as the “specious present.”

Here is the saddle-back (Figure 15.1). The middle of the saddle would be the present. But we somehow represent the two flanks of this bell-shape, as well. The question is: how?

This raises some deep issues about the nature of perception. How is it possible to perceive something that is not present? The ball a split second ago is no longer

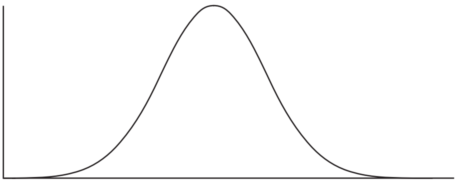

Figure 15.1 *A pictorial representation of the specious present. The peak of the saddle back would be the present moment, and our overall experience extends to a lesser degree to the past and future.*

present, and the ball in a split second is not present yet. According to an influential line of thought in philosophy, we can only perceive what is there to be perceived (e.g. Grice, 1961). But when it comes to time, only the present is present (sic!). So the past, let alone the future can't be perceptually represented.

There are various sophisticated ways of dealing with this problem – for example, extending the temporal dimension of not just the content, but also the vehicle of perceptual representations (see Phillips, 2010). But if we take the concept of temporal mental imagery seriously, then there is no need to complicate things unnecessarily.

We represent the flanks of the bell-shape by means of temporal mental imagery. We know that the early cortical processing of a temporal event has a much wider temporal profile then the retinal event. So some of this perceptual processing will be triggered by corresponding sensory stimulation (the middle), but most of it will not be. It will count as temporal mental imagery in which the early cortical processing is not triggered by temporally corresponding sensory stimulation. It is triggered by sensory stimulation that is either too early or too late.

The further away we veer from the sensory stimulation-driven perceptual processing (the middle of the saddle-shape), the bigger the role temporal mental imagery will play.

Amodal Completion and the Specious Present

To understand the role of mental imagery in an account of the specious present, it is useful to revisit the amodal completion case. When we see the cat hiding behind the picket fence, our sensory systems are only in causal contact with the unoccluded parts of the cat and the slats of the fence. Yet we do not take ourselves to only perceive undetached cat slices, but instead we take ourselves to see a complete cat partly hidden by the fence. This alone, however, does not establish that mental imagery is at play. Our sense of seeing an entire but partly occluded cat could be the result of a post-sensory belief about the world. It was an empirical discovery, not introspection, that showed that our early perceptual processes do not merely represent a cat-picket-cat pattern that corresponds with sensory stimulation, but instead these early perceptual processes (as early as V1 and V2) represent the entire partly occluded shape of the cat. In this way, amodal completion (in at least some central cases) employs perceptual processes that represent cat contours but that are retinotopically associated with regions of the retina that are receiving picket-stimulation. As a result, we have empirical evidence for the role of mental imagery in amodal completion.

Now, turn to the case of the specious present. When we see a football ricochet off of a goalpost, our current sensory stimulation is driven by the current state of the world (taking into account the time it takes the light to travel from the ball to the retina). Yet, just as in the amodal completion case, we do not have the sense that we are only perceiving the static snapshot of the world that is causally impinging on our retinas, but instead, we perceive a complete dynamic temporal interval. Following

James's characterization of the specious present, we seem to not only perceive the current location of the ball, but we are also perceptually aware of where the ball was a split second ago, of the ball's impact with the goalpost, and where the ball will likely be in a split second from now. In this way, the specious present is a temporal analog of the amodal completion case. While in the amodal completion case perception fills in spatial details of the perceived scene, the specious present fills in temporal details of the perceived scenario (i.e. the temporal structure of how events unfold over a period of time). What is left to establish is whether this sense of being aware of a temporally extended interval, the specious present, involves mental imagery or not.

On one philosophical account of temporal perception, *extensionalism*,[1] accounts of the specious present need not employ any appeal to mental imagery. According to extensionalism, the experience of a sequence of events in the world, say B as following A, requires that the overall experience can be decomposed into an experience of A and an experience of B, and that the experience of A, understood as a mental/neural event, precedes the experience of B. In the case of seeing the ball ricochet off the goalpost, our overall experience must decompose into distinct experiences of the ball at various locations on its trajectory, including an experience of the impact, and that these experiences must stand in the appropriate temporal relations to one another. In cases of straightforward veridical perception, the sequence of sensory stimulations will lead to a sequence of perceptual processes that underpin these distinct experiences. If, as the extensionalist would suggest, this suffices for an account of the specious present, then no temporal mental imagery would be needed. Every step of the processing sequence would appropriately correspond to the temporal structure of the incoming sequence of sensory stimulation.

However, a mere sequence of experiences cannot by itself account for the perception of temporal order. In order to perceive the ball *as ricocheting* off the goalpost, it is not enough that we recently perceived the ball approaching the goalpost, then impacting the goalpost, and that now we are perceiving the ball as having some distance from the goalpost. Instead, for us to currently be perceiving the ball as ricocheting off of the goalpost, we must somehow retain information about earlier perceptual states that represented the ball's approach and impact with the goalpost. As Geoffrey Lee (2014b) put it, our prior sensory responses to the world must leave behind traces in the current state of the perceptual system that can be usefully integrated with the current incoming sensory signals. To use another example, when we perceive a crash of thunder as following a flash of lightning, we can only perceive this temporal relation once the thunder influences our sensory receptors. But by that time we will have already processed the signal from the flash of lightning. If the prior perceptual processes representing the lightning left no trace in the current state of the brain, then we would be unable to perceive the thunder as following the lightning. The thunder would simply appear to occur in temporal isolation.

---

[1] Terminology in the literature has not been settled. In calling the view extensionalism we are following usage in (Lee, 2014a); however, the view has also been called molecularism (Hoerl, 2009), the process-view (Lee, 2014b), and the naïve-theory of temporal perception (Phillips, 2010; 2014).

There must be some means, then, by which prior perceptual states leave a trace in the current workings of the perceptual/cognitive system. The question, then, is whether temporal mental imagery has a role to play in how these traces are retained. That is, is this information retained in perceptual processes that do not temporally correspond with the current sensory stimulation? This is not something that any armchair philosophical analysis can determine. Instead, just as in the case of amodal completion, we must turn to the sciences to see whether the relevant representational mechanisms amount to mental imagery.

Since the research on temporal perception is still very much in its infancy our claims here will be circumscribed. We cannot make general claims to the end that in all cases the specious present requires temporal mental imagery. In fact, given the variety of means by which perceptual systems keep track of the temporal structure of our world, it may very well turn out that the speciousness of the present will have several distinct accounts that apply in different contexts. However, what will follow are cases in which the empirical evidence clearly points to the role of temporal mental imagery in accounting for the specious present.[2]

Prediction and Mental Imagery

As James noticed, the breadth of the specious present extends into the past as well as the future. It involves predictive elements that provide us with expectations about what is to come. When we listen to a familiar song, for instance, we notice immediately if the musician misses a note. Or when we see two cyclists approaching an intersection from different directions, we grimace in expectation of the coming collision. Why do we react in these ways? A natural explanation is that we form predictions or expectations of what is going to occur. Any representation of these expectations would seem to precede the relevant sensory stimulation that would lead to perceptual processing of these expected scenarios. The question is, once again, whether these expectations are based in temporal mental imagery – that is, perceptual processes that do not temporally correspond to sensory stimulation (see Zatorre and Halpern, 2005), or whether these expectations are due to non-perceptual capacities that represent the expected scenarios.

While some violations of expectation might be the result of non-perceptual processes, in a recent study by Ekman, Kok, and de Lange (2017), they show that in at least some cases expectations about object trajectories employ temporal mental imagery. In their study they familiarized subjects with a particular dot sequence (a dot moving from the top-left of a monitor to the top-right) and using high speed fMRI they were able to

---

[2] The examples discussed in this chapter are primarily visual ones. The general point that Geoffrey Lee describes, however, applies to all modalities. To experience succession in any non-visual modality, for instance hearing a tone increase in pitch over a temporal interval, requires the retention of information over time. More specifically, however, postdictive phenomena similar to those described below in vision can be found in touch (e.g. the cutaneous rabbit illusion [Geldard and Sherrick, 1972; Grush 2005, 2007]) and in audition (e.g. the auditory flash-lag effect (Alais and Burr, 2003)). However, in both cases the required imaging studies have not been conducted to see whether or not there is the appropriate temporal non-correspondence between the timing of sensory stimulation in these other modalities and the relevant perceptual processes.

successfully measure BOLD responses in V1 for retinotopic locations corresponding with the incoming sensory stimulation. In this way, they were able to map out the trajectory of the dot sequence in V1. They then ran subjects in two distinct conditions. In the control sequence they presented subjects with an initial display in which the dot was located at the end-location of the familiarization sequence (i.e. the dot was shown in the top-right). The corresponding BOLD response only showed activity in V1 areas with receptive fields for the location of the presented stimulus. In the *preplay* condition, subjects were presented with an initial display in which the dot was located at the start location of the familiarization sequence (i.e. the dot was shown in top-left). Interestingly, in this condition subjects showed BOLD responses that corresponded to a time-compressed trajectory of the entire familiarization sequence. That is, they showed BOLD responses for a dot moving from the top-left to the top-right of the display. The response was time-compressed in that the cortical processes traced out the expected trajectory of the dot quicker that they would if they were responding to the actual dot sequence. Furthermore, it was shown that when this cortical preplay was elicited by the initial dot display, subsequent detection performance for the location of the dot along that trajectory was enhanced.

In this case, we have predictive (or anticipatory) temporal mental imagery in that we have early cortical perceptual processes that do not temporally correspond with the relevant sensory stimulation. The V1 activity occurs *prior to* the relevant sensory stimulation.[3]

Postdiction, Apparent Motion, and Mental Imagery

Postdictive perception picks out a range of perceptual phenomena in which the perception of an earlier stimulus is modulated by the perception of a later stimulus. A classic example of postdiction is apparent motion. A standard setup for eliciting apparent motion is the following: A flash of light is presented at T_1 at location L_1, then a second light is presented at T_3 at location L_3. When presented with this sequence, subjects do not perceive the display as consisting of two spatiotemporally separated flashes of light, but instead, they perceive the display as consisting of a single light that smoothly travels from L_1 to L_3 while passing through the intermediary location of L_2 at T_2.

Researchers have found that in early perceptual cortices (as early as V1) (Larsen et al., 2006; Muckli et al., 2005) there is retinotopic activity that corresponds with the apparent trajectory and not just with the presentation of the two spatially separate flashes of light. In this way, the perceptual processes that are associated with activity of retinal regions encoding for the location L_2 are active despite not corresponding to the retinal activity at that retinal location. However, this finding only shows that there is a spatial form of mental imagery that underpins the perception of apparent motion.

---

[3] The authors of this study deny that mental imagery plays a role in preplay. However, they are assuming that mental imagery must be driven through the deliberate top-down process of recreating a visual percept. Since we needn't build the idea of deliberate control into our notion of mental imagery, we needn't accept their conclusion that this is not a form of mental imagery.

Furthermore, since apparent motion may be entirely driven by bottom-up feedforward mechanisms (see Shimoji [2014] for a discussion), it is possible that the sensory stimulation does not require any temporal mental imagery in which there is temporal non-correspondence between perceptual processes and sensory stimulation.

However, in follow-up studies, it was found that area hMT+/V5 mediates the activation of V1 processes corresponding to the apparent motion trajectory (Larsen et al., 2006; Muckli et al., 2005; Sterzer, Haynes, and Rees, 2006). Given that there is a feedback mechanism at work in apparent motion, it shows that there must be a temporal non-correspondence between perceptual activity and the timing of sensory stimulation. In order for this feedback to influence the perception of the light as being at location L_2 at T_2, the perceptual system must first receive the modulatory sensory stimulation of the light being at location L_3 at T_3. Since the activation of V1 areas representing location L_2 is mediated by hMT+/V5 activity, then this activation must occur after the initial activation of V1 to the stimulus at location L_3. Since the activation of L_3 initially corresponded with the sensory stimulation, then it must be the case that any subsequent activation of V1 in response to the apparent motion display must no longer correspond to the temporal pattern of sensory stimulation. That is, in addition to spatial mental imagery, postdictive apparent motion employs temporal mental imagery in order to integrate the traces of past perceptual processes with the current incoming sensory stimulation. Our experience of the light as moving from L_1 to L_3, including the speciousness of the light having just been at locations L_1 and L_2, relies on temporal mental imagery.

Multimodal Temporal Mental Imagery

Apparent motion and perceptual preplay phenomena show that within individual sensory modalities temporal mental imagery plays a role in how the perceptual system fills in the details about the temporally extended world in a way that goes beyond a mere reflection of sensory stimulation. Temporal mental imagery also plays a role in the production of a coherent multisensory world.

It has been widely established that activity in sensory cortices can be entrained to rhythmic patterns in sensory stimulation (Rees, Green, and Kay, 1986; Regan, 1966). A rhythmic flashing of light will cause activity in the visual cortices to oscillate in phase with the flashing light. Similar entrainment can be found in the other sensory systems. In the unimodal cases, perceptual processes will naturally temporally correspond with the timing of sensory stimulation.

It has also been widely established, as a perceptual phenomenon, that when a rhythmic visual stimulus and a rhythmic auditory stimulus are shown to an individual, and the rhythms share a common frequency but are slightly out of phase with one another, the perceptual system will quickly adapt to the discrepancy in the stimuli and will come to perceive the auditory and visual stimuli as being in phase (for classic papers see Fujisaki et al., 2004; Vroomen et al., 2004; and see Vroomens and Keetels, 2010, for a review). There is a perceptual shift in the timing of the events in the world. Just as in the cases of apparent motion and amodal completion, the appearance of the auditory and visual stimuli as being in phase with one another could be accounted for without appealing to any temporal mental imagery. The perceptual recalibration could simply reflect the operation of a post- (or late) perceptual mechanism.

In recent series of experiments, Kösem, Gramfort, and van Wassenhove (2014) showed that this sort of temporal recalibration involved a shift in the timing of initial sensory processes. By using MEG imaging techniques, they were able to isolate the auditory and visual perceptual responses and show that individually the modality specific processes were entrained by their relevant stimuli. Then, after perceptual recalibration occurred, they found that there was a corresponding shift in the timing of the modality specific sensory processes. The timing of the perceptual processes no longer corresponded with the temporal pattern of sensory stimulation, but instead was shifted as the perceptual system tried to produce a coherent representation of the world in which a single audiovisual rhythm was impinging on the sensory receptors. Once again, we find a role for temporal mental imagery in our perception of the temporal structure of the world around.

Limits of Temporal Mental Imagery

At this point, a word of caution is needed. While temporal mental imagery seems to play a role in how we perceive the temporally structured world around us, it leaves certain important aspects of the perception of time unanswered. In order to see what gets left out, we can make a three-layer distinction between different aspects of our perceptual engagement with the world around us.

Let's begin with the case of visual space. First, (1) there is the spatial structure of our sensory stimulation. That is, the spatial distribution of activity on our retinas. Second, (2) there is the spatial structure of early cortical processes – i.e. the retinotopic structure of these early visual cortices. Third, (3) there is the spatial content of early vision. As it has been defined, cases of spatial mental imagery have been characterized as cases in which there is a failure of correspondence between (1) and (2) – that is, a failure of correspondence between the retinopically structured perceptual activity in early visual cortices and the spatial distribution of retinal activity. Given how spatial content of early vision is encoded via the retinotopic structure of the early cortical maps, a non-correspondence between (1) and (2), i.e. the existence of mental imagery, implies a non-correspondence between (1) and (3). Mental imagery in this case provides us with a straightforward account of the spatial content of visual mental imagery.

However, when we translate this three-layer distinction to the temporal case, the same transition does not hold. First, (1) there is the temporal structure of our sensory stimulation – i.e. the temporal sequence of activity on the sensory receptors. Second, (2) there is the temporal structure of early cortical processes. Third, (3) there is the temporal content of perception – e.g. the durations and temporal relations attributed to perceived events. While the existence of retinotopic structure, and its role in spatial representation, allowed for a non-correspondence between (1) and (2) to imply a non-correspondence between (1) and (3), the same cannot be said in the case of time. The timing of perceptual processes can be divorced from their temporal contents (Viera, 2019).

To see how this might be so, consider cases of visuomotor temporal recalibration. In a study by Stetson et al. (2006), subjects were asked to press a button and then after a 35-millisecond delay a flash of light would appear on the screen in front of them. After subjects did this for a while, the experimenters then inserted an extended delay of 135 milliseconds between the button press and the flash of light. It was determined that subjects would quickly adapt to the inserted delay and as a result the perceived interval between the button press and the flash of light would be misperceived as being shorter than it in fact was. However, the interesting part comes when experiments removed the extended delay and returned to the 35-millisecond delay condition. Subjects would report seeing the flash of light as coming before the button press even though the stimuli were identical to those used in the initial portion of the study. Somehow the perceptual system had recalibrated its perception of temporal order.

One possibility for what might have occurred here was that through recalibration there was a shift in the timing of the early sensory processes. Either there was a replaying of the button press after the flash of light (a type of filling in) or both the processing of the flash of light and button press were delayed and played out in an order that fit the perceived order of events. However, neither of these options were borne out by the imaging data. Instead, the initial perceptual processes had the same timecourse prior to and post recalibration.[4] In both cases, the timing of the perceptual processes corresponded with the timing of sensory stimulation. As a result, no temporal mental imagery was at play.[5]

What is posited as making the difference in the perceived order is a late perceptual mechanism that takes the events represented by earlier cortical processes and puts them into temporal relations with one another. The take home message, however, for this study is that unlike the representation of visual space in early visual processes, there is no straightforward way of moving from claims about the temporal structure of perceptual processes to their temporal contents. Temporal mental imagery might provide us with a form of representational access to events that are not currently stimulating our sensory receptors, but an explanation of how those events are perceived as occurring in time may need to appeal to further resources.

Finally, it is important to stress that we have been focusing on temporal mental imagery in the visual sense modality. Given that we know a lot about the neural correlates of mental imagery in other sense modalities, one important new research direction would be to examine the similarities and differences between temporal mental imagery in the different sense modalities. Auditory temporal mental imagery is an especially important subject of research given the importance of temporality in audition. And given the deep multimodality of our perceptual system, another important question is how the temporal mental imagery in all these different sense modalities interact.

---

[4] For evidence see Cai et al. (2012) and Stekelenburg, Sugano, and Vroomen (2011).

[5] The possible difference between the mechanism for recalibration in this case and the mechanism for recalibration for the rhythmic audiovisual stimuli might have to do with the fact that in this case the stimuli were non-rhythmic.

References

Alais, D., and Burr, D. (2003). The "Flash-Lag" Effect Occurs in Audition and Cross-Modally, *Current Biology*, *13*(1), 59–63.

Bakin, J., Nakayama, K., and Gilbert, C. (2000). Visual Responses in Monkey Areas V1 and V2 to Three-Dimensional Surface Configurations. *Journal of Neuroscience*, *20*, 8188–8198.

Ban, H., Yamamoto, H., Hanakawa, T., et al. (2013). Topographic Representation of an Occluded Object and the Effects of Spatiotemporal Context in Human Early Visual Areas. *Journal of Neuroscience*, *33*, 16992–17007.

Beck, J. (2018). Marking the Perception-Cognition Boundary. *Australasian Journal of Philosophy*, *96*, 319–334.

Bullier, J. (2004). Communications between Cortical Areas of the Visual System. In L. M. Chalupa and J. S. Werner (eds.), *The Visual Neurosciences*. Cambridge, MA: MIT Press, 522–540.

Bushnell, B. N., Harding, P. J., Kosai, Y., and Pasupathy, A. (2011). Partial Occlusion Modulates Contour-Based Shape Encoding in Primate Area V4. *Journal of Neuroscience*, *31*, 4012–4024.

Cai, M., Stetson, C., Eagleman, D. M. (2012). A Neural Model for Temporal Order Judgments and their Active Recalibration: A Common Mechanism for Space and Time? *Frontiers in Psychology*, *3*, 470.

Ekman, M., Kok, P., and de Lange, F. (2017). Time-Compressed Preplay of Anticipated Events in Human Primary Visual Cortex. *Nature Communications*, *8*.

Emmanouil, T., and Ro, T. (2014). Amodal Completion of Unconsciously Presented Objects. *Psychonomic Bulletin & Review*, *21*(5), 1188–1194.

Esterman, M., and Yantis, S. (2010). Perceptual Expectation Modulates Category-Selective Cortical Activity. *Cerebral Cortex*, *20*, 1245–1253.

Fujisaki,W., Shimojo, S., Kashino,M., and Nishida, S. (2004). Recalibration of Audiovisual Simultaneity. *Nature Neuroscience*, *7*, 773–778.

Geldard, F. A., and Sherrick, C. E. (1972). The Cutaneous "Rabbit": A Perceptual Illusion. *Science*, *178*, 178–179.

Grice, H. P. (1961). The Causal Theory of Perception. *Proceedings of the Aristotelian Society, Supplementary Volume*, *35*, 121–153.

Grill-Spector, K., and Malach, R. (2004). The Human Visual Cortex. *Annual Review of Neuroscience*, *27*, 649–677.

Grush, R. (2007). Time and Experience. In T. Muller (ed.), *Philosophie der Zeit: Neue Analytische Ansatze*. Frankfurt, Germany: Verlag Vittorio Klostermann.

(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), S209–218.

Hazenberg, S. J., Jongsma, M. L. A., Koning, A., and van Lier, R. (2014). Differential Familiarity Effects in Amodal Completion: Support from Behavioral and Electrophysiological Measurements. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(2), 669–684.

Hedgé, J., Fang, F., Murray, S. O., and Kersten, D. (2008). Preferential Responses to Occluded Objects in the Human Visual Cortex. *Journal of Vision*, *8*, 16–35.

Hoerl, C. (2009). Time and Tense in Perceptual Experience. *Philosopher's Imprint*, *9*(12).

James, W. (1890). *The Principles of Psychology*. In 2 volumes. New York, NY: Henry Holt and Company.

Komatsu, H. (2006). The Neural Mechanisms of Perceptual Filling-in. *Nature Review Neuroscience*, *7*, 220–231.

Kosslyn, S. M., Behrmann, M., and Jeannerod, M. (1995a). The Cognitive Neuroscience of Mental Imagery. *Neuropsychologia*, *33*, 1335–1344.

Kosslyn, S. M., Thompson, W. L., Kim, I. J., and Alpert, N. M. (1995b). Topographical Representations of Mental Images in Primary Visual Cortex. *Nature*, *378*: 496–498.

Kösem, A., Gramfort, A., and van Wassenhove, V. (2014). Encoding of Event Timing in the Phase of Neural Oscillations. *NeuroImage*, *92*, 274–284.

Kovacs, G., Vogels, R., and Orban, G. A. (1995). Selectivity of Macaque Inferior Temporal Neurons for Partially Occluded Shapes. *Journal of Neuroscience*, *15*, 1984–1997.

Larsen, A., Madsen, K. H., Lund, T. E., and Bundesen, C. (2006). Images of Illusory Motion in Primary Visual Cortex. *Journal of Cognitive Neuroscience*, *18*(7), 1174–1180.

Leaver, A. M., van Lare, J., Zielinski, B., Halpern, A. R., and Rauschecker, J. P. (2009). Brain Activation during Anticipation of Sound Sequences. *Journal of Neuroscience*, *29* (8), 2477–2485.

Lee, G. (2014a). Extensionalism, Atomism, and Continuity. In N. Oaklander (ed.), *Debates in the Metaphysics of Time*. London: Bloomsbury.

(2014b). Temporal Experience and the Temporal Structure of Experience. *Philosopher's Imprint 14*(3).

Lee, S. H., Kwan, A. C., Zhang, S., et al. (2012). Activation of Specific Interneurons Improves V1 Feature Selectivity and Visual Perception. *Nature*, *488*, 379–383.

Lee, T. S., and Nguyen, M. (2001). Dynamics of Subjective Contour Formation in the Early Visual Cortex. *Proceedings of the National Academy of Sciences*, *98*(4), 1907–1911.

Lerner, Y., Harel, M., and Malach, R. (2004). Rapid Completion Effects in Human High-Order Visual Areas. *NeuroImage*, *21*, 516–526.

Muckli, L., Kohler, A., Kriegeskorte, N., and Singer, W. (2005). Primary Visual Cortex Activity Along the Apparent-Motion Trace Reflects Illusory Perception. *PLoS Biology*, *3*(8), 1501–1510.

Nanay, B. (2010). Perception and Imagination: Amodal Perception as Mental Imagery. *Philosophical Studies*, *150*, 239–254.

(2012). Perceptual Phenomenology. *Philosophical Perspectives*, *26*, 235–246.

(2018a). Multimodal Mental Imagery. *Cortex*, *105*, 125–134.

(2018b). The Importance of Amodal Completion in Everyday Perception. *I-Perception*. https://doi.org/10.1177/2041669518788887

(Forthcoming). *Seeing Things You Don't See*. Oxford, UK: Oxford University Press.

Page, J. W., Duhamel, P., and Crognale, M. A. (2011). ERP Evidence of Visualization at Early Stages of Visual Processing. *Brain and Cognition*, *75*(2), 141–146.

Pan, Y., Chen, M., Yin, J., et al. (2012). Equivalent Representation of Real and Illusory Contours in Macaque V4. *Journal of Neuroscience*, *32*, 6760–6770.

Pearson, J., and Westbrook, F. (2015). Phantom Perception: Voluntary and Involuntary Nonretinal Vision. *Trends in Cognitive Sciences*, *19*, 278–284.

Pearson, J., Naselaris, T., Holmes, E. A., and Kosslyn, S. M. (2015). Mental Imagery: Functional Mechanisms and Clinical Applications. *Trends in Cognitive Sciences*, *19*, 590–602.

Phillips, B. (2019). The Shifting Boundary between Perception and Cognition. *Noûs*, *53*, 316–346.

Phillips, I. (2010). Perceiving Temporal Properties. *European Journal of Philosophy*, *18*(2), 176–202.

(2011). Perception and Iconic Memory: What Sperling Doesn't Show. *Mind & Language*, 281–311. doi.org/10.1111/j.1468-0017.2011.01422.x.

(2014). Experience of and in Time. *Philosophy Compass*, *9*, 131–144.

Rauschenberger, R., Liu, T., Slotnick, S. D., and Yantis, S. (2006). Temporally Unfolding Neural Representation of Pictorial Occlusion. *Psychological Science*, *17*, 358–364.

Rees, A., Green, G., and Kay, R. H. (1986). Steady-State Evoked Responses to Sinusoidally Amplitude-Modulated Sounds Recorded in Man. *Hearing Research*, *23*(2). 123–133.

Regan, D. (1966). Some Characteristics of Average Steady-State and Transient Responses Evoked by Modulated Light. *Electroencephalography and Clinical Neurophysiology*, *20*(3), 238–248.

Rolls, E. T., and Tovee, M. J. (1994). Processing Speed in the Cerebral Cortex and the Neurophysiology of Visual Masking. *Proceedings of the Royal Society: Biological Sciences*, *257*, 9–15.

Scherzer, T. R., and Ekroll, V. (2015). Partial Modal Completion under Occlusion: What do Modal and Amodal Percepts Represent? *Journal of Vision*, *15*, 1–20.

Shepard, R. N. (1978). Mental Images. *American Psychologist*, *33*, 125–137.

Shibata, K., Watanabe, T., Sasaki, Y., and Kawato, M. (2011). Perceptual Learning Incepted by Decoded fMRI Neurofeedback without Stimulus Presentation. *Science*, *334*, 1413–1415.

Shimoji, S. (2014). Postdiction: Its Implications on Visual Awareness, Hindsight, and Sense of Agency. *Frontiers in Psychology*, *5*, 196.

Slotnick, S. D., Thompson, W. L., and Kosslyn, S. M. (2005). Visual Mental Imagery Induces Retinotopically Organized Activation of Early Visual Areas. *Cerebral Cortex*, *15*, 1570–1583.

Smith, F. W., and Muckli, L. (2010). Nonstimulated Early Visual Areas Carry Information About Surrounding Context. *PNAS*, *107*, 20099–20103.

Stekelenburg, J., Sugano, Y., and Vroomen, J. (2011). Neural Correlates of Motor-Sensory Temporal Recalibration. *Brain Research*, *1397*, 46–54.

Sterzer, P., Haynes, J. D., and Rees, G. (2006). Primary Visual Cortex Activation on the Path of Apparent Motion is Mediated by Feedback from hTM + V5. *NeuroImage*, *32*, 1308–1316.

Stetson, C., Cui, X., Montague, P. R., and Eagleman, D. M. (2006). Motor-Sensory Recalibration Leads to an Illusory Reversal of Action and Sensation. *Neuron*, *51*(5), 651–659.

Sugita, Y. (1999). Grouping of Image Fragments in Primary Visual Cortex. *Nature*, *401*, 269–272.

Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of Processing in the Human Visual System. *Nature*, *381*, 520–522.

Viera, G. (2019). The Fragmentary Model of Temporal Experience and the Mirroring Constraint. *Philosophical Studies*, 176, 21.

Vroomen, J., Keetels, M., de Gelder, B., and Bertelson, P. (2004). Recalibration of Temporal Order Perception by Exposure to Audio–Visual Asynchrony. *Cognitive Brain Research*, *22*, 32–35.

Vroomen, J., and Keetels, M. (2010). Perception of Intersensory Synchrony: A Tutorial Review. *Attention, Perception, and Psychophysics*, *72*, 871–884.

Zatorre, R. J., and Halpern, A. R. (2005). Mental Concert: Musical Imagery and Auditory Cortex. *Neuron*, *47*, 9–12.