Cognitive Critique

TIME, OBSERVATION, MOVEMENT

Myrka Zago

Laboratory of Neuromotor Physiology
Santa Lucia Foundation, Rome

EMAIL: mzago@hsantalucia.it

Mauro Carrozzo

Laboratory of Neuromotor Physiology
Santa Lucia Foundation, Rome
Institute of Cell Biology and
Neurobiology Institute National Research Council, Rome

Alessandro Moscatelli

Laboratory of Neuromotor Physiology
Santa Lucia Foundation, Rome
Department of Neuroscience
University of Rome Tor Vergata, Rome

Francesco Lacquaniti

Laboratory of Neuromotor Physiology
Santa Lucia Foundation, Rome
Department of Neuroscience
Centre of Space BioMedicine
University of Rome Tor Vergata, Rome

EMAIL: lacquaniti@caspur.it

ABSTRACT

Traditionally, a sharp distinction was made between conscious perception of elapsed time, considered a key attribute of cognition, and automatic time processes involving basic sensory and motor functions. Recently, however, this dichotomous view has been challenged on the ground that time perception and timed actions share very similar features, at least for events lasting less than a second. For both perception and action, time estimates require internally generated and/or externally triggered signals, because there is no specific sensory receptor for time in the nervous system. We argue that time can be estimated by synchronizing a neural time base either to sensory stimuli reflecting external events or to an internal simulation of the corresponding events. We review evidence in favor of the existence of distributed, specialized mechanisms, possibly related to brain mechanisms which simulate the behavior of different categories of objects by means of distinct internal models. A critical specialization is related to the animate-inanimate distinction which hinges on different kinematic and kinetic properties of these two different categories. Thus, the time base used by the brain to process visual motion can be calibrated against the specific predictions regarding the motion of biological characters in the case of animate motion, whereas it can be calibrated against the predictions of motion of passive objects in the case of inanimate motion.

INTRODUCTION

Humans are able to analyze time information across a wide range of intervals, from the microsecond timing of sound localization (delay between sound arrival at the left and right ear) to the 24-h period of circadian rhythms and beyond. The relative precision in the discrimination of time intervals is equivalent to a Weber fraction of psychophysics. It corresponds to the ratio of the difference threshold in discrimination of time estimates and the magnitude of the base time interval. On average, this ratio varies to a limited extent over almost the full range of time intervals relevant for our daily activities. Thus, the average ratio remains roughly close to about 10% over more than ten orders of magnitude of the base time interval (Gibbon 1977). However, as we will discuss later, the accuracy and precision of individual time estimates may differ greatly as a func-

tion of the specific task and context. Moreover, the neural substrates and mechanisms most likely differ as a function of the specific time scale in question (Buhusi and Meck 2005; Lewis and Miall 2003; Mauk and Buonomano 2004).

Given the breadth and heterogeneity of the issues at stake when dealing with time processing in general, here we focus only on events unfolding over the scale of tens to hundreds of milliseconds. Moreover, we limit our discussion to timing issues in visual and motor tasks. Time processing over the sub-second scale is crucial for several such tasks. Just to name a few examples, challenging activities such as those involved in performing sports, dancing, or playing music, but also (apparently) less demanding activities such as walking, speaking, or watching a movie, all require accurate time estimates. In brief, time processing over the sub-second scale is essential for most motor and visual activities, and is a prerequisite for successful interactions with people or objects in the environment.

Generally, time cannot be directly measured at a given moment because, in contrast with many other physical quantities, it does not have any specific sensory receptor in the nervous system. Instead, time estimates in the brain might be derived by integrating appropriate information over discrete intervals, but this integration requires internally generated and/or externally triggered signals over the interval to be estimated. These signals would be used to calibrate unit time intervals. According to the scalar expectancy theory (Gibbon 1977), an intrinsic pace-making mechanism generates a series of pulses (metaphorically equivalent to the ticks of a clock) that are concatenated together in larger time bins. These bins would be subsequently counted, and a comparing system would connect the output of the counting process with memorized time representations to estimate the overall duration of the time epoch under scrutiny. This idea is tightly related to the hypothesis of a centralized internal clock, in which an oscillator beating at a constant rate generates pulses that are detected by a counter (Creelman 1962; Treisman 1963). This hypothesis assumes that timing is centralized, in other words it assumes that the same neural circuitry is used irrespective of the sensory modality through which the stimulus has been acquired, for example, irrespective of whether the duration of an auditory tone or the duration of a visual flash are monitored. However, despite the venerable history of the central clock hypothesis and some psychophysical data in its support, no convincing neurophysiological evidence has been provided so far for a central clock.

Therefore, the view that events are timed by a centralized supra-modal clock recently has been called into question in favor of distributed, specialized mechanisms. According to the latter view, several brain regions would be capable of temporal processing and the specific region or network involved would depend on the task and sensory modality being used. The counter model described above is also applicable within the context of specialized time encoding, but in the latter case there would exist several such counters distributed in the brain instead of a unique, centralized one.

It should be noted, however, that on theoretical grounds the hypothesis of a neural clock – whether a centralized clock or one of several distributed clocks – might require physiologically implausible values of oscillation frequency for discriminations with millisecond resolution (Mauk and Buonomano 2004). To circumvent this objection, it has been proposed that time is encoded by the spatio-temporal patterns of activity in multiple neural populations (Mauk and Buonomano 2004). When a stimulus arrives in the network, it will engage hundreds of excitatory and inhibitory neurons. Some of them will fire action potentials, but time-dependent processes (such as short-term synaptic plasticity) might also be triggered. Because of these time-dependent properties, the network will be in a different state when another stimulus arrives some tens of milliseconds later. Consequently, the same input can activate different subpopulations of neurons as a function of the recent stimulus history of the network, and the difference in the network activity produced by the second and first stimulus can be used to code for the time interval separating the two stimuli. Another theory posits that subjective duration mirrors the amount of neural energy (or the total amount of neural activity) used to encode a stimulus (Pariyadath and Eagleman 2007). It should be noted, however, that these theories remain difficult to test experimentally. We will return later to the issue of possible neural substrates of time estimates in the brain.

TIME DISTORTIONS

Observers asked to judge the duration of brief events are prone to errors: perceived duration often does not match the physical duration of the event. As in the case of visual illusions, distortions of perceived time represent an important field of research because they can reveal some of the neural mechanisms underlying the processes

involved in encoding time. For instance, it has been shown that the apparent duration of a dynamic stimulus can be reduced in a local region of visual space following motion adaptation (Johnston et al. 2006), and the effect of this adaptation may be spatially selective either in retinal (Bruno et al. 2010) or real-world coordinates (Burr et al. 2007), allowing to separately time targets placed in different locations of space. Although the retinocentric (retinal coordinates) and the allocentric (real-world coordinates) representation of time have been cast by these two groups as two contrasting hypotheses, they are not mutually exclusive; time encoding might be retinocentric or allocentric depending on the context. Indeed, Johnston et al. (2006) and Bruno et al. (2010) showed that local visual adaptation to gratings induces a shortening of perceived duration only for the adapted retinal position, demonstrating directly a retinocentric selectivity of the timing mechanisms. Burr et al. (2007), instead, showed that the alteration is not related to an adaptation of the early analysis of the visual system but occurs only when the adapter and the event to be timed occupy the same position in external space, not on the retina, demonstrating an allocentric selectivity of the timing mechanisms. Interestingly, in both cases, the effect is selective for the temporal frequency of the adapting grating, being stronger at high temporal frequencies (Johnston et al. 2006). These observations point to the existence of localized, essentially visual components involved in sensing the duration of visual events. The spatial specificity (whether retinocentric or allocentric) of the adaptation effect cannot be easily explained in the context of the hypothesis of a single, supramodal central clock, while it can be accounted for by a mechanism in which separate temporal signals are scaled and integrated to deliver duration estimates locally to a particular visual position. Because there is evidence that spatially localized timers might not be monitored simultaneously (Morgan et al. 2008), these studies suggest the surprising notion that the sense of elapsed time might be fragmented in discrete, modality-specific spatial patches.

Apparent duration depends on several factors which are specific to the stimulus or the context (Eagleman 2008). Thus, it depends on the stimulus visibility (Terao et al. 2008), speed (Kaneko and Murakami 2009), temporal frequency (Kanai et al. 2006), predictability (Pariyadath and Eagleman 2007), as well as the level of attention (Tse et al. 2004), the intention to perform an action (Haggard et al. 2002), or saccadic eye movements (Morrone et al.

2005). The effects of predictability can be demonstrated by show-
ing a given stimulus repeatedly; the first stimulus is typically judged
to have a longer duration than the successive stimuli (Pariyadath
and Eagleman 2007). Similarly, a stimulus different from all the
others in a repeated series (a so-called *oddball* stimulus) is judged
to have lasted longer than the other stimuli with the same physical
duration. In the framework of the counter model described above,
an increment of arousal determined by the appearance of the unex-
pected oddball stimulus would lead to a transient increase in the rate
of an internal clock. Therefore, the accumulator would sum a larger
number of ticks in a given time epoch, and the stimulus duration
would be judged accordingly longer for the oddball. The same phe-
nomenon can also be explained in the context of the energy readout
model (Pariyadath and Eagleman 2007). In higher cortical areas,
neuronal firing rates tend to diminish in response to repeated pre-
sentations of stimuli. For instance, functional magnetic resonance
imaging (fMRI) studies have revealed a similar decrease in BOLD
(blood oxygenation level dependent) responses to repeated presen-
tations of stimuli as compared with the responses to novel stimuli.
Therefore, one can interpret the results as if successive stimuli were
contracted in duration, rather than the oddballs being expanded in
duration.

Saccadic chronostasis refers to the subjective temporal length-
ening of a visual stimulus perceived after an eye movement (Yarrow
et al. 2001). Observers are asked to judge the duration of a stimulus
that changes shape or color at an intermediate time during a saccade,
so that it can only be perceived in its new state at re-fixation because
of the known decreased visibility of stimuli during saccades. Test
stimulus duration must be judged by comparison with a subsequent
reference stimulus, allowing a matched estimate to be derived (the
point of subjective equality). In this type of study, it is consistently
found that observers overestimate the time for which they have seen
the saccadic target compared to constant fixation conditions.

The effects of visual motion on subjective duration are espe-
cially interesting. It has long been known that the perceived dura-
tion of a rapidly moving stimulus is generally longer than that of
a slower or stationary stimulus having the same physical duration,
a phenomenon known as *time dilation* (Brown 1995; Lhamon and
Goldstone 1974; Kanai et al. 2006; Kaneko and Murakami 2009).
Time dilation has been interpreted in the context of the idea that
the brain estimates time based on the number of events that occur

(Fraisse 1963; Brown 1995). In other words, the occurrence of a greater number of events would be taken by the brain as evidence for a longer duration. Salient events, such as changes over time of the visual stimuli, index the passage of time, so that we know how much time has passed by counting these indices. Stimulus motion is accompanied by continuous changes in spatial location and thus provides important temporal indices. There is some controversy, however, as to whether stimulus speed or its temporal frequency or spatial frequency represents the critical element in the time dilation phenomenon. Thus, Kanai et al. (2006) showed that the time distortion could be determined simply by a flickering stimulus, consistent with the idea that the temporal frequency is the key factor. However, Kaneko and Murakami (2009) found that the speed of the stimulus, rather than temporal frequency or spatial frequency per se, best described the perceived duration of a moving stimulus, with the apparent duration proportionally increasing with the logarithm of speed.

Time perception (and possible distortions) associated with real visual motion might be affected crucially by low-level factors, although the critical role of attention in time estimates has already been mentioned. For instance, it is well established that prolonged exposure to a moving pattern influences the perceived speed of subsequent moving patterns (Thompson 1981). This effect has been explained as resulting from speed encoding by distinct temporal filters (corresponding to low-speed and high-speed channels) whose sensitivities decay exponentially over time (Smith and Edgar 1994; Hammett et al. 2005). Time estimates for a moving stimulus might be affected by similar mechanisms. Indeed, the localization of temporal duration effects to the adapted region of the visual field demonstrates that essentially peripheral, spatially localized processes are also involved in the encoding of duration (Johnston et al. 2006). However, rather artificial stimuli are often used to investigate low-level visual mechanisms: random dots, sinusoidal gratings or Gabor functions have been mainly used as real motion stimuli. When one employs more naturalistic images, higher-order cognitive factors also become relevant, as we will see in the next sections.

OBSERVATION OF BIOLOGICAL ACTIONS

Human and animal movements exhibit a number of different regularities. For instance, hand movements exhibit a systematic relationship between the curvature of the path and the movement velocity, formalized as a 2/3 power law (Lacquaniti et al. 1983). Also the temporal coordination of the angular motion of different limb segments obeys simple laws, so that the motion at each segment is linearly coupled with that at other segments for both the upper limbs (Soechting and Lacquaniti 1981) and the lower limbs (Borghese et al. 1996). None of these regularities is dictated by physics alone, but arises due to both neural control and biomechanical constraints. Our visual system is sensitive to the presence of these motion regularities in an observed scene, inasmuch as alterations of the normal biological kinematics are easily detected (Bidet-Ildei et al. 2006; Cook et al. 2009; Chang and Troje 2009). Moreover, when observers are asked to determine which motion appears to be most uniform, they tend to choose motions close to obeying the 2/3 power law instead of truly uniform motions (Viviani and Stucchi 1992).

There is now evidence that also time perception is tuned to biological movements. First, spontaneous movement tempo is influenced by the observation of rhythmical actions (Bove et al. 2009). Indeed, the frequency of self-paced finger opposition movements, which is normally around 2 Hz, was biased towards the frequency of previously observed, similar movements which were performed at either 1 or 3 Hz, and such changes were lasting, being still present two days after the visual exposure. Second, movement timing can be modified according to others' movements even when the observed and the to-be-executed movements are completely unrelated. Watanabe (2008) generated sequences of point-light biological motion from videotapes of an individual performing twenty-six different actions (jumping, running, walking, kicking, throwing a ball, etc.). For comparison, scrambled versions of these biological motion sequences and solid object motion were also generated. Observers viewed one of these stimulus sequences presented at three different (half, normal, and double) rates. After a blank variable period, the observers performed a simple choice reaction-time task that was unrelated to the presented stimulus sequence. The observation of the biological motion produced a negative correlation between reaction time and stimulus rate (i.e., faster biological motions resulted in shorter reaction times), but no

such tendency was found with the scrambled or solid object motion. Watanabe (2008) interpreted these results as a speed contagion from observed actions to own actions.

Carrozzo et al. (2010), instead, investigated in different experiments the manual interception of a moving target or the discrimination of the duration of a stationary flash. In both cases, the observers were presented with different background scenes prior to and during the execution of the task. In separate sessions, the scene displayed characters which could differ in terms of human or artificial appearance and kinematics, while the low-level features of the stimuli were matched as much as possible across conditions. The human character performed a sequence of steps from classical ballet (recorded from a real dancer), while the non-biological character consisted of a rigid object consisting of fourteen disconnected rods whose motion matched the first few harmonics of the human dancer. It was found that time estimates were systematically shorter in the sessions involving human characters moving in the background than in those involving non-biological moving characters. Strikingly, the animate/inanimate context also affected randomly intermingled trials which always depicted the same still character, demonstrating that dynamic cues were not necessary to elicit the interference, but that the biological or non-biological context was crucial. Importantly, the time distortion induced by watching the human figure was graded as a function of the biological plausibility of its motion. Thus, the distortion was greatest when the human figure was animated using the original human data measured by means of motion capture, and it decreased when the same figure was shown in an unrealistic, upside-down configuration, as well as when the original data were phase-shuffled violating the 2/3 power law and severely degrading the impression of a smooth dance. These results therefore provide evidence for an influence of human animacy on time estimates: visual event timers might be tuned to targets in space according to the specific natural features of the stimuli, their animacy being one of the most basic features, implicating high-level mechanisms for time modulation. Animacy affects neural time, but conversely specialized temporal entrainment might contribute to animacy attribution.

In sum, these studies show that, when a background animated scene is displayed during the concurrent or subsequent execution of an unrelated task, the responses of the unaware observer become driven, at least in part, by the dynamics of the background charac-

ter. One may argue therefore that the time base used by the brain to process visual motion is calibrated against the specific predictions regarding the motion of biological figures. This idea is rooted in the so called motor resonance theory. When we see someone moving, our brain may covertly simulate the observed action (Shepard 1984; Jeannerod 1994). A neural correlate of motor simulation or motor resonance has been described in premotor and posterior parietal areas of the monkey, where 'mirror' neurons respond when the monkey performs or views a specific action (Rizzolatti and Craighero 2004). The studies on timing reviewed above are compatible with two different possibilities. A motor resonance might be obtained by synchronizing neural time to a time base intrinsically linked to the internal simulation of the observed action. Alternatively, motor resonance might involve the synchronization of the internal simulation with the observed action, and this could be achieved through the recalibration of a specific internal time base.

The specialization of the neural time estimates as a function of animacy and biological motion would enhance the temporal resolution of visual processing and the ability to predict critically timed events. This specialization might explain why people are so accurate at predicting the timing of others' actions (Sebanz and Knoblich 2009). Interestingly, it has been shown that visual discrimination of point-light motion of two interacting agents is worse when the two actions are desynchronized (Neri et al. 2006), suggesting that time-locking in a behaviorally meaningful way between interacting agents provides an implicit temporal cue and the additional agent can be used to predict the expected trajectory of the relevant agent with better precision.

In Carrozzo et al. (2010), observed action (dance) and performed action (ball interception) were totally unrelated in terms of both the movement type involved and the action goal. Similarly, in time discrimination tasks (Carrozzo et al. 2010; Orgs and Haggard 2011) the target is typically unrelated to the background action. But what happens when the observed action is instrumental for performing the task? It is known for instance that, in ball games involving two-player interactions, performers are able to identify key information about the forthcoming ball trajectory from the movement patterns of the opponent player (Abernethy 1990; Bahill et al. 2005; Huys et al. 2008; Shim et al. 2005). Zago et al. (2011) asked the question of whether the observation of a human action helps to decode environmental forces during the interception of a decelerating

target within a brief time window, a task intrinsically very difficult. The target (a moving ball) decelerated under virtual gravity while reaching the interception point. Zago et al. took advantage of the scene inversion paradigm which consists in presenting the visual scene upright or upside-down. It is well known that the recognition of scenes, people and actions is faster and more accurate when they are right-side up, that is, when they are aligned with the observer (see, for example, Chang et al. 2010; Reed et al. 2003; Yin 1969). Thus, when a digitally edited photograph of a face is presented upside-down relative to the observer, the ability to detect gross distortions and abnormalities is strongly impaired and the responses are slowed down (Thompson 1980), mainly because of a deficit in coding configural information (Freire et al. 2000). Similar viewer-centered inversion effects have been described for the discrimination of static whole-body postures (Reed et al. 2003) and of biological motion in point-light walk stimuli (see, for example, Chang et al. 2010). Zago et al. (2011) employed a factorial design to evaluate the effects of scene orientation (normal or inverted) and target gravity (normal or inverted) on an interception task. In different protocols, button-press responses triggered the motion of a bullet shot by a static human character (Bullet protocol), of a sliding piston (Piston), or of a punching arm of a dynamic human character (Arm). By design, biological movements of the Arm protocol preceded and partly overlapped target ball motion. The most obvious outcome of these experiments should be that the scene orientation prevailed, given the established potency of featural and configural cues built-in to polarized visual scenes (Reed et al. 2003; Yin 1969). Accordingly, one would expect the interception to be more successful with normal upright scenes than with upside-down scenes, independently of the direction of target gravity, with perhaps a further performance bias in favor of the default condition where both scene and gravity are normally oriented.

Consistent with this expectation, it was found that the average timing errors were smaller for upright scenes irrespective of gravity direction in the Bullet protocol, while the errors were smaller for the default condition of normal scene and gravity in the Piston protocol. In the Arm protocol, instead, performance was better when the directions of scene and target gravity were concordant, irrespective of whether both were upright or inverted. Strikingly, average performance in the latter protocol was superior when both the scene and target gravity were inverted than when the scene was upright

but target gravity was inverted. These results suggest that the use of classical viewer-centered reference frames is binding with in-animate scenes, such as those of the Bullet and Piston protocols. Instead, the presence of biological movements in animate scenes (as in the Arm protocol) may help processing target kinematics and timing the interceptive action under the ecological conditions of co-herence between scene and target gravity directions, downplaying the relevance of viewer-centered reference frames.

Finally, Viviani et al. (2011) considered the special case of speech movements by focusing on the accuracy and precision with which observers detect departures from the normal speech rate in observed silent speech. Muted video-clips of the lower face of speaking actors were shown to the observer at a variable rate, both faster and slower than the original rate. Observers had to identify the speech rate closest to the natural one. It was found that esti-mates were accurate when the video-clips were shown as recorded (i.e., played in the forward direction). However, speech rate was significantly underestimated when the video-clips were shown in the backward direction. Because the duration of articulatory ges-tures was the same irrespective of the direction of movement, the bias found experimentally implies a difference in the way durations are perceived depending on movement direction. In other words, perceived duration might depend on features of phono-articulatory kinematics that are not invariant with respect to time-reversals. Also this series of studies points to a specific tuning of time perception to biological movements.

APPARENT AND IMPLIED MOTION

In addition to real motion, apparent motion and implied motion are also able to drive neural timers, pointing to top-down modulatory effects on time estimates. Orgs and Haggard (2011) presented two pictures of the initial and final positions of a human movement separated by different time intervals. At the appropriate inter-stimulus interval, this presentation typically evokes the sensation of apparent motion. The shortest movement path between two positions was always biomechanically impossible. At longer inter-stimulus intervals, participants tended to see a longer, feasible movement path, consistent with previous reports (Chatterjee et al. 1996). At these same long time intervals, participants underestimated the duration of presentation of a square surrounding

the picture sequence, compared to trials displaying degraded body pictures, consistent with the results of Carrozzo et al. (2010) with real motion. Therefore, the perception of apparent biological motion may involve temporal binding of two static pictures into a continuous motion. Temporal binding might depend on top-down mechanisms generating a percept of biological motion in the absence of any retinal motion.

Time-frozen images of an action may convey dynamic information about previous and subsequent moments of the same action, thus providing an impression of motion (so-called implied motion). For example, the overall body posture of people depicted in a photograph may suggest that they were either in motion or still while the photograph was shot. Several psychophysical studies have shown that human observers take into account the presence of dynamic information implicit in action photography. Thus, images which imply motion cause a forward displacement in spatial memory, a phenomenon known as representational momentum (Freyd 1983; Hubbard 2005). In one of these studies (Freyd 1983), subjects were asked to decide whether two photographs presented in succession were the same or different. Each pair of images was taken from a film showing a moving target (human, animal or object). In different trials, the second image in the pair was the same as the first one, or it was a slightly earlier image or a slightly later one. It was found that the observers responded more sluggishly and made more errors in the discrimination when the pair was in forward (real-world) temporal order than when it was in backward (reversed) order, consistent with the idea that the observers memorized the previously depicted target as continuing in motion. Another study (Winawer et al. 2008) showed that viewing a series of photographs with implied motion in a given direction produces motion after-effects in the opposite direction.

Because the perception of real and implied motion may share some common properties and neural substrates, it is reasonable to ask whether implied motion affects the perceived duration of a stimulus, just as real motion does. Two recent studies independently addressed this issue. In one study (Nather et al. 2011), the participants were presented with pictures depicting two different sculptures of ballet dancers made by the impressionist artist Edgar Degas. One sculpture represented a low-arousal body posture judged to require no movement, while the other one represented a high-arousal body posture judged to require considerable movement. The question was

whether the perception of presentation durations of the different body postures was distorted as a function of the embodied movement that originally may have produced these postures. In a temporal bisection task with two ranges of standard durations (0.4-1.6 s and 2-8 s), the participants had to judge whether the presentation duration of each of the pictures was more similar to the short or to the long standard duration. Nather et al. (2011) found that the duration was judged longer for the posture requiring more movement than for the posture requiring less movement. The magnitude of this overestimation was relatively greater for the range of short durations than for that of longer durations, therefore resulting in a significant interaction between the duration of the stimulus and the body posture of the picture. These authors interpreted the result in the framework of the internal clock model, and suggested that this lengthening effect was mediated by an arousal effect of limited duration on the speed of the internal clock system. In this respect, then, implied motion produces time dilation just as real motion does (see above).

In the second study (Moscatelli et al. 2011), observers were asked to compare the variable presentation duration of test photographs with the reference duration of their scrambled version. The test stimulus was a photograph with implied motion or a photograph without implied motion, both taken from the sport field (martial arts, skating, soccer, diving, volley, rugby). Forty different photographs were used (twenty for each type). The two sets of photographs (with and without implied motion) were matched between each other in terms of critical low-level features (luminance and salience), as well as high-level features related to the content (sport discipline, sex and number of depicted athletes), to avoid introducing potential attentional biases. It was found that the duration of photographs with implied motion was discriminated better than the duration of photographs without implied motion. It was also found that the average reaction time for the discrimination of photographs with implied motion was longer than that for photographs without implied motion, suggesting that the processing of implied motion involves longer and/or slower neural routes to compute time duration. This longer processing may depend on the engagement of two visual systems in parallel, one for processing form and the other one for processing implied motion. The perceptual decision about time duration would occur after the convergence of signals from these two pathways. The finding of slower response times with implied motion is also consistent with the report already mentioned

by Freyd (1983). By the same token, one can interpret the results of a better discrimination and slower processing time with implied motion pictures as indicative of an active processing of the implied motion cues of the images for mental imagery of the associated action. In a similar vein, Krekelberg et al. (2003) suggested that an observer might take advantage of implied motion cues to deal with the complexity of motion perception in a natural environment.

The psychophysical results on implied motion can be reconciled with observations derived from brain imaging and stimulation studies. Several neuroimaging (fMRI) studies showed that observation of implied motion images activates the cortical area hMT/V5+, a key region for processing real visual motion (Kourtzi and Kanwisher 2000; Krekelberg et al. 2003; Lorteije et al. 2006). In addition, by transiently disrupting the activity of hMT/V5+ by means of trans-cranial magnetic stimulation (TMS), it has recently been suggested that this area contributes to keep the time of visual events in both motor (Bosco et al. 2008) and perceptual tasks (Bueti et al. 2008). In particular, Bueti et al. (2008) showed that TMS of hMT/V5+ interferes with the discrimination of the duration of either static or moving visual stimuli by reducing the precision of the temporal judgments. One can then speculate that temporal discrimination of photographs with implied motion may engage neural populations in area hMT/V5+ (and possibly other visual motion areas) to a greater extent than the same task with photographs without implied motion.

CALIBRATION OF TIME ESTIMATES AGAINST PHYSICAL LAWS OF MOTION

We discussed above the existence of significant distortions in time estimates due, for instance, to attentional demands, eye movements, target motion. This then raises the question of whether there are internal mechanisms to keep time estimates calibrated so that distortions are the exception rather than the rule. One possibility is that the brain constantly calibrates its time estimation against physical laws from the outside world (Eagleman 2004). This hypothesis is especially relevant in the case of estimates of the duration of a target motion. In fact, the position of a moving object at a given time in the near future might be predicted by a forward internal model (Davidson and Wolpert 2005; Zago et al. 2009) and compared with sensory feedback to keep the perceived time calibrated (Eagleman 2004). Evidence reviewed in the previous

section suggests that we are most probably endowed with models of biological movements. Here we consider the possibility that the brain is also endowed with some kind of model of Newtonian mechanics. In particular, target motion under gravity represents a highly predictable event. Because Earth's gravity is locally constant, the motion of any object accelerated by gravity has a fixed duration over a given path (neglecting air friction). In fact, the main timekeepers used by mankind for centuries – the water clock and the pendulum clock – relied on gravity. Implicit knowledge about the direction and the effects of gravity could be used by the brain for consistent time-keeping. The ability to detect unnatural features in vertical visual motion related to gravity can be demonstrated early in life. Between 5 and 7 months, infants begin to implicitly expect a downwardly moving object to accelerate and an upwardly moving object to decelerate, as shown by an abnormally prolonged attention to displays of the unnatural motion (Kim and Spelke 1992; Friedman 2002). In adults, the manual interception of a vertically falling ball is accurately timed (see Zago et al. 2009), as is the indication of the time of landing of a computer-animated target that rolls off a horizontal surface and falls hidden from view (Huber and Krist 2004). Furthermore, the final position of a horizontally moving target which is suddenly halted is misremembered as being displaced downward below the path of motion, consistent with the idea that gravity effects are implicitly assumed by the observers (Hubbard 1995).

Notice in this regard that a critical role of kinetics in shaping time estimates has been suggested in the context of cognitive motor control (Georgopoulos 2002). Georgopoulos (2002) argued that spatial and temporal aspects of movement involve separate neural processes, and that temporal aspects may be controlled by coupling an action to an intrinsic temporal function such as the tau-guide (Lee 1998). Consistent with this idea, a tau-guide model taking into account gravity was able to account for the assessment of temporal duration made by subjects instructed to throw a ball in the air and to predict when the ball would hit the floor by tapping the floor with a stick, with or without vision (Grealy et al. 2004).

Moscatelli and Lacquaniti (2011) considered the possibility that an internal model of gravity also affects the perceptual judgment of temporal duration of a visual motion. If relative time in perceptually structured displays is efficiently encoded because of the availability of models of physics, time discrimination should be more precise

when the motion of an object complies with gravity constraints than when it artificially violates such constraints. They tested this hypothesis by comparing the time discrimination for linear motion of a virtual object across the four cardinal directions: downward, upward, rightward, and leftward. With a constant positive value of target acceleration, target kinematics was congruent with the effects of gravity only for the downward direction, and one would expect a corresponding anisotropy in time discrimination: durations should be discriminated more precisely during downward motion than during the other motion directions. The results confirmed this prediction, irrespective of whether target motion was embedded in a pictorial scene, including several metric cues (familiar size, linear perspective, shading, and texture gradient), or in a quasi-blank scene lacking any metrics. However, the gravity-related anisotropy was more pronounced in the former than in the latter case.

Next, Moscatelli and Lacquaniti (2011) addressed the issue of whether the sensitivity to gravity constraints is tied to retinal (or other egocentric) coordinates, to Earth's gravity, or to visual references intrinsic to the scene. To this end, the same stimuli were used, but the observer was tilted by 45° relative to the monitor and Earth's gravity in one experiment, while the observer was upright and the monitor was tilted by 45° in another experiment. In both cases, pictorial downward was tilted relative to the retinal vertical meridian. Nevertheless, it was found that the discrimination precision was still higher for targets directed downwards relative to the pictorial vertical, although tilting decreased the size of the effect. By contrast, the anisotropy essentially disappeared with the non-pictorial scene once target motion was oblique.

Overall, the difference in precision between downward and upward motion was not constant across experiments, but varied in a graded manner as a function of the conditions, being highest when both the observer and the pictorial scene were upright and lowest when the target direction in the non-pictorial scene was tilted by 45° relative to an upright observer. To model this graded behavior, a linear combination of the 3 types of cues was used: pictorial cues (P), orientation of the observer (O) and orientation of target motion (T) relative to the physical vertical. The resulting weighing coefficients were 43, 37 and 20% (of the overall response) for O, T and P, respectively. The observation that egocentric cues specifying the observer's orientation (O) dominate is in line with much previous work on the perceptual discrimination of scenes, people and actions

(Chang et al. 2010). On the other hand, the substantial contribu-
tion of visual references intrinsic to the scene, such as the direction
of target motion (T) and the presence of additional pictorial cues
(P), agrees with the previous observation that viewing a photograph
with strong polarization cues, which indicate relative up and down
directions in the picture, can alter the perceived direction of abso-
lute up and down directions in the real world (Jenkin et al. 2004).

These results suggest that spatial representations for computing
time are flexible, and may be anchored to a variety of different
egocentric and allocentric references. A similar viewpoint has
recently emerged from studies on adaptation-based duration
compression, showing that visual event timers may remain anchored
to retinal coordinates (Bruno et al. 2010) or may exhibit a genuine
spatial tuning in external space (Burr et al. 2011). In fact, adaptation
to high temporal frequency induces spatially specific reductions in
the apparent duration of sub-second intervals containing medium
frequency drift or flicker.

The neural bases of the described anisotropy may arise at the
processing stage at which the analysis of visual motion (direction
and speed) has been combined with the internal model of gravity.
A candidate network is given by a set of (possibly inter-connected)
cortical regions: the putative human homolog of middle-temporal
area (hMT+), parieto-insular vestibular cortex (PIVC), and posterior
parietal cortex (PPC). All three regions are involved in processing
visual motion and encoding time, but they play different roles vis-à-
vis gravitational stimuli. Thus, while neural populations in hMT+ can
feed direction and speed information for each motion direction into
downstream regions (Born and Bradley 2005), neural populations
of PIVC appear to be tuned preferentially to object motion related
to gravity. Indeed, fMRI studies (Indovina et al. 2005; Maffei et al.
2010; Miller et al. 2008) showed activation of hMT+ with object
motion either coherent with or violating gravity in a distant scene.
Instead, PIVC was activated selectively by motion coherent with
gravity. Conversely, trans-cranial magnetic stimulation (TMS) of
hMT+ affected the interception timing for both gravity-coherent
and gravity-incoherent target motion in the vertical or horizontal
direction, whereas TMS of TPJ affected only the interception
timing for vertical motion coherent with gravity (Bosco et al. 2008).
Internal time signals may arise directly in visual-motion regions
through the modulation of local horizontal connections, resulting
in trailing inhibition left behind by the moving object (Sundberg et

al. 2006). In monkeys, MT/MST feed visual information into PPC regions, such as lateral intra-parietal (LIP) area and 7a, which in turn are inter-connected with PIVC. Direct correlates of elapsed time in the sub-second range have been found in PPC, where populations of neurons exhibit ramping activities whose slope tightly correlates with the perceived duration in a time discrimination task (Leon and Shadlen 2003) or with motor response timing in an interception task (Merchant et al. 2004). The slope is probably shaped by spatio-temporal integration of excitatory and inhibitory inputs related to visual-motion, motor intention, and high-order contextual signals. We conjecture that neural attributes of the internal model of gravity fed by PIVC may affect both the slope and its temporal variability, thereby contributing to the internal time estimates.

CONCLUSIONS

We reviewed evidence for the existence of distributed, specialized mechanisms for time encoding in the sub-second range. Specialization could be related to brain mechanisms which simulate the behavior of different categories of objects by means of distinct models. One such specialization we considered is related to the animate-inanimate distinction. This distinction is a foundational one, because it arises early in infancy, is cross-culturally uniform, and is critical for causal interpretations of actions and events. The animate-inanimate or living-nonliving distinction hinges upon expected kinetic differences. Animate entities are endowed with internal energy sources which allow self-propelled motion. By contrast, inanimate entities are driven by external energy sources only and are incapable of self-propelled motion. In particular, people's expectations from daily life regarding how human beings move in the environment differ considerably from expectations regarding the motion of inanimate objects. Thus, animate targets in the visual field move according to biological-dynamic laws. Instead, the dynamics of inanimate targets simply obey Newton's laws. Expected kinetics may also account for the finding that time appears to run faster in an animate context than in an inanimate one (Carrozzo et al. 2010; Orgs and Haggard 2011). Indeed, in the ancestral world where action monitoring presumably evolved, animate targets tended to move more frequently than inanimate targets, and their behavior was more time-sensitive. Accordingly, changes in animate targets are detected faster than those in inanimate

targets (New et al. 2007).

Although time perception is not unitary, nevertheless there are some factors which can affect disparate time estimates in the same manner. In particular, we remarked that the entrainment of neural time basis by biological or non-biological (e.g., gravitational) cues holds true for both perceptual judgments of elapsed time (as in time discrimination tasks) and in automatic motor control tasks (as in the interception of a moving target). In the former case, time estimates are relevant to reconstruct (postdict) the temporal history of an event which occurred in the recent past. In the latter case, instead, time estimates are necessary to predict the occurrence of events in the near future. An organism that can predict accurately is able to plan a response to foreseen events, while one that cannot predict can only react after the event. Moreover, purely reactive behavior may be rendered useless by the sensori-motor delays; neural transmission, muscle force generation and effector inertia, each contribute to considerable time delays. Also, the reliance on kinetics (Georgopoulos 2002) is a factor which affects both perceptual and automatic time estimates.

REFERENCES

Abernethy B (1990) Expertise, visual search, and information pick-up in squash. Perception 19:63-77

Bahill AT, Baldwin DG, Venkateswaran J (2005) Predicting a baseball's path. Am Sci 93:218-225

Bidet-Ildei C, Orliaguet JP, Sokolov AN, Pavlova M (2006) Perception of elliptic biological motion. Perception 35:1137-1147

Borghese NA, Bianchi L, Lacquaniti F (1996) Kinematic determinants of human locomotion. J Physiol (Lond) 494:863-879

Born RT, Bradley DC (2005) Structure and function of visual area MT. Annu Rev Neurosci 28:157-189

Bosco G, Carrozzo M, Lacquaniti F (2008) Contributions of the human temporoparietal junction and MT/V5+ to the timing of interception revealed by transcranial magnetic stimulation. J Neurosci 28:12071-12084

Bove M, Tacchino A, Pelosin E, Moisello C, Abbruzzese G, Ghilardi MF (2009) Spontaneous movement tempo is influenced by observation of rhythmical actions. Brain Res Bull 80:122-127

Brown SW (1995) Time, change, and motion: the effects of stimulus movement on temporal perception. Percept Psychophys 57:105-116

Bruno A, Ayhan I, Johnston A (2010) Retinotopic adaptation-based visual duration compression. J Vis 10(10):30, 1-18

Bueti D, Baharami B, Walsh V (2008) Sensory and association cortex in time perception. J Cogn Neurosci 20:1054-1062

Buhusi CV, Meck WH (2005) What makes us tick? Functional and neural mechanisms of interval timing. Nat Rev Neurosci 6:755-765

Burr D, Tozzi A, Morrone MC (2007) Neural mechanisms for timing visual events are spatially selective in real-world coordinates. Nat Neurosci 10:423-425

Burr D, Cicchini GM, Arrighi R, Morrone MC (2011) Spatiotopic selectivity of adaptation-based compression of event duration. J Vis 11(2):21, 1-9

Carrozzo M, Moscatelli A, Lacquaniti F (2010) Tempo rubato: animacy speeds up time in the brain. PLoS One 5(12):e15638

Chang DH, Troje NF (2009) Acceleration carries the local inversion effect in biological motion perception. J Vis 9(1):19, 1-17

Chang DH, Harris LR, Troje NF (2010) Frames of reference for biological motion and face perception. J Vis 10(6):22, 1-11

Chatterjee SH, Freyd JJ, Shiffrar M (1996) Configural processing in the perception of apparent biological motion. J Exp Psychol Hum Percept Perform 22:916-929

Cook J, Saygin AP, Swain R, Blakemore SJ (2009) Reduced sensitivity to minimum-jerk biological motion in autism spectrum conditions. Neuropsychologia 47:3275-3278

Creelman CD (1962) Human discrimination of auditory duration. J Acoust Soc Am 34:528-593

Davidson, P R, Wolpert, D M (2005) Widespread access to predictive models in the motor system: a short review. J Neural Eng 2:S313-S319

Eagleman DM (2004) Time perception is distorted during slow motion sequences in movies. J Vis 4(8):491, 491a

Fraisse P (1963) The psychology of time. Harper and Row, New York, NY

Freire A, Lee K, Symons LA (2000) The face inversion effect as a deficit in the encoding of configural information: direct evidence. Perception 29:159-170

Freyd JJ (1983) The mental representation of movement when static stimuli are viewed. Percept Psychophys 33:575-581

Friedman WJ (2002) Arrows of time in infancy: the representation of temporal-causal invariances. Cogn Psychol 44:252-296

Georgopoulos AP (2002) Cognitive motor control: spatial and temporal aspects. Curr Opin Neurobiol 12:678-683

Gibbon J (1977) Scalar expectancy theory and Weber's law in animal timing. Psychol Rev 84:279-325

Grealy MA, Craig CM, Bourdin C, Coleman SG (2004) Judging time intervals using a model of perceptuo-motor control. J Cogn Neurosci 16:1185-1195

Haggard P, Clark S, Kalogeras J (2002) Voluntary action and conscious awareness. Nat Neurosci 5:382-385

Hammett ST, Champion RA, Morland AB, Thompson PG (2005) A ratio model of perceived speed in the human visual system. Proc Biol Sci 272:2351-2356

Hubbard TL (1995) Environmental invariants in the representation of motion: Implied dynamics and representational momentum, gravity, friction, and centripetal force. Psychonomic Bull Rev 12:822-851

Huber S, Krist H (2004) When is the ball going to hit the ground? Duration estimates, eye movements, and mental imagery of object motion. J Exp Psychol Hum Percept Perform 30:431-444

Huys R, Smeeton NJ, Hodges NJ, Beek PJ, Williams AM (2008) On the dynamic information underlying visual anticipation skill. Percept Psychophys 70:1217-1234

Indovina I, Maffei V, Bosco G, Zago M, Macaluso E, Lacquaniti F (2005) Representation of visual gravitational motion in the human vestibular cortex. Science 308:416-419

Jeannerod M (1994) The representing brain: neural correlates of motor intention and imagery. Beh Brain Sci 17:187-245

Jenkin HL, Jenkin MR, Dyde RT, Harris LR (2004) Shape-from-shading depends on visual, gravitational, and body-orientation cues. Perception 33:1453-1461

Johnston A, Arnold DH, Nishida S (2006) Spatially localized distortions of event time. Curr Biol 16:472-479

Kanai R, Paffen CL, Hogendoorn H, Verstraten FA (2006) Time dilation in dynamic visual display. J Vis 6:1421-1430

Kaneko S, Murakami I (2009) Perceived duration of visual motion increases with speed. J Vis 9:1-12

Kim IK, Spelke ES (1992) Infants' sensitivity to effects of gravity on visible object motion. J Exp Psychol Hum Percept Perform 18:385-393

Kourtzi Z, Kanwisher N (2000) Activation in human MT/MST by static images with implied motion. J Cogn Neurosci 12:48-55

Krekelberg B, Dannenberg S, Hoffmann KP, Bremmer F, Ross J (2003) Neural correlates of implied motion. Nature 424:674-677

Lacquaniti F, Terzuolo C, Viviani P (1983) The law relating the kinematic and figural aspects of drawing movements. Acta Psychol (Amst) 54:115-130

Lee DN (1998) Guiding movement by coupling taus. Ecol Psychol 10:221-250

Lhamon WT, Goldstone S (1974) Studies of auditory–visual differences in human time judgment: 2 More transmitted information with sounds than lights. Percept Mot Skill 39:295-307

Leon MI, Shadlen MN (2003) Representation of time by neurons in the posterior parietal cortex of the macaque. Neuron 38:317-327

Lewis PA, Miall RC (2003) Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. Curr Opin Neurobiol 13:250-255

Lorteije JA, Kenemans JL, Jellema T, van der Lubbe RH, de Heer F, van Wezel RJ (2006) Delayed response to animate implied motion in human motion processing areas. J Cogn Neurosci 18:158-168

Maffei V, Macaluso E, Indovina I, Orban G, Lacquaniti F (2010) Processing of targets in smooth or apparent motion along the vertical in the human brain: an fMRI study. J Neurophysiol 103:360-370

Mauk MD, Buonomano DV (2004) The neural basis of temporal processing. Annu Rev Neurosci 27:307-340

Merchant H, Battaglia-Mayer A, Georgopoulos AP (2004) Neural responses during interception of real and apparent circularly moving stimuli in motor cortex and area 7a. Cereb Cortex 14:314-331

Miller WL, Maffei V, Bosco G, Iosa M, Zago M, Macaluso E, Lacquaniti F (2008) Vestibular nuclei and cerebellum put visual gravitational motion in context. J Neurophysiol 99:1969-1982

Morgan MJ, Giora E, Solomon JA (2008) A single "stopwatch" for duration estimation, a single "ruler" for size. J Vis 8(2):14

Morrone MC, Ross J, Burr D (2005) Saccadic eye movements cause compression of time as well as space. Nat Neurosci 8:950-954

Moscatelli A, Lacquaniti F (2011) The weight of time: gravitational force enhances discrimination of visual motion duration. J Vis 11(4):5, 1-17

Moscatelli A, Polito L, Lacquaniti F (2011) Time perception of action photographs is more precise than that of still photographs. Exp Brain Res 210:25-32

Nather FC, Bueno JL, Bigand E, Droit-Volet S (2011) Time changes with the embodiment of another's body posture. PLoS One 6(5):e19818

Neri P, Luu JY, Levi DM (2006) Meaningful interactions can enhance visual discrimination of human agents. Nat Neurosci 9:1186-1192

New J, Cosmides L, Tooby J (2007) Category-specific attention for animals reflects ancestral priorities, not expertise. Proc Natl Acad Sci USA 104:16598-16603

Orgs G, Haggard P (2011) Temporal binding during apparent movement of the human body. Vis Cogn 19:833-845

Pariyadath V, Eagleman D (2007) The effect of predictability on subjective duration. PLoS One 2(11):e1264

Reed CL, Stone V, Bozova S, Tanaka J (2003) The body inversion effect. Psychol Sci 14:302-308

Rizzolatti G, Craighero L (2004) The mirror-neuron system. Annu Rev Neurosci 27:169-192

Sebanz N, Knoblich G (2009) Prediction in joint action: what, when, and where. Topics Cogn Sci 1:353-367

Shepard RN (1984) Ecological constraints on internal representations: Resonant kinematics of perceiving, imagining, thinking, and dreaming. Psychol Rev 91:417-447

Shim J, Carlton LG, Chow J W, Chae WS (2005) The use of anticipatory visual cues by highly skilled tennis players. J Motor Beh 37:164-175

Smith AT, Edgar GK (1994) Antagonistic comparison of temporal frequency filter outputs as a basis for speed perception. Vis Res 34:253-265

Soechting JF, Lacquaniti F (1981) Invariant characteristics of a pointing movement in man. J Neurosci 1:710-720

Sundberg KA, Fallah M, Reynolds JH (2006) A motion-dependent distortion of retinotopy in area V4. Neuron 49:447-457

Terao M, Watanabe J, Yagi A, Nishida S (2008) Reduction of stimulus visibility compresses apparent time intervals. Nat Neurosci 11:541-542

Thompson P (1980) Margaret Thatcher: a new illusion. Perception 9:483-484

Thompson P (1981) Velocity after-effects: the effects of adaptation to moving stimuli on the perception of subsequently seen moving stimuli. Vis Res 21:337-345

Treisman M (1963) Temporal discrimination and the indifference interval: implications for a model of the 'internal clock'. Psychol Monographs 77:1-31

Tse PU, Intriligator J, Rivest J, Cavanagh P (2004) Attention and the subjective expansion of time. Percept Psychophys 66:1171-1189

Viviani P, Figliozzi F, Lacquaniti F (2011) The perception of speech rate and of the arrow of time in visual speech. Exp Brain Res (in press)

Viviani P, Stucchi N (1992) Biological movements look uniform: evidence of motor-perceptual interactions. J Exp Psychol Hum Percept Perform 18:603-623

Watanabe K (2008) Behavioral speed contagion: automatic modulation of movement timing by observation of body movements. Cognition 106:1514-1524

Winawer J, Huk AC, Boroditsky L (2008) A motion aftereffect from still photographs depicting motion. Psychol Sci 19:276-283

Yarrow K, Haggard P, Heal R, Brown P, Rothwell JC (2001) Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. Nature 414:302-305

Yin RK (1969) Looking at upside-down faces. J Exp Psychol 81:141-145

Zago M, La Scaleia B, Miller WL, Lacquaniti F (2011) Observing human movements helps decoding environmental forces. Exp Brain Res, Sep 27 [Epub ahead of print]

Zago M, McIntyre J, Senot P, Lacquaniti F (2009) Visuo-motor coordination and internal models for object interception. Exp Brain Res 192:571-604