The Time of Perception and the Other Way Around

Marcus Vinícius C. Baldo, André M. Cravo, and Hamilton Haddad
University of São Paulo (Brazil)

The world we perceive is delayed in relation to its flowing content, as well as the outcome of our actions on the world in relation to the moment we decide to act. This mosaic of different latencies permeating both perception and action has to be taken into account critically in order for us to cope with the temporal challenges constantly imposed by the environment. Fundamental notions, such as the sense of agency and causality, depend on the temporal relationship of events occurring in well-defined windows of time. Here, we offer a broad, yet abridged, historical view of some thought-provoking issues concerning the time of perception and action. From the pioneering work of Wundt, Titchener, and Libet to recent findings and ideas related to the employment of visual illusions as psychophysical probes (such as the flash-lag effect), we have tried to expose some problems inherent to the act of measuring the time of both perception and action, and devise possible solutions as well.

Keywords: flash-lag effect, time of perception, time of action, intentional binding

Correspondence concerning this article should be addressed to Marcus Vinícius C. Baldo, “Roberto Vieira” Sensory Physiology Lab, Department of Physiology and Biophysics, Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil. E-mail: baldo@icb.usp.br
When we look at a starry sky, we see a world that never existed in that way: many stars we still see died long ago, before others were born. The time elapsed since a photon leaves a star until it reaches our retinas varies from a few to billions of years. Surprisingly, although on a different scale, the same holds true for our daily life: the face we see, the voice we hear, and the hand we touch at the same time, engender sensory signals that take different times to travel throughout our neural processing stages, from receptors to the brain, before a bound and unitary percept is created. Moreover, transmission delays are by no means a feature inherent to perceptual processing only but are an integral component of motor behaviors as well.

On the one hand, behavioral timing is a crucial issue in survival, as adaptive success may depend on reasonably accurate judgments of temporal order and durations, as well as on precisely timed motor commands. On the other hand, differential transmission and processing latencies create a mosaic of temporal patterns spreading through the pathways and networks underlying both perception and action. A fundamental question thus immediately arises: How can we deal with such unrelated latencies in order to cope with the temporal challenges constantly imposed by the environment? We could split the above question into two. The first half addresses the generation of prompt skilled actions, which might involve predictive and automatic mechanisms, usually quite independent of conscious scrutiny but strongly dependent on previous learning experience. The second half of the question tackles the recognition that perception not only lags behind sensory reality but does it by spreading itself over a blend of temporal delays. For example, in a large concert hall, not only is all we perceive delayed in relation to the stimulating sources but also the asynchrony between what we see and what we hear might reach up to 400 ms (Durgin & Sternberg, 2002; Pöppel, 1997).

The psychology—and metaphysics—of time perception has a long and enthralling history. In the eighteenth century, astronomers experienced better than anyone else the difficulties in measuring the very instant of time when an event takes place. Using the so-called “eye and ear” method, they recorded the precise positions of celestial events. This method consisted of measuring the moment when a particular star crossed the meridian marker of a telescope, with the time instant of stellar crossing being estimated by counting the audible ticking of a pendulum clock. Spatial interpolation was then used to estimate the moment of transit in tenths of seconds.

Large discrepancies between trained astronomers often resulted in significant measurement errors and this variability led to the formulation of a “personal equation” to try to correct this problem (Boring, 1929). The link between astronomy and psychology was made by the father of experimental psychology (or physiological psychology, as he named it), Wilhelm Wundt (Schmidgen, 2003, 2005). Using a pendulum apparatus (Figure 1), he designed an experimental setup inspired by the eye and ear method. In a series of experiments, subjects had to judge the position of a fast moving clock hand when a brief click was presented. Wundt showed that participants systematically misperceived the position of the hand, perceiving it ahead or before the true position (the position where the hand physically was when the click was presented), depending on the experimental conditions. This was the prime example of a “complication experiment,” characterized by involving the perceptual comparison of a continuous and a sudden event. Wundt’s results could be accounted for by invoking either differences between visual and auditory perceptual latencies or different strategies adopted by the observers; more precisely, to which of the two events, visual or auditory, the observer was paying attention during the task. Among others, these findings led Titchener (1908) to formulate the law of prior entry: “the object of attention comes to consciousness more quickly than the objects which we are not attending.”

As we shall see, these two kinds of accounts—differential latencies and attention—are nowadays still offered as explanations of the outcomes of related psychophysical tasks used to assess the spatiotemporal features of sensory perception, such as the flash-lag effect (FLE) and temporal order judgments (TOJ). In psychophysical tasks involving either the FLE or TOJ, we can often observe large shifts in the psychometric functions, in comparison to the accompanying thresholds, which characterize them as true illusions (Klein, 2002). Whereas an illusion arising from a TOJ is purely temporal, the FLE represents, as we will see next, a perceptual illusion in space and time.

![Figure 1. Wundt’s pendulum apparatus used in “complication experiments” (from www.psichi.org/pubs/articles/article_457.asp)](image-url)
Illusions in Space-Time

The Flash-Lag Effect

As mentioned above, recurring disagreement between the measurements made by different astronomers at that time, led to the proposal that perceptual discrepancies might arise from the relative sharing of attention split between visual and auditory stimuli (Spence, Shore, & Klein, 2001). In a more contemporary and prosaic scenario, assistant referees often make mistakes during soccer games when judging the relative location of an attacking player at the moment the ball touches or is passed by one of his or her teammates (Baldo, Ranvaud, & Morya, 2002).

The common root of the above examples (and of any complication experiment in general) is the need to perform a dual task: where is the moving object (star or player) when the time-marker was presented (a given tick or pass). This question is intrinsically spatial and temporal in its very essence. In the midst of a variety of other visual illusions arising from the same question, the flash-lag effect (FLE) stands out as a still intriguing perceptual phenomenon, hotly debated over the last 12 years (Baldo & Klein, 1995; Eagleman, 2001; Eagleman & Sejnowski, 2000; Kirekelber & Lappe, 2001; Lappe & Keekelberg, 1998; Nijhawan, 1994, 2002; Nijhawan & Khurana, 2000; Purushothaman, Patel, Bedell, & Ogmen, 1998; Schlag & Schlag-Rey, 2002; Whitney, 2002). In the FLE, a stationary flash is perceived as spatially lagging behind a moving object, despite their being physically aligned to each other at that moment (Figure 2).

The FLE, in fact an almost 100-year-old illusion (Fröhlich, 1923; Mackay, 1958; Metzerger, 1931), was rediscovered in new clothing by Romi Nijhawan (Nijhawan, 1992, 1994). Nijhawan originally interpreted the flash-lag effect as resulting from a spatial extrapolation of the moving object (“motion extrapolation”), owing to the predictability of its trajectory. The perceptually extrapolated position of a moving object would thus compensate for the spatial error introduced by delays occurring throughout the visual system.

In 1995, the first challenge to Nijhawan’s explanation was put forward by Baldo and Klein. It was then demonstrated that the magnitude of the FLE was dependent not only on the features of the moving object (such as its speed), as previously shown by Nijhawan, but on the visual eccentricity of the stationary flash as well. This finding was at odds with the underpinnings of the motion extrapolation hypothesis. At that time, we interpreted the FLE “as resulting from a longer time delay involved in the visual processing of the flashing dots” (Baldo & Klein, 1995). We proposed that some amount of time, dependent on eccentricity, is required to bring the flashing stimulus to a sufficiently high degree of perceptual processing for a snapshot of the moving stimulus to be taken. Three years later, three independent groups reported empirical findings that, yet again, did not fit in the motion extrapolation scheme. These authors showed, for instance, the robust dependence of the FLE on the trajectory of the moving stimulus (Lappe & Kirekelberg, 1998; Whitney & Murakami, 1998) as well as on the luminance of both moving and stationary stimuli (Lappe & Kirekelberg; Purushothaman et al., 1998). From then on, the way was paved for a multitude of different approaches, views, and accounts concerning the flash-lag phenomenon (for reviews, see Eagleman, 2001; Kirekelberg & Lappe, 2001; Nijhawan, 2002; Nijhawan & Khurana, 2000; Schlag & Schlag-Rey, 2002; Whitney, 2002).

Our first demonstration of a putative role of visual attention in forging the flash-lag effect was based on a rather hazy empirical foundation (Baldo & Klein, 1995). The main idea was centered on the time delay required to shift attention from the flashing stimulus’ location to the moving stimulus’ pathway, in order to make a spatial judgment of their relative locations. The influence of visual eccentricity on attentional shifts was, however, confounded with its even stronger influence on basic sensory issues, such as temporal latencies and spatial resolution. Later on, we were able to refine our analysis, providing more convincing support to the contribution of visual attention to the FLE. The magnitude of the FLE was shown to be modulated by the spatial predictability of the flashing stimulus (Baldo, Kihara, & Klein, 2000; Baldo, Kihara, Namba, & Klein, 2002; Baldo & Namba, 2002) and its symbolic cueing as well (Namba & Baldo, 2004). Moreover, a modulatory role of either bottom-up or top-down deployment of attention was recently implemented in a simple feed-forward neural network (Baldo & Caticha, 2005), which was able not only to replicate a long list of empirical findings but also to encapsulate, in a unifying perspective, the main current accounts concerning the FLE.

Temporal Order Judgment

If we cannot determine the absolute timing of a subjective experience, all that can be done is to compare the relative timing of different experiences. This comparison is carried out in perceptual tasks involving temporal order...
judgments (TOJ), in which a pair of near-simultaneous stimuli are presented and observers are asked to judge which of them came first (Allan, 1975; James, 1890; Jaskowski, 1991; Spence et al., 2001; Sternberg & Knoll, 1973; Ulrich, 1987). The idea behind TOJ manipulation is that one stimulus of the pair could receive some kind of perceptual priority that would alter the time it takes to reach awareness. The factors that establish this perceptual priority could be grouped into pre-attentive and attentive processes (Haddad, Carreiro, & Baldo, 2002). Pre-attentive, or sensory, factors comprise some basic features of the sensory stimulus, such as its physical intensity or visual eccentricity (Haddad, Klein, & Baldo, 1999).

Attention is also assumed to influence the speed with which a stimulus will reach consciousness (Stelmach & Herdman, 1991), in line with the idea that the sensory information originating from an attended stimulus reaches some critical point in neural processing prior to the information arising from an unattended stimulus (Jaskowski, 1993; Stelmach & Herdman). Recently, the role of attention in temporal order perception has been investigated by means of the simultaneous measures of TOJ and event-related potential (ERP) recordings in the human visual cortex (McDonald, Teder-Sälejärvi, Russo, & Hillyard, 2005). These investigators observed that attention enhanced the amplitude of neural activity in the visual cortex, showing that attention-induced shifts in time-order perception probably arise from modulations of signal strength rather than from processing speed in the early sensory pathways.

The Time of Perception

The archetype of striking results obtained by employing TOJ and complication experiments (such as those leading to the FLE) is found in the seminal work carried out by Benjamin Libet and collaborators (Libet, 2002, 2004; Libet, Wright, Feinstein, & Pearl, 1979). Working with patients undergoing neurosurgical procedures, Libet and his group assessed latencies involved in both perception and action (see also the next section, The Time of Action). Their findings have been pungently debated since then, which has contributed to a clarification of concepts and interpretations related to the phenomenology of time in perception as well in action.

In a series of experiments, Libet and his group (Libet, Wright, Feinstein, & Pearl, 1979) compared the time of sensations arising from the electrical stimulation of the subject’s hand with the time of sensations caused by the direct electrical stimulation of the exposed somatosensory cortex (Figure 3). The delay involved in eliciting a sensation by means of a peripheral and natural stimulus includes the signal transmission through sensory pathways and the completion of a sequence of neural events in higher areas of the nervous system (which Libet called “neural adequacy”). Given that this latency, measured in previous experiments, is of hundreds of milliseconds, Libet and his colleagues expected the patients to report, for the stimuli presented simultaneously, that the cortical stimulation could be felt first. In fact, the inverse was observed, even in instances in which the cortex was stimulated before the hand. Libet and colleagues interpreted their results by proposing that the sensations evoked by the skin stimulation were referred “backward” to the time of the stimulus (Libet, 2002, 2004; Libet et al.; see also, on this topic, the special issue of Conscious and Cognition, Volume 11, 2002).

Libet’s findings and interpretations instilled some furor in the scientific arena, and were taken at face value by some scientists either to substantiate exotic physical proposals (Penrose, 1989) or to lend support to a dualist view of the neurosciences (Popper & Eccles, 1983). On the other hand, Libet’s work has been also severely criticized on both methodological and interpretational grounds (Churchland, 1981, 1986; Glynn, 1990; Gomes, 1998, 2002; Klein, 2002; Pocket 2002). A special kind of criticism, akin to the present discussion, concerns the representation of time by our brains (Dennett, 1991; Dennett & Kinsbourne, 1992; Glynn).

As sustained by the philosopher Daniel Dennett (Dennett, 1991; Dennett & Kinsbourne, 1992) and by the physiologist Ian Glynn (1990), the brain does not have to use time to represent time itself. Therefore, it would be improper to identify the perception of an event’s occurrence time with the time when we become conscious of it. That is, it is not necessary to postulate an isomorphism between the temporal order of the brain processes and the represented temporal order of events in the world. Instead of this, the nervous system would simply add time markers (or “stamps”) on
sensory inputs early in the series of neural events related to
the conscious perception of these inputs. When asked to
judge the time of occurrence of a given stimulus, subjects
could use these markers to order a sequence of mental
events. This approach remits to Dennett’s “content-vehicle”
distinction. As an example, we can consider what happens
when we receive a letter from the post office. If we are
interested in the time at which the events (content) in the
letter (vehicle) occurred, we look at the date on the letter
(time of content) rather than consider the date at which it
was delivered (time of vehicle). This line of reasoning is
quite similar to the “postdiction” account, devised by
Eagleman and Sejnowski (2000) in order to explain the
flash-lag effect. According to these views, perception is not
an “on-line” process, but “postdicted” to the instant of
relevant time-markers.

The Time of Action

In 1983, based on his earlier research, Libet and
collaborators started a sequence of classical experiments
that aimed to measure the occurrence time of a voluntary
motor decision (Libet, 1985, 2004; Libet, Gleason, Wright,
& Pearl 1983). They did this by asking the subject to report
the perceived location of a clock’s hand at the moment the
urge to move was felt and also when the movement was
felt to be executed. Their main findings were that a specific
electrical signal observed in brain activity, called readiness
potential (RP), occurred about 500 ms before the voluntary
act, whereas the subjective feeling of the urge to act was
reported roughly 150 ms before the voluntary act. According
to Libet, these results indicate that the brain starts the voluntary
process before the subject becomes consciously aware of the urge or wish, and the RP would be correlated
with the conscious intention. Haggard and Eimer (1999),
using a modification of Libet’s paradigm found that the lateralized readiness potential (LRP) is more likely to be
the correlate of conscious intention, rather than the RP. The
LRP is the point at which activity in the hemisphere contralateral to the selected hand exceeds the ipsilateral
activity associated with the non-selected hand. These results
would indicate that conscious intention is linked to the
specific preparation to perform a particular movement and
not to the earliest neural preparation to action. Though the
results found by Libet have had a long and often controversial history, an important contribution of these
results is the suggestion that before we became aware of
our intentions, several processes have already taken place
in our brain.

More recent psychophysical studies, employing a
procedure similar to Libet’s, have demonstrated that being
agent of an event might affect the perceived time of such
event (Haggard, Clark, & Kalogerias, 2002). In their
experiments, Haggard and colleagues asked subjects to
indicate the perceived time of either their actions or a beep
evoked by their action after a 250 ms delay. These results
were compared with blocks in which the subjects judged
either an action that did not evoke a beep, or a beep that
occurred without action. They found that subjects perceived
their action as shifted in time towards the beep that it caused,
and perceived the beep as shifted earlier in time towards
the action that caused it (Figure 4). In further experiments
they added a delay between the key press and the evoked
beeps. The results showed that this binding decreases as the
interval between action and effect increases. Additionally,
this binding effect is modulated by temporal contiguity and
temporal predictability and these authors refer to this linkage
as intentional binding (for a review, see Haggard, 2005).

An issue that arises is what mechanisms might be involved
in intentional binding. A possible explanation would be based
on feed-forward models. In 1950, Roger Sperry proposed that
when a motor command is sent to the muscles, an efferent
copy (also known as “corollary discharge”) is also sent to the
sensory areas of the brain. This efferent copy would reduce
the latency of the stimulus caused by the subject. Apparently,
even organisms as simple as arthropods have such a
mechanism (Webb, 2004). More recently, several feed-forward
models based on efferent copies have been proposed in order
to explain certain aspects of motor performance, such as motor
planning, prediction, and learning (Wolpert, 1997). The main
assumption of these models is that the preparation of a
movement would include an efferent copy sent to an internal
predictive model (Figure 5). This model estimates the effect
of the motor commands, allowing more rapid adjustments
and more fluent movement. Although these models were
originally proposed to explain motor performance, they
became valuable to understand which components of action
are conscious, and which are unconscious.

Other possibilities for intentional binding might be
suggested. As we discussed previously, attention is known
to have an important influence on the perception of an
event’s occurrence time, shifting the perception of an

![Figure 4. Intentional binding between voluntary actions and their
effects. When action and tone occur together, the perceived time
of the action shifts forwards in time from the baseline value,
towards the tone. The perceived time of the tone shifts earlier in
time, towards the action. Action and tone are bound together across
time, implying a reduction in the perceived interval between them
(modified from Haggard, 2005).](image-url)
attended event to an earlier instant in time. When one causes a stimulus, its temporal predictability can be considered as the highest possible (Eagleman & Holcombe, 2002), which might reduce its perceptual latency. This attentional hypothesis differs from Haggard’s proposal. According to him, the binding of action and its effect is a result of specific cognitive functions of our central nervous system, rather than a consequence of attentional processes. As for the attentional account, the preparation of a movement would include the allocation of attentional resources to both the expected sensory consequences and the time of execution of the movement itself. Again, we enter the question as to whether our nervous system utilizes time to encode time: Does the stimulus caused by the subject reach some neuronal population faster than a computer-generated stimulus because of some neural facilitation, or does our brain represent a self-produced stimulus in such a way that is later interpreted as occurring before in time? Although these processes might occur later in our perception, the facilitation may be due to processes that begin with the specific preparation of an action. Here too, we can see that the flash-lag effect can also be used as a useful tool to investigate the relationship between perception and action. As recently shown by Joan López-Moliner and Daniel Linares (2005, 2006), and also by us (Haddad & Baldo, 2005), when the time-marker (flash) is triggered by an active observer, the magnitude of the flash-effect, given by the relative location between the flash and a moving stimulus, is reduced when compared to a passive condition, in which the flash is not actively triggered. These results make clear the influence of action on space-time perception, although still not uncoupling attentional modulation from other cognitive mechanisms.

Figure 5. Schematic representation of a forward model. The forward model assumes the operation of an internal loop (broken lines) that leads the motor command to generate the anticipation of sensory changes, which can be then used to modulate the processing of the ongoing input (adapted from Webb, 2004).

Conclusion

Besides acting on the world we live in, we struggle to make sense of it. Actually, to perceive the world is to build mental models (representations) that should be useful in planning future actions. Therefore, we could say that our actions (as well as our ancestors’ actions) shape our perceptions. Similarly to the paramount role of timing in planning and executing adaptive actions, timing is also crucial to mold the final content of our perceptions. For instance, the natural and crystal-clear sense of causality depends on narrow temporal correlations between perceived events. Another vital notion, our sense of agency also depends on temporal correlations between what we perceive from the world and what we perceive from our own actions. Therefore, the physical timing of events is translated to the neural timing of neuronal processes, which is then somehow represented as a particular mental timing of experiential percepts (van de Grind, 2002).

In this paper, we have tried to show some problems inherent to the act of measuring the time of perception and action. Except for TOJ procedures, in which the comparisons are made between two abrupt events, in most cases, the comparison takes place between an abrupt event and a moving one (even when motion does not happen in the usual physical space, but within a more abstract perceptual space). It happened in Wundt’s, Libet’s and Haggard’s experiments, as described above, as well as in all procedures leading to the flash-lag phenomenon: the instant of occurrence of a sudden stimulus (the occurrence instant that one wants to measure) must be matched with the course of some temporal variable in continuous flow. It happens all the time in daily situations: for instance, as when someone asks us what time is it now. To provide this information, we have to look at the clock (which takes time) and then the very instant we intended to grasp has already passed. In the same way, in the experiments discussed here, our nervous system had to deal with the task of comparing the neural activity related to an abrupt event with the ongoing neural activity related to an event “moving” in sensory space. At least, this process should require an attentional shift from the abrupt-neural-activity to the moving-neural-activity and would, therefore, spend some amount of time. Hence, we are faced to a general problem concerning every attempt to measure the particular instant of occurrence of a perceptual event. This is a tremendous challenge because, as we saw, fundamental difficulties arise whenever subjects have to perceptually crystallize an instant of time.

As a last remark, we should not overlook the very essence of time, with its physical and metaphysical meanings. These aspects are to be included in any serious debate on the relationship between time and perception; otherwise we could come up, for example, with the following original conclusion: “there are three times, […] for these three exist in the mind, and I find them nowhere else: the
present of things past is memory, the present of things present is sight, the present of things future is expectation.” Although these words may sound quite contemporary and, as a matter of fact, they may express current views on time perception, they were stated by St. Augustine about 1600 years ago.

References


James, W (1890). The principles of psychology. New York: Dover.


Received December, 11, 2006
Revision received May, 16, 2007
Accepted July, 24, 2007