

Perceived temporal asynchrony between sinusoidally modulated luminance and depth

Gojko Žarić

Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands



Arash Yazdanbakhsh

Department of Psychological & Brain Sciences, Center for Computational Neuroscience and Neural Technology, Boston University, Boston, MA, USA



Shigeaki Nishina

Honda Research Institute Japan Co., Ltd., Wako-Shi, Saitama, Japan



Peter De Weerd

Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands



Takeo Watanabe

Department of Cognitive, Linguistic and Psychological Sciences, Brown University, Providence, RI, USA



Simultaneously presented visual events lead to temporally asynchronous percepts. This has led some researchers to conclude that the asynchronous experience is a manifestation of differences in neural processing time for different visual attributes. Others, however, have suggested that the asynchronous experience is due to differences in temporal markers for changes of different visual attributes. Here, two sets of bars were presented, one to each eye. Either the bars were moving or their luminance was gradually changing. Bars moved horizontally in counterphase at low frequencies along short trajectories and were presented stereoscopically, such that the horizontal movements were perceived as back-and-forth motion on a sagittal plane, or monocularly to a dominant eye, preserving a perception of the horizontal movements on a frontal plane. In a control condition, bars were stationary and their luminance was modulated. The changes in stimulus speed or luminance occurred sinusoidally. When asked to adjust the phase of one stimulus to the other to achieve synchronous perception, participants showed a constant phase offset at the lowest frequencies used. Given the absence of abrupt transitions and the presence of similar gradual turning points in our stimuli to control for attentional effects, it can be concluded that asynchronous percepts in multimodal stimuli may at least in part be a manifestation of difference in neural

processing time of visual attributes rather than solely a difference in the temporal markers (transitions versus turning points).

Introduction

From the moment light hits the retina until we become aware of the image, on the order of 50–150 ms may elapse (Lamme & Roelfsema, 2000; Nowak, Munk, Girard, & Bullier, 1995; Vanni et al., 2004). This is due to the fact that the signal travels from the receptors in our eyes through subcortical nuclei and cortical areas, triggering interactions between feed-forward and feedback processes (Tootell, Hadjikhani, Mendola, Marrett, & Dale, 1998). Furthermore, the stimuli in our natural environment are typically defined by multiple cues. A number of studies have reported that simultaneous visual events defined in different stimulus domains can lead to temporally asynchronous percepts (Harris, Duke, & Kopinska, 2006; Ishii, Seekkuarachchi, Tamura, & Tang, 2004; MacKay, 1958; Moutoussis & Zeki, 1997; Nijhawan, 1994; Nishida & Johnston, 2002). These observations likely reflect a difference in the visual neural circuitry

Citation: Žarić, G., Yazdanbakhsh, A., Nishina, S., De Weerd, P., & Watanabe, T. (2015). Perceived temporal synchrony between sinusoidally modulated luminance and depth. *Journal of Vision*, 15(15):13, 1–11, doi:10.1167/15.15.13.

involved in processing different visual cues (Roelfsema, Tolboom, & Khayat, 2007).

The asynchrony effect was first observed by MacKay (1958). He placed his participants in a stroboscopically lit room and asked them to look at a self-luminous object and intermittently press their eyeballs. They reported a discrepancy between the self-luminous object, which seemed to be moving, and the rest of the room, which stayed still. In another experiment, Nijhawan (1994) coined the term “flash-lag effect” to describe the spatial lagging of the flashing end segments of a rotating line whose central part was illuminated constantly. Another group (Ishii et al., 2004) found a 3-D flash-lag effect using a mirror stereoscope to simulate two thin white sticks approaching the subject. During the approach of the two moving sticks, a third white stick was flashed between them. The researchers found that the flash stimulus appeared to be behind the moving stimuli. Harris et al. (2006) stereoscopically presented luminance-defined discs, dynamic random-dot stereograms, and random-dot stereograms to manipulate disparity (difference in the image location of an object seen by the left and right eyes), looming (rapid expansion in size), and lateral motion. Participants had to judge the position of the 3-D moving stimuli relative to a reference stimulus with a fixed position at the moment of a flash (reference stimulus changed luminance for ~30 ms). The perceived 3-D moving stimuli were perceived ahead of their actual location as defined by their relative position to the reference stimulus at the time of the flash. The size of the effect increased when a disagreement between monocular cues (looming and lateral motion) and a disparity increased, and when the stimuli were perceived as moving towards the viewer. Similar to the flash-lag studies, Moutoussis and Zeki (1997) used color change instead of flashes and asked their participants to pair the color of the pattern (red or green) to its direction of vertical motion (upwards or downwards). The squares changed both color and motion direction following a square-wave function. The results showed that the color change is perceived earlier than the motion-direction change.

Two hypotheses have been proposed to account for the asynchrony effect. The first states that asynchronous perception of simultaneously presented stimuli is the result of different processing times for different visual attributes (Moutoussis & Zeki, 1997). The alternative hypothesis is that the asynchronous experience stems from different subjective marker assignments which depend on the temporal structure of stimulus presentation, including transitions and turning points as the sources of asynchrony (Nishida & Johnston, 2002). Nishida and Johnston define temporal markers as “temporally localized representations of salient temporal features” (2002, p. 360). In this view,

the asynchronous perception in these and similar experiments would be a consequence of different temporal markers assigned to the compared stimulus changes (Arnold & Wilcock, 2007; Harris et al., 2006; Ishii et al., 2004; López-Moliner & Linares, 2006; Moutoussis & Zeki, 1997; Nijhawan, 1994; Nishida & Johnston, 2002). To test the importance of such markers, Nishida and Johnston (2002) defined transitions as the first-order temporal changes (abrupt changes requiring measurements at only two points in time) and turning points as the second-order temporal changes (reversal of a gradual change, requiring measurements at three points in time). Two stimulus parameters were used—color and direction motion of a plaid pattern; they either changed both in transitions, both in turning points, or separately in transitions and turning points, yielding a 2×2 design. Motion perception was delayed when the motion change was a turning point and the color change a transition, and conversely, color perception was delayed when the color change was a turning point and the motion change a transition. The effect was found in a frequency range of 0.5–2 Hz, with the strongest effect around 2 Hz and smaller effects below that frequency. Importantly, this delay dropped significantly when both changes were of the same kind (transition or turning point), suggesting that the asynchrony effect can be explained by different temporal structures rather than a difference in the visual attributes. These findings suggest that independent of the stimulus parameter used, the presence of a turning point in a parameter seems to delay the processing of that parameter. Furthermore, the residual asynchrony when the temporal markers were matched was attributed to attentional gating.

In the present study, we investigated whether a smooth temporal profile would prevent asynchronous perception of the two cues used or whether asynchronous perception would remain despite the absence of abrupt stimulus transitions and similar turning points. Importantly, we presented two complementary conditions to control for attentional gating—i.e., bars whose luminance or position/depth changed over time were presented in the center of the vision. The participants saw bars either with two eyes and through a haploscope/stereoscope (an optical device for presenting one image to one eye and another image to the other eye, allowing fusion and 3-D perception) or with only one eye (monocularly) and no stereopsis. We used a smooth (sinusoidal) temporal profile for the modulation of two visual stimulus attributes (luminance and position). The participants were instructed to judge “the point of nearest (farthest) depth/position and luminance” of the stimulus. Thus, they judged position (not motion). Importantly, measurements at three points in time are required to perform this task,

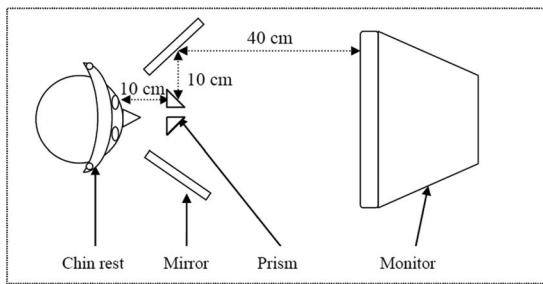


Figure 1. Experimental setting.

regardless of stimulus (depth/position or luminance), and the temporal profiles of the turning points for luminance, position, and depth are the same and again sinusoidal. Our experiments extended previous findings by showing that two stimulus parameters undergoing smooth sinusoidal changes at the same frequency and at matched phases can nevertheless be perceived with a delay relative to each other even when attentional gating is ruled out. We therefore suggest that the resulting perceived phase offsets or delays in our experiments, measured in the absence of attentional advantage and abrupt stimulus transitions and in the presence of similar turning points, may offer a better estimate of neural processing differences between different stimulus parameters than are obtained in typical flash-lag experiments that include abrupt stimulus changes.

Methods

Participants and experimental setting

Eight adults (two women and six men; age: $M[SD] = 28.23[2.52]$ years; six right-eye dominant) participated in the experiment.

The experiment consisted of four sessions (two for the binocular conditions and two for the monocular). Left-eye and right-eye stimuli were independently viewed through a haploscope by the respective eyes (Figure 1). For the binocular conditions, stereoscopic vision was assessed by an independent static image consisting of five circles: four in the angles of the imaginary equilateral diamond and the fifth in the vicinity of the intersection of the diagonals. When viewed through the haploscope together with its mirror image, the middle circle would seem to be closer to the participant. Participants were asked to describe the figure, and if they used any term that would describe the middle circle as closer to them (e.g., “popping out,” “closer,” “coming out of the screen”), this was taken as a sign of intact 3-D vision. For the monocular condition, there was a black screen in front of one of

the prisms, and thus participants could see only one image with the uncovered eye. The uncovered eye was the dominant eye as measured by the “hole-in-card” test (Seijas et al., 2007). The test card was downloaded from <http://www.usaeyes.org/lasik/library/Dominant-Eye-Test.pdf>. A chin rest was used to maintain the position of the head and the distance from the computer and haploscope. The complete viewing distance was 60 cm: 40 from the monitor to the mirrors of the haploscope, 10 from the mirrors to the prism of the haploscope, and 10 from the prism to the eye.

Stimuli

The stimuli were custom-made, and the stimulus presentations and randomizations were all programmed in MATLAB R2009a (MathWorks, Natick, MA) using the Psychophysics Toolbox Version 3 extension (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007). Stimuli were presented on a Vision Master Pro 514 computer screen (20.4-in. diagonal screen size, 100-Hz refresh rate, 1024×768 resolution; Iiyama, Hoofddorp, the Netherlands) that was connected to a laptop (Satellite L500-128, Toshiba, Neuss, Germany).

For the binocular conditions, each stimulus pair consisted of left and right single stimuli presented on the left and the right side of the screen and viewed through a haploscope independently by the left and right eyes. This enabled disparity-induced 3-D depth perception. In the monocular condition, a black screen covered one prism and mirror between the participant and the haploscope; thus only one member of the stimulus pair was presented (to the dominant eye). In this way, the physical input to the viewing retina was the same as in the binocular condition.

A single member of a stimulus pair is illustrated in Figure 2. In each member, a small fixation point (red square) was located above a small central square, itself flanked by two additional small squares. All stimuli were presented against a gray background (25.9 cd/m^2) (Figure 2a). Squares were $14'$, and the distance between the lateral squares and the central one was 2° . The fixation point was $48'$ above the central square (Figure 2b). We refer to square stimuli as bars. The small central and lateral bars were used to present either a luminance modulation or motion (Figure 2c through e). In the case of luminance modulation, the bar changed from 1 to 138.3 cd/m^2 . In the case of motion, the amplitude of the horizontal displacement was $19'$ to each side—in other words, $38'$ between the extreme left and right points, equivalent to $38'$ crossed ($+38'$) and uncrossed ($-38'$) disparity between the extreme near and far depths in stereo viewing. In stereo-viewing conditions, a Dc condition was tested (with central motion in depth, within $\pm 38'$ crossed and uncrossed disparities, flanked

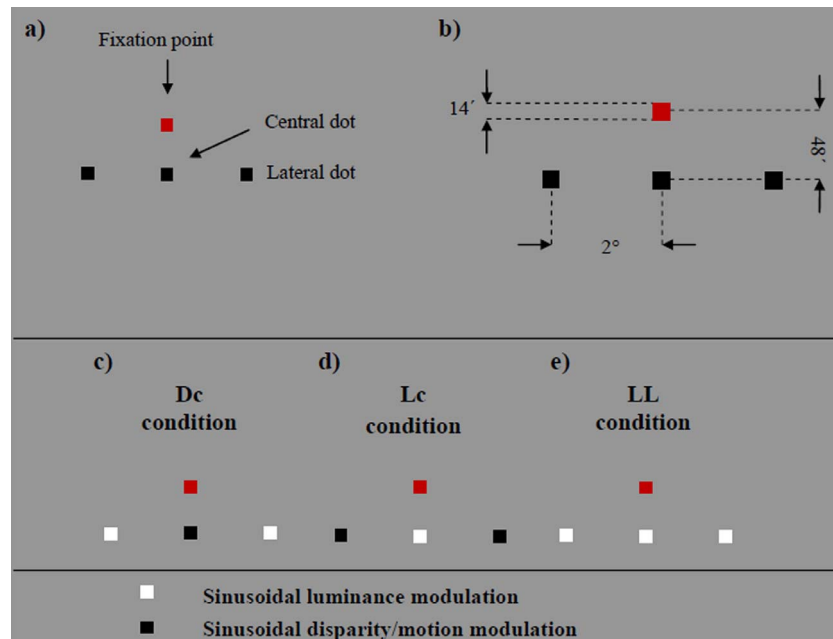


Figure 2. Stimuli single sets and experimental conditions. (a) Locations of the fixation point and the central and lateral bars. (b) Size of the bars and their distance from each other. (c–e) Three experimental conditions with different luminance and disparity modulation of the bars: White bars represent the luminance modulation and black bars represent disparity modulations. See text for further details of stimuli presentation.

by luminance-modulating stimuli), as were as an Lc condition (with central luminance modulation flanked by stimuli modulated in depth within $\pm 38'$ crossed and uncrossed disparities) and an LL condition (all three bars below fixation were used to present a luminance modulation). In the monocular condition, we presented all three conditions (Dc, Lc, LL) again, with the difference that in this case a perception of the horizontal movements on a frontal plane was preserved (though there was no movement in depth, for simplicity we kept the Dc and Lc labels for central motion flanked by luminance-modulating stimuli). The geometry and dynamics of our stimuli translate to the minimum zero velocity for displacement and disparity at extreme locations, and the maximum velocities of $2^\circ/s$, $4^\circ/s$, and $6^\circ/s$ for displacement and $4^\circ/s$, $8^\circ/s$, and $12^\circ/s$ for disparity at 1, 2, and 3 Hz, respectively.

Dc condition

In the Dc condition, the left and right pair-member central bars were moved sinusoidally to the left and right (Figure 2c). The central-bar position for the left eye was determined by

$$c_L = A \sin(2\pi ft + \varphi) \quad (1)$$

and for the right eye by

$$c_R = -A \sin(2\pi ft + \varphi) \quad (2)$$

where A represents the amplitude of the maximum displacement from the center of oscillation, which equaled $19'$ visual angle. The modulation frequencies f were 1, 2, and 3 Hz, and φ represents the phase difference of central-bar displacement relative to the phase of the luminance modulations of the lateral bars. At stimulus onset, φ was initially drawn from a uniform random distribution from 0 to 2π .

The sinusoidal horizontal motion of the central bar in each member of a pair occurred in mirror symmetry, thereby generating a sinusoidal disparity modulation, which in turn induced a perceived sinusoidal motion in depth (hence the abbreviation Dc). In Figure 2c and d, the bars showing just sideways motion are rendered in black (1 cd/m^2).

In the Dc condition, the lateral bars were modulated in luminance with the same frequency as was used for motion in depth in the central location. Luminance for the left and right bars L_1 was modulated as follows:

$$L_1 = L_{\min} + \frac{(L_{\max} - L_{\min})(1 + \sin(2\pi ft))}{2} \quad (3)$$

where L_{\min} is the minimum luminance (1.0 cd/m^2) and L_{\max} the maximum luminance (138.3 cd/m^2). The luminance modulation was gamma-corrected. If φ is zero, this means that the maximum luminance and maximum crossed disparity of the central bars coincide. Matching phase for these two physical stimulus measures does not imply that the percept of the two

modulations is perceived as simultaneous. To match perception, participants were asked to vary the relative phase of φ in either direction by pressing one key or the other until they achieved a synchronized depth percept of the depth modulation in the central bar with respect to the luminance of the lateral ones.

Lc condition

In the Lc condition (Figure 2d), the lateral bars were perceived to have depth modulation and the central bar to have luminance modulation (hence the abbreviation Lc). The lateral bars were moving sinusoidally sideways to modulate horizontal disparity. The lateral-bar positions of the left- and right-eye stimuli from their central points as a function of time are given by

$$x_L = A \sin(2\pi ft) \quad (4)$$

for the left eye and

$$x_R = -A \sin(2\pi ft) \quad (5)$$

for the right eye.

The central-bar luminance modulation for the left- and right-eye stimuli is given by

$$L_c = L_{\min} + \frac{(L_{\max} - L_{\min}) \left(1 + \sin(2\pi ft + \varphi)\right)}{2} \quad (6)$$

LL condition

In the LL condition (luminance center/luminance periphery), both lateral and central bars showed a luminance modulation (Figure 2e). The central-bar luminance modulation followed Equation 6, and the lateral-bar luminance modulation followed Equation 3.

Note that in all conditions the adjustable phase φ belongs to the central bar, irrespective of whether it showed disparity (Dc condition) or luminance (Lc and LL conditions) modulations. Furthermore, stimulus modulations in the lateral bars were always identical and in phase.

Participants' task

Each participant performed 21 trials in a block design for each of the three stimulus conditions (Dc, Lc, LL) and at each of the three frequencies. In addition, the stereo and monocular conditions were run, yielding a $3 \times 3 \times 2$ experimental design. The

sequence of conditions was pseudorandomized, to avoid having the same condition or frequency in two consecutive blocks, and counterbalanced across the participants.

To make the instruction of synchrony as concrete as possible, participants were asked in the Dc and Lc conditions to change the initially randomized phase φ of the central stimulus back and forth until they perceived the point of nearest/farthest (minimal/maximal) perceived depth and the point of maximal/minimal perceived luminance to coincide in time. In the monocular condition, synchrony was defined as a perceived concurrence in time of the point of most lateral position relative to the center of motion and the point of maximal perceived luminance. To keep the condition similar to the binocular ones, participants who viewed stimuli with the right eye had to adjust the most leftward lateral position with luminance, and participants who viewed with the left eye, most rightward position. In the LL condition, the participants had to adjust the phase φ of the luminance modulation in the central bar until they perceived the extreme levels of the central and lateral bars to coincide in time. The set phase φ for each condition was recorded when the participant pressed a key to indicate perceptual synchrony. The task was not time limited.

Statistical analysis

Data were analyzed using MATLAB and SPSS 20. For the analysis we included only data between $-\pi/2$ and $\pi/2$ radians (corresponding to time delays of ± 500 ms for 1 Hz, ± 250 ms for 2 Hz, and ± 166 ms for 3 Hz; see the nongray part of the histogram in Figure 3), as the larger offsets could represent synchronization with the counterphase at higher frequencies. This filtering step was supported by the finding that there was a higher percentage of trials with φ responses outside the $[-\pi/2, \pi/2]$ range in the 3- and 2-Hz conditions than in the 1-Hz condition. In these conditions, φ reports showed a dip around $-\pi/2$ and $\pi/2$ radians and a rise towards the counterphase (Figure 3). Anecdotal reports from the participants also support the presence of counterphase errors. For the statistical analysis, we multiplied results for the Lc condition by -1 (negLc) to obtain the differences only from the effect sizes and not from the expected opposite delays in the Dc and Lc conditions. We used a three-way repeated-measures ANOVA. Although the number of trials differed across the conditions, due to the filtering step, this should not pose a problem for linear models such as ANOVA, as the trials within the chosen interval still represent a random sample. If sphericity was violated (Mauchly's

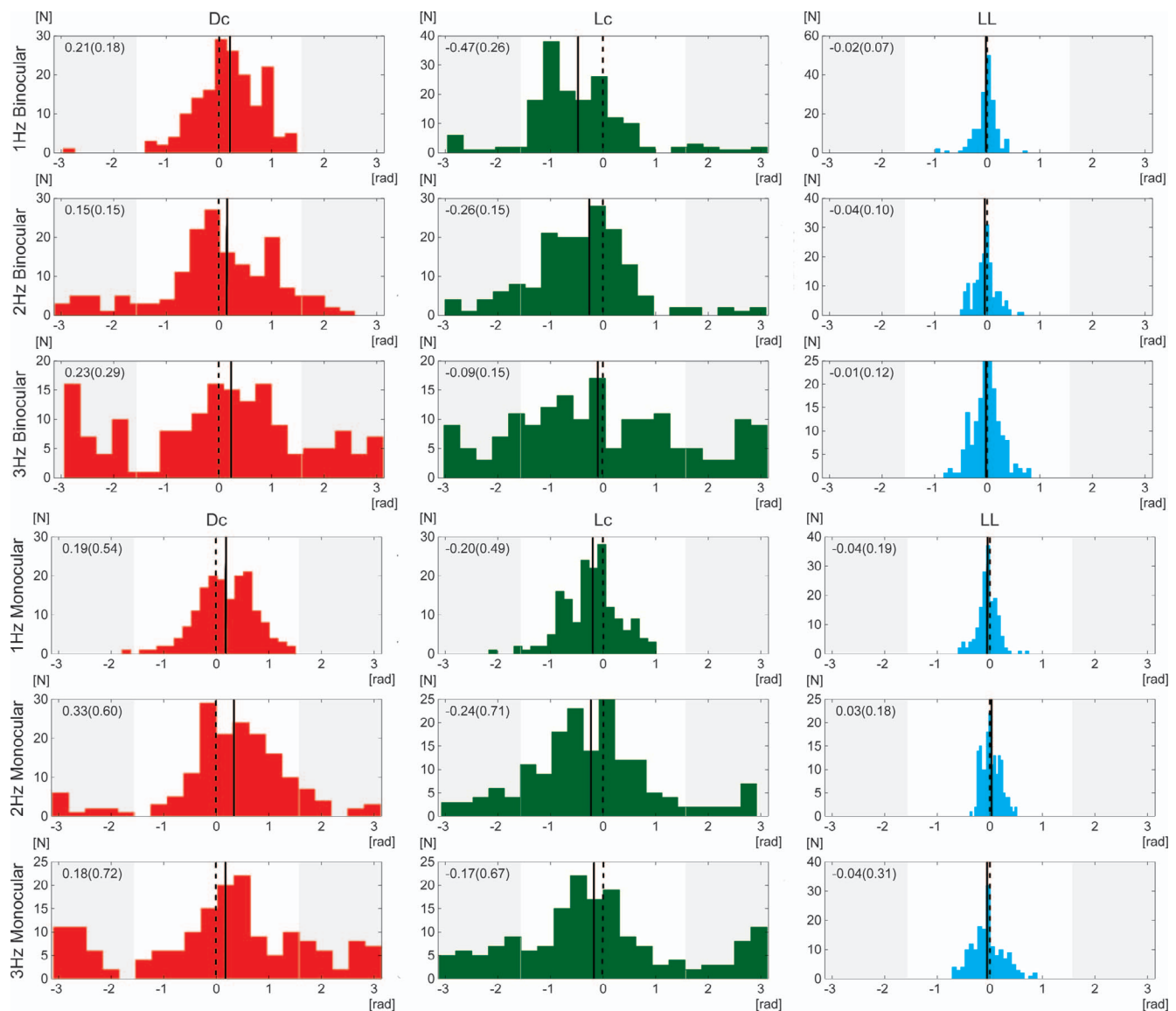


Figure 3. Group histograms: red = Dc, green = Lc, blue = LL. Gray areas show excluded trials outside the $[-\pi/2, \pi/2]$ interval. The vertical black line represents the mean of the trials in the $[-\pi/2, \pi/2]$ interval. The mean (standard deviation) of the trials between $-\pi/2$ and $\pi/2$ is given in the upper left corner of each histogram.

test), we reported Greenhouse–Geisser-corrected values. The within-subject factors were trial type (Dc, negLc, LL), frequency (1, 2, 3 Hz), and ocularity (binocular, monocular). Next to significance values (p), we report partial eta-squared (η^2) for effect sizes. Follow-up analysis included one-sample t -test statistics for each condition with reported one-tailed significances, as we expected Dc delay to be positive, Lc delay to be negative, and LL delay to be nonsignificant. The multiple-comparisons problem was controlled by false-discovery-rate (FDR) correction based on a uniform distribution of p values under the null hypothesis (Benjamini & Hochberg, 1995).

Results

As mentioned in Methods, for the analysis we included data between $-\pi/2$ and $\pi/2$ radians to diminish the effects of erroneous counterphase synchronization. As can be seen in Figure 3 (right column), the data histogram in the LL condition fell fully into the $[-\pi/2, \pi/2]$ range, and hence all trials could be used (see white portion of Figure 3). The same is true for the vast majority of trials in the Dc and Lc conditions at 1 and 2 Hz, which fell in the $[-\pi/2, \pi/2]$ range of the histograms. Even at 3 Hz, almost two thirds of trials occurred in the $[-\pi/2, \pi/2]$ range. Table 1 shows the

	Binocular			Monocular		
	1 Hz	2 Hz	3 Hz	1 Hz	2 Hz	3 Hz
Dc	1	29	67	1	28	61
Lc	21	30	62	3	34	59
LL	0	0	0	0	0	0

Table 1. Number of excluded trials. Notes: The excluded data represented 13% of the original total number of trials.

numbers of trials in all conditions excluded after data filtering.

A repeated-measures ANOVA was performed with three within-subject factors: trial type (Dc, negLc, LL), frequency (1, 2, 3 Hz), and ocularity (binocular, monocular). This analysis yielded a significant main effect of trial type, $F(2, 14) = 8.40, p = 0.004, \eta^2 = 0.546$, and a marginally significant main effect of frequency, $F(2, 14) = 3.49, p = 0.059, \eta^2 = 0.333$, as well as a significant interaction between frequency and ocularity, $F(2, 14) = 4.69, p = 0.028, \eta^2 = 0.401$. Other effects were not significant. Thus, there was an effect of trial type on synchronicity perception, which depended on ocularity.

A follow-up repeated-measures ANOVA showed no differences between Dc and negLc over frequency and ocularity, with the smallest p value of .073 for the Frequency \times Ocularity interaction and all other $ps \geq 0.104$. Conversely, comparison of Dc and LL trial types yielded only a main effect of trial type, $F(1, 7) = 17.99, p = 0.004, \eta^2 = 0.720$, with all other effects yielding $ps > 0.105$. Comparison between negLc and LL yielded main effects of trial type, $F(1, 7) = 18.85, p = 0.003, \eta^2 = 0.729$, and frequency, $F(1, 7) = 8.61, p = 0.004, \eta^2 = 0.552$, as well as interactions between trial type and frequency, $F(2, 14) = 3.79, p = 0.048, \eta^2 = 0.352$, and frequency and ocularity, $F(1, 7) = 5.79, p = 0.015, \eta^2 = 0.453$.

As we obtained a Frequency \times Ocularity interaction, we compared perceptually adjusted phases for the Lc trial type in binocular and monocular conditions at each frequency separately. Confirming the visual

		1 Hz		2 Hz		3 Hz	
		$t(7)$	p	$t(7)$	p	$t(7)$	p
Binocular	Dc	3.235	0.007	2.899	0.012	2.188	0.033
	Lc	-5.205	0.000	-4.907	0.001	-1.760	0.061
	LL	-0.899	0.199	-1.248	0.126	-0.355	0.367
Monocular	Dc	2.079	0.038	3.302	0.007	1.576	0.080
	Lc	-2.344	0.026	-2.664	0.016	-1.181	0.138
	LL	-1.650	0.072	1.382	0.105	-1.414	0.100

Table 2. Tests of phase adjustments by participants compared to zero (one-sample t tests). Notes: Italics indicate tests with one-tailed $p < 0.05$. Bold indicates tests that are significant after FDR correction, $q(\text{FDR}) = 0.016$.

inspection of mean offsets for the Lc in Figure 4, we obtained a significant difference only at 1 Hz, $t(7) = -3.181, p = 0.015$, with no difference at other frequencies ($p > 0.876$).

Thus, our results show the differences in effect sizes between the control (LL) and each of the experimental trial types but not between the two experimental trial types (Lc and Dc). Furthermore, only analysis including the Lc trial type yielded significant differences between the binocular and monocular conditions, and this was due to the Lc trial type at 1 Hz, with the offset in the binocular condition being atypically large compared to the other offsets.

In testing for which trial types had a phase adjustment significantly different from zero (Table 2; Figure 4), we found that the phase in the Dc and Lc trial types differed from zero at 1 and 2 Hz for both binocular and monocular conditions, but only Dc differed at 3 Hz, and only for disparity.

To obtain the temporal difference (Δt) in milliseconds between the central and lateral bars in subjective phase adjustment, we used the following relation:

$$\Delta t = \frac{1000 * \varphi}{(2\pi f)} \text{ [ms]} \tag{7}$$

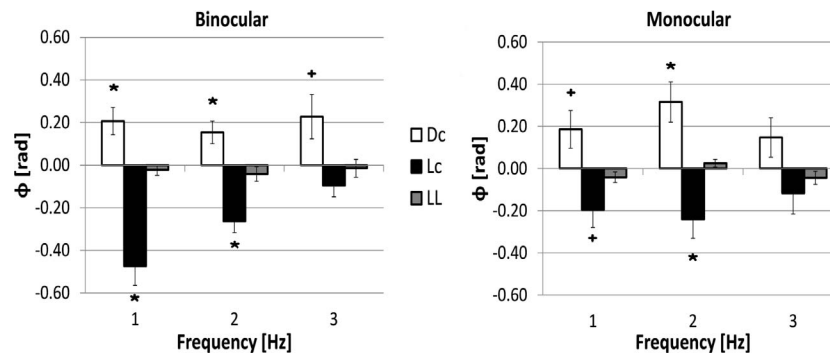


Figure 4. Central-bar adjusted peak phase difference. Conditions that are significant after FDR correction are marked *; significant uncorrected p values that do not survive correction are marked +.

	Binocular			Monocular		
	1 Hz	2 Hz	3 Hz	1 Hz	2 Hz	3 Hz
Dc	33 ± 29	12 ± 12	12 ± 16	30 ± 40	25 ± 22	8 ± 14
Lc	−75 ± 41	−21 ± 12	−5 ± 8	−31 ± 38	−19 ± 20	−6 ± 15
LL	−4 ± 11	−3 ± 8	−1 ± 6	−7 ± 11	2 ± 4	−2 ± 5

Table 3. Central-bar adjusted peak temporal difference (Δt , in milliseconds) relative to the lateral bars in perceptual synchrony. In both the Dc and Lc conditions, motion phase led luminance phase (i.e., luminance changes were processed faster). Since participants always adjusted the phase of the central bar, this led to positive time offsets in the Dc condition and negative offsets in the Lc condition.

Table 3 summarizes the temporal differences for all three conditions. Positive and negative values of Δt correspond to positive and negative values of φ , respectively. The positive temporal difference in the Dc condition means that participants on average set the perceived motion-change time ahead of the perceived luminance-change time to achieve perceptual synchrony, as if the percept of motion change lagged behind the percept of luminance change. By contrast, in the Lc condition the participants on average set the luminance-change time after the motion-change time, as if the percept of luminance change led the percept of motion change. Overall, this means that luminance is processed faster than motion (in both binocular/stereo and monocular conditions).

Discussion

We found that in temporal modulation of motion and luminance, participants set the phase of motion modulation ahead of that of the luminance modulation to achieve perceptual synchrony. This occurred irrespective of whether the motion stimulus was monocular on a frontal plane or 3-D on a sagittal plane. The data suggest that luminance processing occurs faster than the processing of both monocular and stereo motion, and hence that luminance is perceived ahead of motion in stimuli showing matched oscillatory modulations. This advantage of luminance was limited to frequencies of 1–2 $c/^\circ$, and disappeared at higher frequencies.

Can we interpret this result as being due to a difference in neural processing time between motion processing (Bayerl & Neumann, 2004; Pack, Born, & Livingstone, 2003; Roe, Parker, Born, & DeAngelis, 2007; von der Heydt, Zhou, & Friedman, 2000) and luminance processing? In a study by Moutoussis and Zeki (1997), moving squares (up and down) and isoluminant colors (green and red) were perceived synchronously only when the motion-direction change preceded the color change—as if the perception of motion-direction change lagged behind the perception of the color change. This finding is similar to ours. Moutoussis and Zeki suggested that the perceived time

discrepancy between synchronous visual events results from a difference in neural processing for each visual attribute.

However, Johnston and Nishida (Johnston & Nishida, 2001; Nishida & Johnston, 2002) found that the perceived delay of changes in motion direction relative to physically synchronous color changes depended on the temporal structure of the stimuli. In particular, their findings depended upon the presence of abrupt temporal changes in color versus smooth temporal changes in motion direction. They concluded that the perceptual asynchrony was due to the assignment of different subjective markers—namely transitions versus turning points—in the stimuli rather than to different neural processing times of stimulus attributes such as luminance and motion. They attributed residual delays that were still present when the temporal markers were equalized to attentional gating rather than processing delays (Nishida & Johnston, 2002).

To test the importance of the temporal structure of cyclical stimulus changes, we introduced sinusoidal temporal profiles for all parameter changes in our experiments. This eliminated the presence of any sharp transitions in our stimuli that could have been used as distinctive first-order temporal markers, and kept the temporal profile of turning points perfectly the same to avoid any confounds with regard to second-order temporal markers. Furthermore, we introduced two complementary configurations (Dc and Lc) to control for attentional gating (Reeves & Sperling, 1986). Nevertheless, we found a phase offset when disparity/horizontal motion and luminance were set to allow a percept of synchronous modulation. The control condition (LL) excluded the possibility that this phase offset was due to differences in visual processing pertaining to differences in retinal positions, namely, between the central and lateral bars. Most importantly, the two complementary configurations (Dc and Lc) in essence yielded the same results. These results render unlikely an explanation in terms of an attentional advantage (Carrasco & McElree, 2001; Hillyard, Vogel, & Luck, 1998) for the cue attended in the task. Participants had to adjust the phase of the central stimulus parameter, and they thus attended to motion

in the Dc conditions and to luminance in the Lc condition. According to an attentional-gating hypothesis, the centrally attended stimulus would always have an advantage, which is incompatible with the systematic advantage that we found for luminance processing.

The processing advantage for luminance and the associated difference in phase settings between luminance and motion could not be observed for frequencies higher than 2 Hz. This resembles previous findings demonstrating that when visual attributes modulate with frequencies above 2 Hz, it is hard to accurately judge which changed before the other when they are presented at a different locations (Holcombe & Cavanagh, 2001). Similarly, for the 3-Hz condition in our study, multiple participants reported difficulties in maintaining the unified percept in the binocular conditions, as well as uncertainty whether the synchronization was made in phase or counterphase. This can be seen also from the histograms for the 2- and 3-Hz conditions (Figure 3) that have a W shape, indicating that there are also peaks in the counterphase region. Although previous research has demonstrated difficulties in accurate judgments above 2 Hz (Holcombe & Cavanagh, 2001), the fact that the participants had no trouble perceiving changes in individual visual attributes (depth/position/luminance) over time at 3 Hz motivated us to include this condition. However, while the participants had no problem perceiving individual visual attributes, our results confirm that the perceptual-synchrony task for 3 Hz was difficult, due to uncertainty whether the synchronization was made in phase or counterphase. One possible explanation for this difficulty is that during stereoscopic viewing, the task is more challenging in the binocular conditions because of the geometry of the stimulus, in which disparity (and its rate of change) at each moment is twice the displacement (and its rate of change; see the Stimuli subsection in Methods).

Measuring the reaction time for the point of maximal (minimal) perceived depth/position and luminance is difficult and less practical, because there is no abrupt change in position or luminance. Moreover, the range of reaction times measured by Nishida and Johnston (2002) is on the order of ~ 100 ms, and in our experiments the range of asynchrony is on the order of ~ 10 ms; hence there is an order of magnitude's difference. Therefore, measuring reaction time (with the mentioned difficulty of a lack of abrupt change), while it is an order of magnitude larger than the targeted temporal offset, probably results in a sample of collected reaction times whose standard deviation is comparable with the ~ 10 -ms range.

Opposite to our luminance lead, flash-lag experiments (as described in the Introduction) have demonstrated a lag of luminance compared to motion (Harris et al., 2006; Ishii et al., 2004). In flash-lag experiments,

the participant's estimation of motion at the moment of a flash leads the physical position in depth (i.e., as defined by luminance). Interestingly, in one of those studies, a condition similar to our Lc condition was used, with flash lagging behind the motion (Harris et al., 2006). The fact that they used flash (impulse temporal structure) and we used sinusoidal luminance modulation could explain the difference in the direction of the lag. Thus, subjective markers such as those provided by abrupt transitions in a stimulus-parameter modulation—i.e., flashes as in flash-lag experiments—can have a dominant effect on temporal perception (Nishida & Johnston, 2002). In the case of abrupt transitions in stimulus-parameter modulations, the advantage may originate from attentional influences; in the case of flash-lag experiments, predictive coding may contribute (Berry, Brivanlou, Jordan, & Meister, 1999; Jancke, Erlhagen, Schöner, & Dinse, 2004; Nijhawan & Wu, 2009). Still, the possibility that both neural processing and the occurrence of temporal markers could contribute to the perception temporal asynchrony, even without abrupt transitions, cannot be ruled out.

In the present study, we observed phase offsets in the percept of slowly modulated visual parameters in different locations of the visual field. Compared to the square temporal profile for color and position that Nishida and Johnston used (2002, figure 6A [panel C1P1]), or the triangular temporal profile where the transition is sharp (similar to the triangular temporal profile of Nishida & Johnston, 2002, figure 6A, panels C1P2 for position, C2P1 for color, and C2P2 for both position and color), in the current study the temporal profile of luminance, position, and depth is sinusoidal and thus gradual instead of abrupt. In other words, the transition is neither square-shaped nor triangular but rather rounded (following a sinusoidal temporal pattern). We suggest that the introduction of the sinusoidal modulation rather than a flash, square, or triangular wave temporal profile is a practical setup to compare neural processing time and control for attentional effects. One limitation of this approach is a need for averaging over multiple trials, as unavoidable internal and external noise can induce variability in the single responses in the absence of an abrupt cue.

One possible explanation for the reported difference between processing time of luminance and position/depth change could be that perception of position/depth change during the change of a luminance-defined bar necessitates integration of luminance within the range of horizontal positions over time—and in particular, that the receptive-field size of neurons in early visual areas needed for localizing or detecting disparity of the bar is limited (Hubel & Wiesel, 1977). This could pose more of a processing

burden, because of the needed integration across the neurons whose limited receptive fields fall within the horizontal range of displacement, resulting in the position/depth lagging behind the luminance. Thus, it can be hypothesized that the two stimuli would have different temporal integration times. Luminance change would have faster temporal integration, as it falls into same receptive fields and is thus processed by the same neurons, while position change would have slower temporal integration, as it requires information from different neurons. Conversely, in the flash-lag experiments temporal integration would only be possible for the moving stimuli, as the information persists over time, while the flash persists only for a short time and is biased by its last-seen position (Krekelberg & Lappe, 2001).

In terms of the implications of these results for studying human vision, the method could be extended to any pair of visual modalities. Although this would have to be validated by neurophysiological experiments, it is possible that the experimental design used may be a tool in such experiments for teasing apart the temporal dynamics in the processing of different stimulus attributes.

Moreover, in practice, the results have implications for more optimized design of visual signals, particularly when temporal judgments are needed, such as traffic signs, airplane cockpits, and runway signals for pilots during landing.

Keywords: asynchronous perception, disparity, neural processing time, temporal markers, sinusoidal modulation

Acknowledgments

This work was supported in part by the Center of Excellence for Learning in Education, Science and Technology (NSF SBE-0354378), the National Institutes of Health/National Eye Institute (R01EY019466), the Office of Naval Research (N00014-11-1-0535), the Air Force Office of Scientific Research (FA9550-12-1-0436), and the Honda Research Institute, Japan. We would like to thank Nozomi Ito for all the administrative support during the progress and final stages of the research.

*GŽ and AY contributed equally to this article. Commercial relationships: none.

Corresponding author: Arash Yazdanbakhsh.

Email: yazdan@bu.edu.

Address: Department of Psychological & Brain Sciences, Center for Computational Neuroscience and Neural Technology, Boston University, Boston, MA, USA.

References

- Arnold, D. H., & Wilcock, P. (2007). Cortical processing and perceived timing. *Proceedings of the Royal Society B*, 274(1623), 2331–2336, doi:10.1098/rspb.2007.0576.
- Bayerl, P., & Neumann, H. (2004). Unified mechanisms in the disambiguation and grouping of visual information in motion, stereo, and monocular depth perception [Abstract]. *Perception*, 33(ECVP Abstract Supplement), 79.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300. Retrieved from <http://www.jstor.org/stable/10.2307/2346101>
- Berry, M. J., Brivanlou, I. H., Jordan, T., & Meister, M. (1999, March 25). Anticipation of moving stimuli by the retina. *Nature*, 398, 334–338, doi:10.1038/18678.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436.
- Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proceedings of the National Academy of Sciences, USA*, 98(9), 5363–5367, doi:10.1073/pnas.081074098.
- Harris, L. R., Duke, P., & Kopinska, A. (2006). Flash lag in depth. *Vision Research*, 46(17), 2735–2742, doi:10.1016/j.visres.2006.01.001.
- Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society B*, 353, 1257–1270, doi:10.1098/rstb.1998.0281.
- Holcombe, A. O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nature Neuroscience*, 4(2), 127–128, doi:10.1038/83945.
- Hubel, D. H., & Wiesel, T. N. (1977). Ferrier lecture: Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society B*, 198(1130), 1–59, doi:10.1098/rspb.1977.0085.
- Ishii, M., Seekkuarachchi, H., Tamura, H., & Tang, Z. (2004). 3D flash lag illusion. *Vision Research*, 44(17), 1981–1984, doi:10.1016/j.visres.2004.03.026.
- Jancke, D., Erlhagen, W., Schöner, G., & Dinse, H. R. (2004). Shorter latencies for motion trajectories than for flashes in population responses of cat

- primary visual cortex. *The Journal of Physiology*, 556, 971–982, doi:10.1113/jphysiol.2003.058941.
- Johnston, A., & Nishida, S. (2001). Time perception: Brain time or event time? *Current Biology*, 11(11), R427–R430, doi:10.1016/S0960-9822(01)00252-4.
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, 36, ECVF Abstract Supplement.
- Krekelberg, B., & Lappe, M. (2001). Neuronal latencies and the position of moving objects. *Trends in Neurosciences*, 24(6), 335–339, doi:10.1016/S0166-2236(00)01795-1.
- Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, 23, 571–579, doi:10.1016/S0166-2236(00)01657-X.
- López-Moliner, J., & Linares, D. (2006). The flash-lag effect is reduced when the flash is perceived as a sensory consequence of our action. *Vision Research*, 46(13), 2122–2129, doi:10.1016/j.visres.2005.11.016.
- MacKay, D. M. (1958, February 15). Perceptual stability of a stroboscopically lit visual field containing self-luminous objects. *Nature*, 181(4607), 507–508, doi:10.1038/181507a0.
- Moutoussis, K., & Zeki, S. (1997). A direct demonstration of perceptual asynchrony in vision. *Proceedings of the Royal Society B*, 264(1380), 393–399, doi:10.1098/rspb.1997.0056.
- Nijhawan, R. (1994, July 28). Motion extrapolation in catching. *Nature*, 370(6487), 256–257, doi:10.1038/370256b0.
- Nijhawan, R., & Wu, S. (2009). Compensating time delays with neural predictions: Are predictions sensory or motor? *Philosophical Transactions of the Royal Society A*, 367, 1063–1078, doi:10.1098/rsta.2008.0270.
- Nishida, S., & Johnston, A. (2002). Marker correspondence, not processing latency, determines temporal binding of visual attributes. *Current Biology*, 12(5), 359–368, doi:10.1016/S0960-9822(02)00698-X.
- Nowak, L. G., Munk, M. H., Girard, P., & Bullier, J. (1995). Visual latencies in areas V1 and V2 of the macaque monkey. *Visual Neuroscience*, 12, 371–384, doi:10.1017/S095252380000804X.
- Pack, C. C., Born, R. T., & Livingstone, M. S. (2003). Two-dimensional substructure of stereo and motion interactions in macaque visual cortex. *Neuron*, 37(3), 525–535, doi:10.1016/S0896-6273(02)01187-X.
- Reeves, A., & Sperling, G. (1986). Attention gating in short-term visual memory. *Psychological Review*, 93(2), 180–206, doi:10.1037//0033-295X.93.2.180.
- Roe, A. W., Parker, A. J., Born, R. T., & DeAngelis, G. C. (2007). Disparity channels in early vision. *The Journal of Neuroscience*, 27(44), 11820–11831, doi:10.1523/JNEUROSCI.4164-07.2007.
- Roelfsema, P. R., Tolboom, M., & Khayat, P. S. (2007). Different processing phases for features, figures, and selective attention in the primary visual cortex. *Neuron*, 56, 785–792, doi:10.1016/j.neuron.2007.10.006.
- Seijas, O., Gómez de Liaño, P., Gómez de Liaño, R., Roberts, C. J., Piedrahita, E., & Diaz, E. (2007). Ocular dominance diagnosis and its influence in monovision. *American Journal of Ophthalmology*, 144(2), 209–216, doi:10.1016/j.ajo.2007.03.053.
- Tootell, R. B., Hadjikhani, N. K., Mendola, J. D., Marrett, S., & Dale, A. M. (1998). From retinotopy to recognition: fMRI in human visual cortex. *Trends in Cognitive Sciences*, 2(5), 174–183, doi:10.1016/S1364-6613(98)01171-1.
- Vanni, S., Warnking, J., Dojat, M., Delon-Martin, C., Bullier, J., & Segebarth, C. (2004). Sequence of pattern onset responses in the human visual areas: An fMRI constrained VEP source analysis. *NeuroImage*, 21, 801–817, doi:10.1016/j.neuroimage.2003.10.047.
- von der Heydt, R., Zhou, H., & Friedman, H. S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Research*, 40(15), 1955–1967, doi:10.1016/S0042-6989(00)00044-4.