Research Focus

Sound and vision

Bart Krekelberg

Vision Center Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

When a brief flash appears at the same position as a moving object, the flash is perceived to lag behind. This so-called flash-lag effect tells us something about the perception of space and time: *where* **is the moving object** *when* **the flash appears? A recent paper by Alais and Burr on auditory and crossmodal flash-lag effects indicates that our (often implicit) models of the perception of space and time might be flawed.**

The flash-lag effect [1] is an illusion of spatial localization that has been studied extensively in the visual domain [2–4]. Alais and Burr have recently shown [5] that it also occurs for moving sounds and even when the location of a moving sound is compared with that of a visual flash (Fig. 1). Capitalizing on the very different properties of the auditory system, Alais and Burr's data reveal inconsistencies in existing models of the flash-lag effect.

Just another illusion?

Why has the flash-lag effect (FLE) received so much attention over the last few years? Is it not just another illusion, showing once more that what you see is not what you get? First, the phenomenon is unusually robust: it is seen not only in carefully controlled laboratory settings, but it can be demonstrated in a dark lecture hall with a strobe light and a continuously visible cigarette tip or, at a conference, by a researcher walking on the stage with a photo-flash. Second, the phenomenon is rich: it is found for visual motion, visual attribute changes (e.g. colour, contrast, entropy), and it is affected by self-motion signals (for recent reviews, see [2–4]). Adding to the richness and complexity, Alais and Burr have now shown that the FLE can be extended from vision to sound: brief tone bursts are perceived to lag behind a moving sound source or moving light source. The ubiquity of flash-lag effects suggests that this is not just another illusion, but instead the signature of a general mechanism that determines snapshots of continuously changing attributes of the sensory world.

Neural latencies

Models of the FLE have proliferated at nearly the same rate as experimental studies. The most popular view, however, remains that the neural latency of the flash is longer than that of the moving object. Given this difference in latency, the flash reaches awareness when the moving object is already farther along its trajectory [6]. This differential-latency view is shown in terms of a schematic neural network in Fig. 2.

Existing neurophysiological data are by no means conclusive on whether flashes or bursts have a longer neural latency than moving objects [4]. Alais and Burr's new data, however, provide an alternative way to assess the validity of the differential-latency model. They show that the FLE exists for sound and cross-modally between sound and vision. By extending the FLE to the auditory domain, which has very different neural processing properties from vision, new constraints on existing views of the FLE arise. Alais and Burr measured the classical FLE between a visual flash and visual motion (condition 'VV'; Fig. 1a), but also lag-effects between an auditory 'flash' and auditory motion ('AA'; Fig. 1b), and crossmodal effects: a visual flash with auditory motion ('VA'; Fig. 1c) and *vice versa* ('AV'; Fig. 1d). In all four cases, subjects reported that the flash lagged behind the motion. The size of the effects, however, was revealing; they are ordered as: AA > AV > VA > VV. If differential latencies are responsible for all these effects, then these data imply an ordering of latencies: $\lambda_{\text{Auditory-motion}} < \lambda_{\text{Visual-motion}} < \lambda_{\text{Visual-flash}} < \lambda_{\text{Auditory-flash}}$. Alais and Burr point out that it is highly improbable that the auditory system's latency for motion is the fastest when its latency for 'flashes' is up to 169 ms slower (from their data). Moreover, neural latencies in the auditory system are known to be much shorter than those in the visual system. This therefore raises the question: what is wrong with the simple and intuitively appealing differential-latency model?
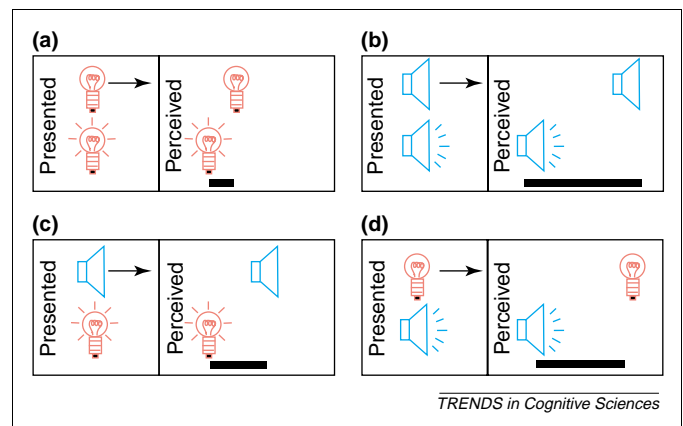
Fig. 1. The flash-lag effect (FLE). (a) When a moving object and a flashed object are shown at the same position at the same time, the flashed object appears to lag behind (condition 'VV', using Alais and Burr's terms). (b) Alais and Burr have now shown that this effect also occurs in audition (condition 'AA'), and is even larger than in vision [5]. (c,d) The FLE also occurs crossmodally. (c) condition 'VA', (d) condition 'AV'. Horizontal black bars show the relative size of the FLE in each of the four conditions. Arrows indicate which object is moving.

Corresponding author: Bart Krekelberg (bart@salk.edu).
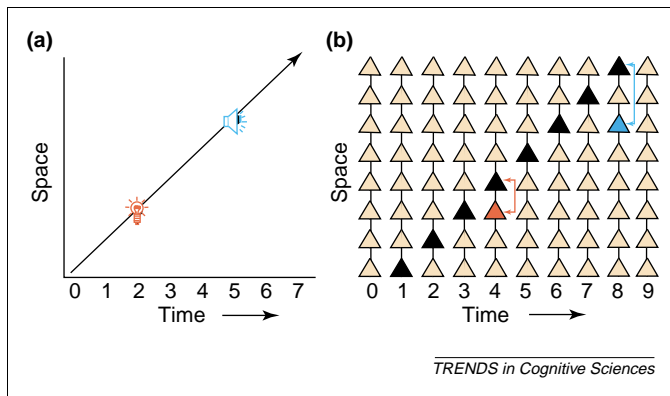
Fig. 2. An explanation of the flash-lag effect that is intuitive but probably wrong. (a) Space-time diagram of a typical flash-lag experiment. Black line = the moving object; red bulb = a flashed visual object; blue speaker = a 'flashed' auditory object. Time in arbitrary units. (b) A model that underlies much reasoning about the FLE. The columns represent a topographically organized neural network at successive times during the FLE experiment. Black triangles = activation caused by the moving object; red triangle = activation caused by the visual flash; blue triangle = activation caused by the auditory 'flash'. The assumptions made here are that neural activation by the moving object is delayed by 1 (arbitrary) time unit, activation by the visual flash by 2 units and activation by the auditory 'flash' by 3 units. By the time the neuron representing the visual flash becomes active (the red neuron at t = 4), another neuron in the topographic map already represents the position of the moving object (the black neuron at t = 4). The red double arrow indicates the mismatch in the neural representation and provides an explanation of the flash-lag effect. The blue arrow shows the FLE for an auditory flash, which is even larger because the (assumed) latency difference between the moving and 'flashed' object is larger.

How are space and time encoded in the brain?

Nobody would suggest that red neurons represent red objects. Nevertheless, an analogous view is often taken for the representation of space and time. In Fig. 2b, neurons represent position by means of their position in the network. Similarly, in Fig. 2, the time of perception is the time at which particular neurons become active. Hence, perceived time is represented by physical time. This is contrary to what is commonly assumed about the coding of other sensory attributes: the spikes of the neurons in Fig. 2 carry no information about space or time. If the brain codes space and time in this manner it must have a clever homunculus that knows not only which neuron in the topographic map represents a particular position in space but also when it fired. Assuming that such a clever homunculus exists leads to many interpretational problems (for a review, see [7]).

To answer the question of *where* the moving object is *when* the flash appeared, the brain must first determine the time of appearance of the flash and then use this information to determine the position of the moving object at that (perceptual) time. This calculation is likely to be complicated, and could be affected by events happening between the physical time of appearance and the physical time the answer is given. By analogy, consider brightness. Perceived brightness is affected by the current amount of light, the light in the spatial surroundings, and the light level that preceded the current moment. Given the ubiquity of lateral and feedback interactions between neurons, such spatio-temporal interactions should really not be all that surprising. Now consider the perceived time of

appearance as just another feature of the visual environment – a feature that has to be determined from the sensory input and that is not in principle different from brightness. In this view, the spatio-temporal surround of an object would be expected to influence its perceived time of appearance. Similarly, if position is viewed as just another feature of an object then the neural algorithm that determines it could well be influenced by positions of other objects in the spatio-temporal surround.

Space perception

Alais and Burr discuss in detail some of the existing flash-lag models that partially incorporate such ideas about the perception of position or time. For instance, the temporal averaging [8] and postdiction models [9] do not determine position as a simple snapshot of the environment, but as the result of a calculation using a stream of position input signals. The positional sampling model [10], on the other hand, stresses that the time when the moving object's position is perceived need not be simply related to the physical time of the flash. In this model, the flash starts a process that eventually samples the position of the moving object, but the sampling process itself takes an unknown amount of time.

Time perception

Recent studies of perceptual asynchrony have approached the time of perception from a different perspective. Moutoussis and Zeki showed that if an object simultaneously changes both its colour and its direction of motion, the colour change is perceived before the motion change [11]. In a model like that in Fig. 2, one would conclude that the latencies for colour perception must be shorter than those for motion perception. This, however, contradicts existing electrophysiological data [12] and such models of perceptual asynchrony run into the same interpretational problem as differential-latency models of the flash-lag effect. Nishida and Johnston recently suggested an interesting model in which the time of perception is not equated with any particular physical time of a neural event. In their time-marker theory, the brain actively assigns temporal markers to events. Perceptual asynchrony can then arise from an erroneous marker assignment or an error in the comparison of two markers, rather than from a difference in neural processing time [13].

Conclusion

In extending the scope of investigation for the flash-lag effect to the auditory domain, Alais and Burr provide new ways to investigate and constrain mechanisms of spatial perception. Additionally, however, their data seem to indicate that our assumptions about the perception of time are in need of revision. Such revision may be guided by a fruitful combination of the more spatially oriented research of the flash-lag effect with the temporal studies that investigate

perceptual asynchrony. Taken together such studies might be able to tell us more about the 'where' pathway and how it interacts with a (so-far entirely speculative) 'when' pathway.

References
1 Metzger, W. (1931) Versuch einer gemeinsamen Theorie der Phäno-mene Fröhlichs und Hazelhoffs und Kritik ihrer Verfahren zur Messung der Empfindungszeit. *Psychol. Forsch.* 16, 176–200
2 Whitney, D. (2002) The influence of visual motion on perceived position. *Trends Cogn. Sci.* 6, 211–216
3 Nijhawan, R. (2002) Neural delays, visual motion and the flash-lag effect. *Trends Cogn. Sci.* 6, 387–393
4 Krekelberg, B. and Lappe, M. (2001) Neuronal latencies and the position of moving objects. *Trends Neurosci.* 24, 335–339
5 Alais, D. and Burr, D. (2003) The 'flash-lag' effect occurs in audition and cross-modally. *Curr. Biol.* 13, 59–63
6 Whitney, D. and Murakami, I. (1998) Latency difference, not spatial extrapolation. *Nat. Neurosci.* 1, 656–657
7 van de Grind, W. (2002) Physical, neural, and mental timing. *Conscious. Cogn.* 11, 241–264.Discussion 308–313
8 Krekelberg, B. and Lappe, M. (2000) A Model of the perceived relative positions of moving objects based upon a slow averaging process. *Vis. Res.* 40, 201–215
9 Rao, R.P. *et al.* (2001) Optimal smoothing in visual motion perception. *Neural Comput.* 13, 1243–1253
10 Brenner, E. and Smeets, J.B.J. (2000) Motion extrapolation is not responsible for the flash-lag effect. *Vis. Res.* 40, 1645–1648
11 Moutoussis, K. and Zeki, S. (1997) A direct demonstration of perceptual asynchrony in vision. *Proc. R. Soc. Lond. B Biol. Sci.* 264, 393–399
12 Schmolesky, M.T. *et al.* (1998) Signal timing across the macaque visual system. *J. Neurophysiol.* 79, 3272–3278
13 Nishida, S. and Johnston, A. (2002) Marker correspondence, not processing latency, determines temporal binding of visual attributes. *Curr. Biol.* 12, 359–368

Number knows no bounds

Elizabeth M. Brannon

Psychological and Brain Sciences Center for Cognitive Neuroscience, Duke University, Box 90999, Durham, NC 7708-0999, USA

Is the approximate representation of large numbers in adult humans bound to a sensory modality or other aspects of the stimulus presentation, or are these non-verbal numerical representations abstract? A recent paper by Barth, Kanwisher and Spelke provides compelling evidence that non-verbal mental magnitudes are modality-independent.

Numerical systems are often considered a crowning achievement of modern-day human cultures. Numbers are the universal language, and mathematics an abstract system whose symbols are far removed from elementary sense data. However, the last few decades of research have shown us that, as would be predicted by a Darwinian perspective, there are evolutionary precursors to the human mathematical mind that do not presuppose language. Research with non-human animals has documented that animals represent number and can reason about numerical quantities [1,2]. Similarly, research with human infants suggests that before productive language, or in some cases even comprehensive language, has blossomed, infants are discriminating, ordering and perhaps even adding and subtracting numerical quantities [2–5].

Non-verbal number representations in adult humans
Non-verbal means of representing number are not limited to animals and preverbal infants. Adult humans also possess a non-verbal system for representing and reasoning about numerical quantities. Moyer and Land-aeuer [6] first showed that when adults make ordinal comparisons with Arabic numerals, their response time and accuracy are systematically related to the numerical distance between choices (distance effect) and the size of the two numbers being compared (size effect). The distance effect is in fact strikingly similar in animals and humans (see Fig. 1) [7]. That adults exhibit distance (and magnitude) effects regardless of whether the stimuli are presented as Arabic numerals or random-dot patterns, suggests that in both cases number is represented as mental magnitudes.
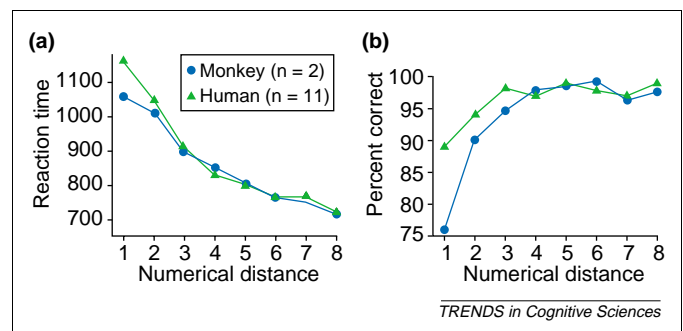

Fig. 1. The effect of numerical distance on accuracy and latency in monkeys and humans. (Monkey data, blue circles; human data, green triangles.) Reaction time (a) and accuracy (b) as a function of the linear distance between two visual arrays that ranged in value from 1–9. The distance effect is strikingly similar in the two species. Redrawn with permission from Ref. [7].

Corresponding author: Elizabeth M. Brannon (brannon@duke.edu).