NEUROSCIENCE of Consciousness

Neuroscience of Consciousness, 2024, **2024(1)**, 1–14

DOI: https://doi.org/10.1093/nc/niad027

Research Article

Audiovisual interactions outside of visual awareness during motion adaptation

Minsun Park[1], Randolph Blake[2,*], Chai-Youn Kim[1,*]

[1]School of Psychology, Korea University, 145, Anam-ro, Seongbuk-gu, Seoul 02841, Republic of Korea
[2]Department of Psychology, Vanderbilt University, PMB 407817 2301 Vanderbilt Place, Nashville, TN 37240-7817, United States

*Corresponding author. Department of Psychology, Vanderbilt University, PMB 407817 2301 Vanderbilt Place, Nashville, TN 37240-7817, USA. E-mail: randolph.blake@vanderbilt.edu; School of Psychology, Korea University, 145, Anam-ro, Seongbuk-gu, 02841, Republic of Korea. E-mail: chaikim@korea.ac.kr.

Abstract

Motion aftereffects (MAEs), illusory motion experienced in a direction opposed to real motion experienced during prior adaptation, have been used to assess audiovisual interactions. In a previous study from our laboratory, we demonstrated that a congruent direction of auditory motion presented concurrently with visual motion during adaptation strengthened the consequent visual MAE, compared to when auditory motion was incongruent in direction. Those judgments of MAE strength, however, could have been influenced by expectations or response bias from mere knowledge of the state of audiovisual congruity during adaptation. To prevent such knowledge, we now employed continuous flash suppression to render visual motion perceptually invisible during adaptation, ensuring that observers were completely unaware of visual adapting motion and only aware of the motion direction of the sound they were hearing. We found a small but statistically significant congruence effect of sound on adaptation strength produced by invisible adaptation motion. After considering alternative explanations for this finding, we conclude that auditory motion can impact the strength of visual processing produced by translational visual motion even when that motion transpires outside of awareness.

Keywords: audiovisual interactions; multisensory integration; visual awareness; visual motion aftereffect; motion adaptation; continuous flash suppression

Introduction

Our everyday perceptual experiences of objects and events originate from energic information harvested by our various sensory modalities and transduced into neural signals by our sensory nervous systems. Amazing as these sensory events are, however, the information embodied in neurosensory signals remains largely implicit and requires interpretative processing by the brain before emerging as explicit perceptual concomitants of their real-world origins. To put it in other words, neural activity within the earliest stages of perceptual processing is ambiguous and incomplete, thus underspecifying what in the world gave rise to them. Detailed, adaptive perception requires reformatting sensory information into neural representations shaped by contextual factors, past experience, and current needs. Implementing these higher-order processes is the brain's job. One useful means for deriving information about what is in the world is to combine information derived from multiple sensory modalities. This realization has spawned the burgeoning field of research on multisensory integration (Macaluso and Driver 2005, Angelaki et al. 2009, Sathian and Ramachandran 2019, Wallace et al. 2020, Dwyer et al. 2022).

Within this research field, the most widely studied form of multisensory integration focuses on the melding of auditory and visual information. Those studies have demonstrated robust audiovisual (AV) interactions in perception (Spence and Sathian 2020), ranging from tasks tapping into plausibly low-level sensory processing to tasks involving high-level semantic judgments. To give some examples, a simple auditory tone can heighten the perceived intensity of a concurrently presented weak light flash (Stein et al. 1996, Chen et al. 2011b), and the perceived direction of auditory motion can impact the perceived direction of visual motion when the AV motion signals are concurrent in time (Lewis and Noppeney 2010, Rosemann et al. 2017) and space (Sekuler et al. 1997, Sadaghiani et al. 2009, Hidaka et al. 2011). By the same token, "semantically" congruent AV events make it easier to understand speech perception (Alsius and Munhall 2013, Plass et al. 2014) and to perform complex visual tasks such as discriminating different tap dancing sequences portrayed by point-light animations (Arrighi et al. 2009) and to form reliable associations between visual shape representations and associated auditory utterances (Heyman et al. 2019).

One phenomenon that has been utilized to investigate low-level neural interactions between hearing and seeing is the visual motion aftereffect (MAE). The MAE is illusory motion experienced in a direction opposite to real "adapting" motion viewed

Received 11 August 2023; Revised 5 December 2023; Accepted 27 December 2023

immediately prior to adaptation (Wohlgemuth 1911). This compelling illusion is one of vision science's most well-studied phenomena (Wade 1994, Anstis et al. 1998), and its perceptual character varies depending on the global configuration of the visual motion experienced during adaptation (Mather et al. 2008). Neural models of the MAE have been developed based on the notion of temporary shifts in patterns of activity within populations of neurons differing in their preferred directions of motion (e.g. Mather 1980), i.e. neurons of the sort identified within areas comprising early stages of the visual hierarchy (Petersen et al. 1985, Tootell et al. 1995, Huk et al. 2001).

So, how can the MAE be utilized to assess AV interactions? One way is to ask whether the visual MAE can be induced by prior listening (i.e. adapting) to sound that appears to be moving in a given direction over time. A few studies have tested this possibility, but the results are mixed. Some studies report that listening to sound perceived to be moving subsequently induces a visual MAE (Hedger et al. 2013, Berger and Ehrsson 2016), while others failed to find that sound can induce the MAE (Jain et al. 2008). Strictly speaking, however, those studies do not address the question of AV interaction, because sound and vision were not presented at the same time.

A more direct test is to ask whether sound presented concurrently with visual motion influences the strength of the visual MAE, and this is what our group did in a study reported a few years ago (Park et al. 2019). We found that auditory motion concurrently presented with visual motion during adaptation lengthened and strengthened the consequent visual MAE if the two bisensory events were congruent with respect to the direction of motion. The MAE was significantly weaker, however, when the sound direction was incongruent with the direction of the visual adaptation motion or when visual adaptation was not accompanied by sound at all. In the experimental design of that study, the critical interactions of AV motion signals occurred during adaptation periods, with the test periods involving visual stimulation only. We confirmed that eye movements did not differ among the three adaptation conditions, but we were left with the possibility that post-adaptation judgments of MAE strength might have been influenced by expectations or response bias associated with the knowledge of whether sound and vision experienced "during" adaptation were congruent or not. Decisions formed in that way would not require reliance on actual melding of visual and auditory-evoked neural activity and that realization led us to design and execute the experiment described in this paper.

By way of preview, we created adaptation conditions where observers were completely unaware of the direction of visual adapting motion and only knew the direction of motion of the sound they were hearing. This rendered it impossible for perceptual awareness of the congruence vs. incongruence between sound and vision to bias performance on the MAE measurement task. We rendered visual motion perceptually invisible during adaptation using continuous flash suppression (CFS), a potent procedure whereby a salient, dynamic visual masking stimulus presented to one eye can produce prolonged suppression of awareness of a more benign visual stimulus presented to the other eye (Fang and He 2005, Tsuchiya and Koch 2005). By using this potent technique for blocking a normally visible stimulus from awareness, we were able to measure the extent to which invisible monocular motion could be potentiated by auditory motion dependent on the congruence between the directions of motion of the auditory and visual stimuli. From earlier work, we knew that a visual stimulus, even when rendered invisible through interocular suppression, can still generate visual aftereffects including the MAE (Lehmkuhle and Fox 1975, Blake et al. 2006, Maruya et al. 2008, Kaunitz et al. 2011). In a similar vein, human brain imaging studies reveal the existence of evoked neural activity in the striate cortex in response to the presentation of visual motion stimuli that are suppressed from awareness (Yuval-Greenberg and Heeger 2013). Considered together, these lines of evidence imply that a visual stimulus suppressed from awareness can still evoke neural activity within visual areas of the brain where motion information is registered.

Based on the results described earlier, we foresaw a way to measure post-adaptation MAE strength under conditions where observers would not be influenced by cognitive knowledge about the congruence of auditory and visual motion. At a broader level, we surmised that results from this experiment might have an important bearing on the question of the conditions under which AV integration can transpire outside of awareness (Chen et al. 2011a, Alsius and Munhall 2013, Faivre et al. 2014, Cox and Hong 2015, Noel et al. 2015).

Materials and Methods

Participants

To determine an appropriate sample size to estimate possible differences in MAE durations dependent on AV conditions, we conducted an a priori power analysis using G*power 3.1 (Faul et al. 2009). We evaluated the effect sizes from two previous studies, one being our previous study with similar procedures except for the visibility of the motion during adaptation (Park et al. 2019) and the other being a study of AV semantic congruency outside visual awareness (Cox and Hong 2015). Power analyses were performed based on the reported effect size (η_p^2) from these studies for within-factor repeated measures analysis of variance (effect size F = ranged between 0.275 and 0.384, alpha = 0.05, power = 0.8). The estimated sample size was proposed between 13 and 23, and we selected a sample size of 23 for the main experiment for sufficient datasets in the case of exclusion during and after data collection. As for Pilot Experiments 1 and 2, different groups of five and seven participants took part, respectively. All observers had normal or corrected-to-normal vision and no auditory deficit. All observers were volunteers recruited from Korea University and received monetary compensation for participation. This study was approved by the Institutional Review Board of Korea University [IRB 1040548-KU-IRB-07-174-A-2(E-A-1)]. All participants completed the written informed consent form before the experiment.

Apparatus

Visual stimuli were displayed on a gamma-corrected cathode ray tube monitor set to display 1024 × 768 pixels at a 100 Hz frame rate and viewed at a distance of 60 cm. Auditory stimuli were delivered via headphones. All auditory and visual stimuli were created and presented using MATLAB and the Psychophysics Toolbox-3 (Brainard 1997, Kleiner et al. 2007). Participants viewed the visual stimuli through a mirror stereoscope that presented left- and right-eye displays on the two halves of the monitor. The observer's head was stabilized by a head/chin rest, and testing was carried out in a quiet, dark room.

Stimuli

Auditory white noise (44.1 kHz sampling rate) was generated and modified using commercially available sound creation/editing software (Sound Studio, Felt Tip Inc). The compelling sense

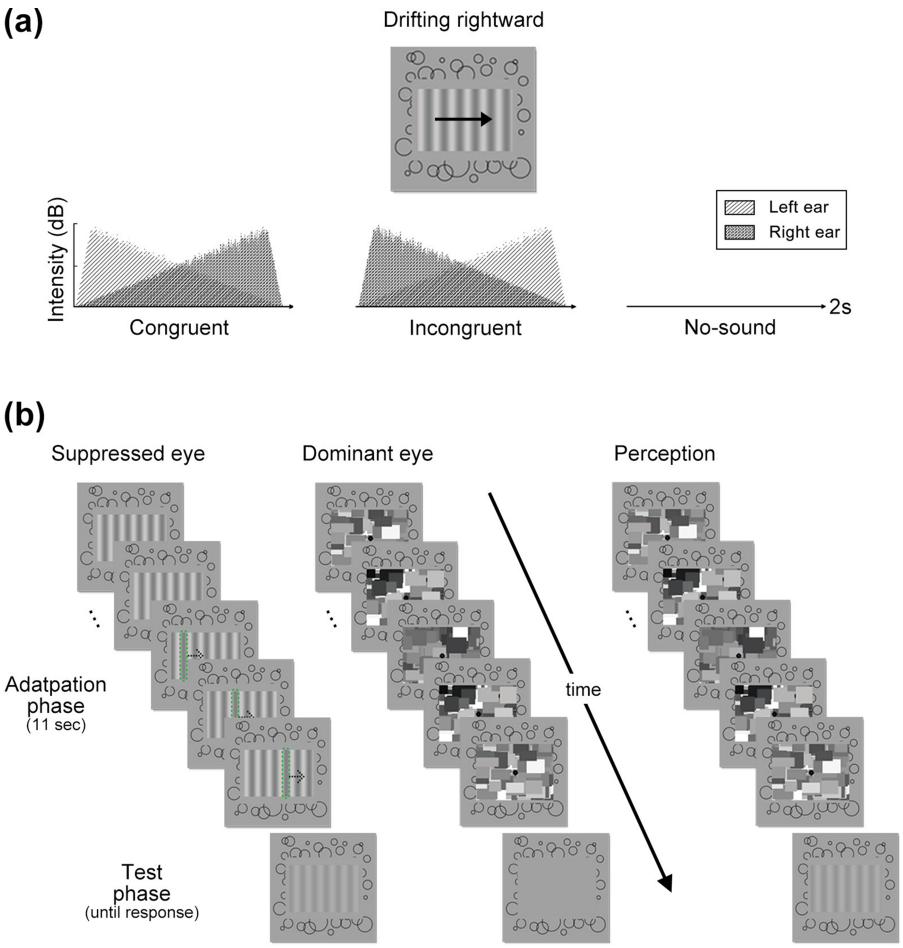

Figure 1. Experimental stimuli, conditions, and procedures in the main experiment. (a) Schematic of the visual and auditory stimuli used during the adaptation phase. The visual stimulus was presented in the suppressed eye. The direction of the adaptation grating was either leftward or rightward (rightward in this example). The auditory motion was generated by simulating the interaural intensity difference between white noise presented to the left and the right channels. According to the direction congruency of AV stimuli during the adaptation phase, there were CON, INC, and NS conditions. (b) In a trial, the 11-s adaptation phase preceded a test phase. During the adaptation phase, each eye viewed a vertical grating moving in the left or right direction (rightward in this example) and the CFS patterns. According to the AV conditions, the directional sounds were accompanied or not accompanied. During the adaptation phase, observers are perceptually aware of the CFS patterns and the directional sounds not the moving grating. During the test phase, the CFS patterns disappeared, and a static vertical grating of lower contrast was presented. The test phase was maintained until observers responded to the direction and the duration of induced MAE.

of auditory motion either leftward or rightward was created by crossfading the intensity of 2-s bursts of white noise which were heard over binaural headphones (Fig. 1a). The intensity of sounds at the beginning and the end of each noise burst was reduced over a 150-ms period, to eliminate abrupt transients at the onset and offset of the 2-s noise burst. A 500-ms quiet interval occurred between each 2-s noise burst, to allow a reset of the unidirectional motion being presented on a given trial.

Visual stimuli were grayscale, sinusoidal gratings presented on a uniform, gray background (28 cd/m^2). Gratings were vertically oriented, 2-cycle/deg, and they were used during the adaptation and test phases of each trial (Fig. 1a). The gratings appeared within a square window subtending a 3° × 2° visual angle. During 11 s of adaptation, the contours of the adaptation grating drifted continuously to the left or to the right accompanied by repeated presentations of the 2-s white noise pulses sandwiched between the 500-ms reset intervals. To promote the perceptual grouping of the discrete sound motions and the continuous visual motions, we modulated the speed of the drifting grating motion; one spatial phase of the grating was shifted in every video frame to produce the appearance of smooth motion to the left or to the right at a steady speed of 30 arcmin/s whenever the 2 s of sound motion was happening. During the 500-ms reset interval between the sound presentations, the drift speed of the grating slowed to 15 arcmin/s. The contrast of the adaptation grating was fixed at 23%, a value derived from the contrast-response function obtained in Pilot Experiment 1 (see "Pilot Experiment 1" in the Procedures section). The test grating was identical in orientation and spatial frequency to the adaptation grating, but the test grating was lower in contrast (i.e. 16%) to promote a more conspicuous MAE (Keck et al. 1976, Nishida et al. 1997).

The CFS stimuli comprised a series of randomly created grayscale Mondrian-like patterns with a central fixation point presented within a rectangular aperture subtending 3.5° × 2.5° of visual angle. Each Mondrian-like pattern was filled with rectangles drawn in variable luminance, location, and size (from 0.2° to 1.2° in length) within the aperture. A total of 1100 grayscale Mondrian-like patterns were generated in advance and randomly updated every 10 Hz during the adaptation phase. The display of the CFS was normalized in mean luminance matching the luminance

of the background. Root mean square contrast of the CFS was normalized at 93%.

To promote stable binocular alignment of eyes during periods of dichoptic stimulation, the stimuli viewed separately by the two eyes (grating and CFS) were presented within bubble-shaped fusion frames (Fig. 1b). The angular dimensions of the bubble-shaped fusion frames were 6° × 6°, the outline circles portraying the bubbles differed in size. With appropriate alignment of the mirrors together with careful positioning of the fusion frames within the two halves of the video monitor (described later), stable binocular alignment was easily maintained.

Procedures

To determine a person's dominant eye, we administered the Miles test to each participant, requiring them to report which eye was seeing a distal object through a small window formed by overlapping the hands of their extended arms. Prior to an experiment, the following procedure was used to align the dichoptic images on the video screen. The fusion frame viewed by one eye was presented continuously, while the fusion frame viewed by the other eye appeared and disappeared once every 2.2 s. Participants adjusted the horizontal and vertical positions of the intermittently seen image using arrow keys on the computer keyboard, until achieving a position where just a single frame was reliably seen within the center of the field of view.

In the main experiment, each trial began by displaying the current trial number, prompting the participant to press the space bar to begin the sequence of events defining a trial. This keypress triggered the 15-s binocular presentation of identical, dynamic noise patterns to each eye (dichoptic presentation). Each noise pattern comprised a 9 × 9 grid of cells arrayed within a square frame 4.6° on a side; the grayscale luminance of each cell varied within the grid and varied over time (25 Hz), with the constraint that the mean luminance remained constant over space and time (28 cd/m^2). These dynamic noise arrays were viewed for 15 s, the purpose being to eliminate any residual MAE or afterimages from the previous trial. Following the visual noise phase, a vertically oriented grating of 23% contrast level appeared within the fusion frame viewed by the eye to be adapted on this trial. At this time, participants assessed whether they experienced any hint of residual, illusory motion of the stationary grating and, if not, they were free to press the space bar to initiate the adaptation phase of the trial sequence. At the beginning of the visual motion adaptation phase, the adaptation grating presented to the suppressed eye drifted either leftward or rightward (Fig. 1b). During the 11-s adaptation period of some trials, 2-s episodes of sound were presented four times with the 500-ms inter-sound interval; on other trials, no sound was presented at all. The speed of the drifting adaptation grating decreased at the same time as the offset of the sound and resumed its original speed at the onset of sound motion. This periodic modulation of the speed of the visual adaptation motion also occurred during the no-sound (NS) condition. The dominant sighting eye viewed the dynamic Mondrian patterns (i.e. CFS stimuli) that appeared centered on a black, central fixation point; successive patterns were updated at a frame rate of 10/s. Participants were instructed to maintain fixation on a central fixation point in the center of the CFS stimulus. Owing to the potency of CFS stimuli, the adaptation grating viewed by the other eye remained continuously suppressed from visual awareness on nearly all trials, meaning that participants were unaware of the directional congruency between auditory and visual motion. On those rare trials where the grating pattern breached interocular suppression, even briefly, participants terminated the trial by pressing the space bar; the trial was repeated later in the experiment.

Immediately following the 11-s adaptation phase, the test phase ensued: (i) the CFS array disappeared from view in the previously dominant eye and was replaced by an uncontoured, uniform gray square within the fusion frame, and (ii) at the same time, a 16% contrast, stationary test grating was presented to the eye previously exposed to the adaptation motion. Participants were tasked with deciding whether this test grating appeared to drift and, if so, to indicate its direction and duration of drift by pressing one of two computer keys when the grating no longer appeared to move. If they failed to see any "motion" of the test grating at the onset of the test phase, participants pressed the third key to report 'no-MAE'. In those trials reporting 'no-MAE', the MAE duration was recorded as 0. The keypress reporting the participant's response also triggered the trial number prompting initiation of the next trial that started with the 15-s period of random noise designed to erase any residual, subthreshold effect of adaptation.

Participants performed a total of 54 trials, 48 of which were devoted to the following six AV (AV) conditions: 2 visual adaptation directions (leftward, rightward) × 3 sound conditions [congruent (CON), incongruent (INC), NS] × 8 repetitions. Randomly inserted among those 48 trials were six catch trials on which a stationary grating was presented during the adaptation phase, their purpose being to estimate the false alarm rate and exclusion criteria for reporting an MAE. Participants were not informed that these catch trials would be included. The order of trials was randomized.

The setup and the procedures of the two pilot experiments that preceded the main experiment closely mirrored those described earlier for the main experiment. The aim of Pilot Experiment 1 was to determine an appropriate level of contrast to be used in the main experiment. To do this, we measured the contrast-response function (MAE duration as a function of adaptation contrast) using the same adaptation/test sequence as that employed in the main experiment with the exception of the omission of CFS and sound. The contrast of the monocularly viewed, visible grating varied from trial to trial over a 1.15 log-unit range in ∼0.25 log-unit steps (i.e. 3.16%, 5.62%, 10%, 17.78%, 28.48%, and 43.29%). There were 24 trials in total: 6 contrast levels × 2 visual adaptation directions (leftward, rightward) × 2 repetitions. The purpose of Pilot Experiment 2 was to determine whether (i) the monocular adaptation grating of the selected contrast level from Pilot Experiment 1 remained fully suppressed from awareness in the presence of the dynamic CFS mask viewed by the other eye and (ii) whether that suppressed monocular adaptation grating still produced reliable, residual MAE. Participants were monocularly adapted to a visible moving grating without CFS (i.e. no-CFS condition) or to an invisible moving grating suppressed by CFS in the other eye (i.e. CFS condition) for 11 s. Sounds were not presented in any of these trials. For the CFS condition, if the suppressed moving grating appeared during adaptation, participants pressed the space bar to indicate a breach in CFS. There were 16 trials in total: 2 visual adaptation directions (leftward, rightward) × 2 adaptation conditions (visible, invisible) × 4 repetitions.

Results

Results from pilot experiments

Pilot Experiment 1: MAE duration is dependent on contrast

The strength of the MAE produced by adaptation to translational visual motion (also known as the waterfall illusion) depends on the contrast of the adaptation stimulus, with this dependence

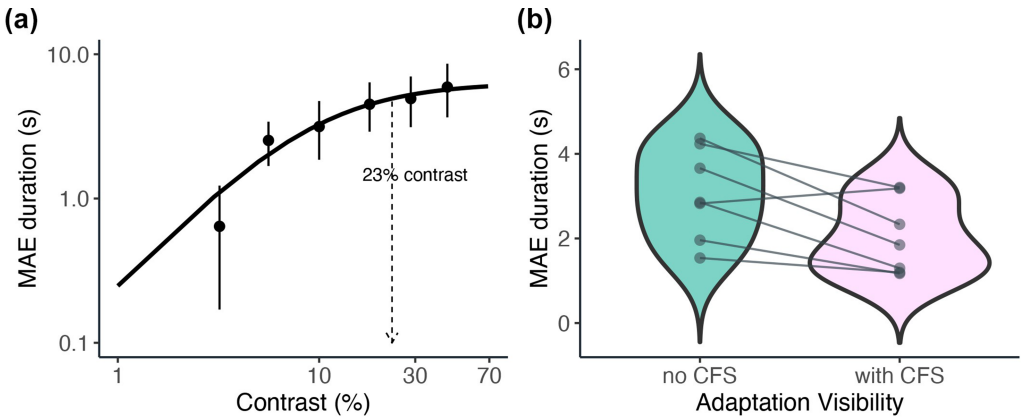

Figure 2. The results of Pilot Experiments 1 and 2. (a) MAE durations produced by different contrasts of a monocularly presented drifting grating viewed for 11 s prior to inspection of a stationary version of that grating. Data are shown in a log scale. Filled circles designate bootstrapping mean MAE durations ($n = 5$), and vertical bars demarcate the 95% confidence intervals derived by bootstrapping (10,000 iterations with replacement). The data were fitted using the Naka-Rushton equation (Naka and Rushton 1966), resulting in the dark, solid curve. From this curve, an adaptation contrast (23%) was selected for use in Pilot Experiment 2 and in the main experiment. (b) Duration of MAE produced by a monocularly presented drifting grating that was either visible during 11-s adaptation period (no CFS) or suppressed from visibility by a presentation of a dynamic interocular mask (with CFS). Seven individuals participated in this pilot experiment none of whom were members of the group tested in Pilot Experiment 1. The pairs of individual circles connected by lines are median values for a given participant.

taking the form of a compressive nonlinearity that reaches an asymptotic ceiling once contrast exceeds an intermediate level (Keck et al. 1976, Nishida et al. 1997, Blake et al. 2006). Because our main experiment seeks to learn whether auditory sound can boost the effective strength of the visual adaptation motion, it is essential that the contrast employed for visual adaptation does not produce an asymptotic level of MAE.

Figure 2a confirms that the contrast-response function generated under the conditions of our experiment exhibits the well-established, canonical form characteristic of the relation between visual contrast and psychophysical results (e.g. Vinke et al. 2022); MAE durations increase with adaptation contrast within an intermediate range of values after which a ceiling duration value is reached. We should stress that these MAE measurements were produced by adaptation periods that were purposefully brief, i.e. 11 s, and they were not preceded by an initial, long period of adaptation as often used in visual adaptation experiments. Moreover, a given trial was never initiated until all traces of an MAE from the previous trial had abated. From the contrast-response curve, we selected for the main experiment an adaptation contrast of 23%, a value producing reliable but non-asymptotic MAEs.

Pilot Experiment 2: interocular suppression weakens but does not abolish motion adaptation

From the earlier work, we knew that when a monocularly viewed motion adaptation stimulus is removed from awareness by interocular suppression, the resulting MAE may be weakened but not necessarily abolished, depending on adaptation contrast (Blake et al. 2006, Maruya et al. 2008, Kaunitz et al. 2011). We needed to ensure that this was true for the conditions employed in this study, particularly given the relatively brief adaptation period deployed in our main experiment. Specifically, we reasoned that for sound potentially to interact with visual motion processing there must be some reliable, residual neural signals of visual motion in order for sound to impact adaptation. So, using the empirically established, non-asymptotic contrast level from Pilot Experiment 1, i.e. 23%, we determined whether a monocular adaptation stimulus could induce a subsequently experienced MAE even when that adaptation stimulus was abolished from awareness by CFS for the entire 11-s period of adaptation.

Figure 2b shows MAE durations for two conditions of adaptation visibility: visible monocular adapting motion (no-CFS condition) and completely suppressed adapting motion (CFS condition). MAE durations for no-CFS and CFS trials averaged 3.0 vs. 2.1 s, respectively. A one-tailed paired t-test comparing MAE durations for these two conditions reveals a statistically significant effect of adaptation visibility on the MAE duration [$t(6) = 3.36$, $P = .015$, Cohen's $d = 1.27$, $BF_{10} = 7.15$]. As expected, visual suppression by CFS does effectively reduce the duration of MAE. At the same time, a residual MAE was reliably found in the CFS condition (i.e. all seven participants reported MAE durations on the large majority of CFS trials), implying that neural motion signals do indeed survive despite complete suppression of the motion stimulus for the entire 11-s period of adaptation. In other words, when a visual stimulus is erased from visual awareness, its adaptation potency is weakened but not abolished. We attribute this reduction in MAE strength to the invisibility of the visual adapting stimulus by interocular suppression from the CFS not to the mere presence of the Mondrian viewed by one eye. This inference is consistent with the previous research showing that the reduced strength of adaptation aftereffects following monocular adaptation only happens when the CFS mask viewed by the other eye appears at the same perceived location as the monocular adapting stimulus presented to the other eye—otherwise the adapting stimulus maintains its effectiveness even if ongoing CFS viewed by the other eye appears at a neighboring regions of visual space (afterimage adaptation, Fig. 3, Tsuchiya and Koch 2005; tilt aftereffect adaptation, Fig. 2, Kanai et al. 2006).

These pilot experiment results set the stage for asking whether the auditory motion heard during visual motion adaptation can modulate the strength of this residual MAE and, if so, whether that modulation is related to the congruence between auditory and visual motion.

Results from the main experiment

Following the dictum ["As soon as you have collected your data, before you compute any statistics, look at your data…if you

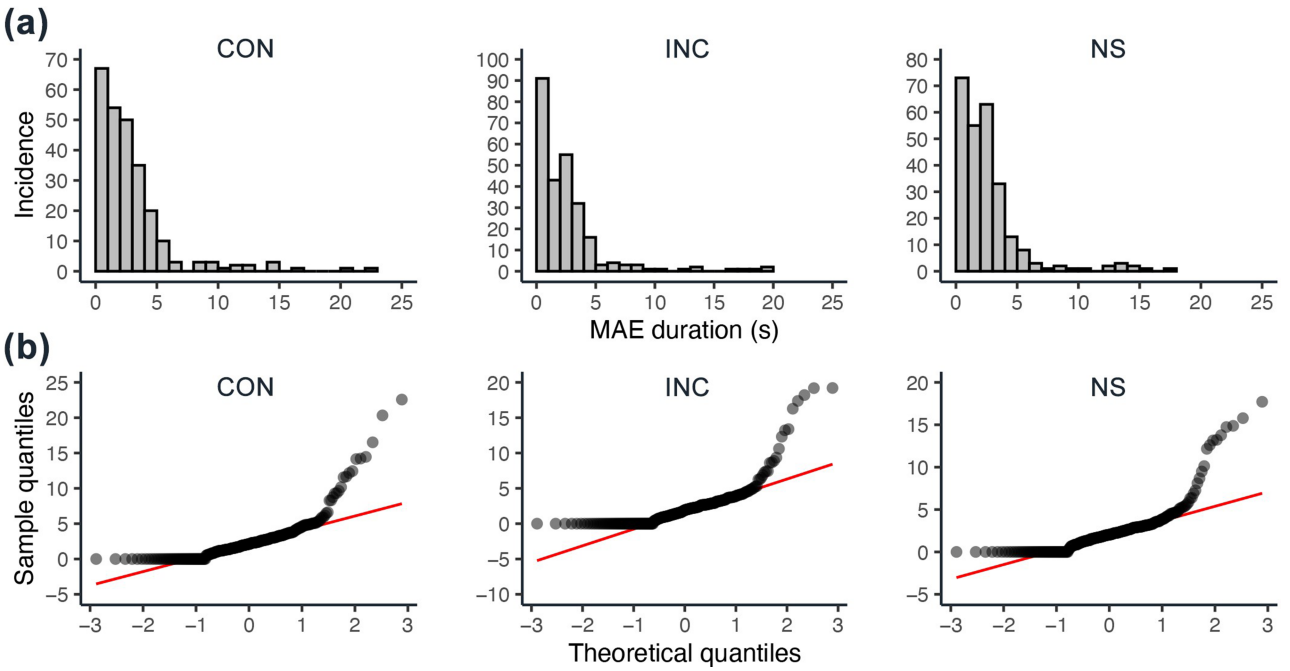

Figure 3. Overall views of the individual MAE duration values comprising results from the CON, INC, and NS conditions. (a) Distribution plots illustrate the incidence of those values using the Freedman–Diaconis method for binning as implemented in JASP and R. (b) It plots data points using the Q–Q procedure that highlights areas within the dataset that deviate from the normality (diagonal line).

assess hypotheses without examining your data, you risk publishing nonsense" (Wilkinson and the APA Task Force on Statistical Inference, 1999, p. 597).] referenced by Fife (2020) on his JASP blog, we began by carefully inspecting the raw data to ensure that we followed best-practice guidelines for testing the questions we set out answer. The following four subsections explain what we learned from those inspections and how we dealt with the questions that emerged concerning data analysis and data pruning.

Breakthrough incidence

Of crucial importance, the CFS display presented to the dominant eye was indeed highly successful in producing complete suppression of the visual adaptation motion presented to the other eye throughout the 11-s adaptation phases; breakthroughs were reported on only 33 trials, i.e. fewer than 4% of the 918 trials comprising the experiment. Among those breakthrough trials, the incidence across the CON, INC, and NS conditions was 4%, 2%, and 5%, respectively. We calculated the nonparametric Spearman correlation between incidence of breakthroughs and duration of MAE across all participants for each AV condition, and none of those three correlations came close to statistical significance [CON: $r_s = -0.085$, P (two-tailed) = .74; INC: $r_s = 0.072$, P (two-tailed) = .78; NS: $r_s = -0.024$, P (two-tailed) = .92]. As a reminder, each breakthrough trial was flagged when it occurred, and a replacement trial was repeated later in the trial sequence. Thus, the total number of successful trials (i.e. no breakthroughs) was constant across conditions and participants.

Trials on which MAE direction was in the "unexpected" direction

The hallmark feature of the MAE is motion in a direction opposite that of the immediately preceding adaptation motion. Yet on a very small fraction of trials (total of 38 out of 816) in our experiment, participants' button press responses signified an experience of illusory motion in the "same" direction as the visual adaptation motion. These rare trials arose with all three AV conditions (6%, 4%, and 4% for CON, INC, and NS trials, respectively), and the average MAE duration on these trials was 2.9 s. We conjecture that at least some of those trials may be attributable to keypress mistakes reporting the perceived direction of illusory motion (there were three keypress options following each adaptation period). Nevertheless, we were unable to see a rational way to incorporate duration values from these trials into the formal analyses, so data from these trials were excluded from the dataset. The downside to doing this is that all three conditions now do not have an equal number of trials (CON has several more exclusions than do INC and NS). Fortunately, the variance estimates within the datasets to be compared are approximately equivalent (homogeneity of variance), which mitigates the impact of the unequal number of samples (hence, degrees of freedom).

Trials on which an MAE was not experienced

On some trials, participants' responses following an 11-s period of adaptation to visual motion indicated that they failed to experience illusory motion following adaptation and, thus, we recorded the MAE duration for that trial as zero; the overall incidence of these kinds of trials on the CON, INC, and NS conditions was 23%. This was not entirely surprising, because during the pilot experiments described earlier, which did not involve sound, there were infrequent trials when adaptation failed to elicit an MAE.

One can imagine why in a blocked set of 54 trials a participant might occasionally fail to experience an MAE. For example, attention may have lapsed during an adaptation period, and it is well known that focused attention strengthens the MAE (Georgiades and Harris 2000, Huk et al. 2001, Rezec et al. 2004, Kaunitz et al. 2011). Alternatively, perhaps one impact of sound on visual motion is to weaken the neural strength of visual motion when sound and

vision are incongruent, in which case those zero-duration trials are meaningful. To see if there is evidence for this speculation, we examined the zero-duration trials to see how many happened on CON trials (in which case we would expect strengthened motion adaptation and hence longer durations) and how many happened on INC trials (in which case we would not expect strengthening of motion adaptation and, if anything, weakening and, perhaps, no-MAE). In fact, the incidence of zero-duration trials for the CON and INC trials was 20% and 26%, respectively, not compelling evidence for the speculation.

All things considered, we are reluctant to invalidate these zero-duration trials just because they run counter to expectation. So, to be on the safe side, we performed and reported below several pertinent statistical analyses on both datasets: one comprising all trials and the other comprising only trials producing measurable MAEs (i.e. zero-duration trials removed). Supplementary Figure S1 gives the subject-by-subject summary of descriptive statistics for data with zero durations included and data with zero values excluded.

MAE reports following adaptation to a stationary grating

One would not expect a "stationary", invisible grating presented during the adaptation period to induce an MAE, and for 80 out of the total 102 trials involving no visual motion that was indeed the reported experience. As for the 22 trials on which participants "did" report experiencing illusory motion after 11 s of exposure to a stationary grating suppressed from awareness, the average reported MAE duration was 2.6 s. Why might these "false alarms" occasionally arise?

One might reasonably wonder whether hearing auditory motion for 11 s prompted appearance of illusory visual motion immediately following that noise exposure. After all, earlier studies have found that the perceived direction of visual motion can be biased by prior exposure to auditory sound that mimics rising or falling musical pitch (Hedger et al. 2013) or lateral sweeps of a sine-wave tone (Hidaka et al. 2011). In our experiment, might audible sound motion on its own cause visual motion adaptation or, alternatively, prompt visual motion priming? For several reasons, visual adaptation to sound alone or priming by sound alone seems unlikely. For one thing, the incidence of these false alarms is low and, for another, the reports of illusory motion direction following a given no-motion adaptation period were not consistently related to the direction of the sound presented during that 11-s exposure period (nine "consistent" reports and seven "inconsistent" reports). Moreover, illusory visual motion was also reported following six of the trials in which the static visual grating was unaccompanied by sound at all.

We surmise, instead, that expectation engendered by task instructions and the actual experiences on the vast majority of trials may have unwittingly encouraged false alarm responses following some trials when visual motion was not presented. After all, participants were not told that the trial sequence would include "catch trials", i.e. trials where they were likely not to experience illusory motion. Indeed, on the vast majority of trials, they did experience and report illusory motion. Thus, participants may have expected to experience illusory motion following each adaptation period, a mindset that could have biased them to report something on trials where the sensory evidence evoked decision uncertainty. In any event, the infrequency of these putative false alarms and the pattern of conditions under which they happen have no obvious bearing on the interpretation of results on trials involving adaptation to visual motion and it is to those results on those 816 trials that we turn next.

MAE durations are not normally distributed

Figure 3a illustrates the distributional profile of all 816 MAE duration values from the CON, INC, and NS conditions (i.e. 17 participants × 3 AV conditions × 16 trials), including the zero-duration values. These plots disclose the distinctly non-normal shape of those frequency distributions, due in part to the presence of 0-s durations within the complete dataset (the issue discussed earlier). Departure from normality of those data is also reflected in the Q–Q plot in Fig. 3b, where kurtosis shows up as deviations of data from the prediction line at the two tails. The results from the Shapiro–Wilk test implemented in JASP also confirm significant departures from normality of these data ($P < .0001$).

Fortunately, there are established ways to address this situation. The most straightforward approach is to use nonparametric inferential statistical tests for which the assumption of normality is relaxed, and this is what we have done in the remainder of the Results section. (Another tactic for dealing with non-normal data is to reduce skewness by applying a log-transformation of the data. But the log transform returns NaN (also known as not a number) for zero values, so this is not practical for the larger dataset.)

Influence of sound motion on MAE strength produced by adaptation to invisible visual motion

We now turn to the central question: does the presence of auditory motion during visual motion adaptation to a stimulus suppressed from awareness by CFS impact the strength of the MAE? And, if so, does that impact depend on the congruence between direction of auditory motion and direction of visual motion? These are the two main questions we set out to answer from our dataset. As we move through those analyses it is important to keep in mind that the three kinds of AV trials were randomly intermixed within a single block. On the vast majority of trials, moreover, participants had no idea what visual motion condition they were experiencing on any given trial, owing to the potency of CFS; those rare trials where the moving grating achieved visibility were not included in these analyses.

A total of six participants were excluded from the following analyses: one participant reported that no movement was visible during the experiment, with an average MAE duration of >1 s. Additionally, for catch trials (i.e. presenting stationary grating during the adaptation phase), five participants reported MAEs exceeding 50% and the average duration surpassing 2 s. With the remaining 17 participants' data, an initial analysis comparing MAE durations for leftward vs. rightward motion using the nonparametric Wilcoxon signed ranks test revealed that the median durations for the two sets of directions were not significantly different ($P = .59$). Thus, we pooled durations over the two directions of motion for all subsequent analyses.

Using this pooled dataset, we analyzed the data for each participant to get a sense of the range of individual differences among the 17 participants. Figure 4 shows the datasets including zero-duration trials in boxplot format for each participant's data combined over the three AV conditions (see Supplementary Fig. S2 for boxplots of the datasets with zero-duration trials removed). From those plots, it is obvious that the median duration (dark horizontal bar in each plot) varies among individuals, but within a fairly narrow range, with one distinct exception: the individual designated as P6 reported MAE durations for which the median value is about five times longer than the median combined over the other 16 participants. Clearly, P6 is a statistical outlier by standard criteria (e.g,

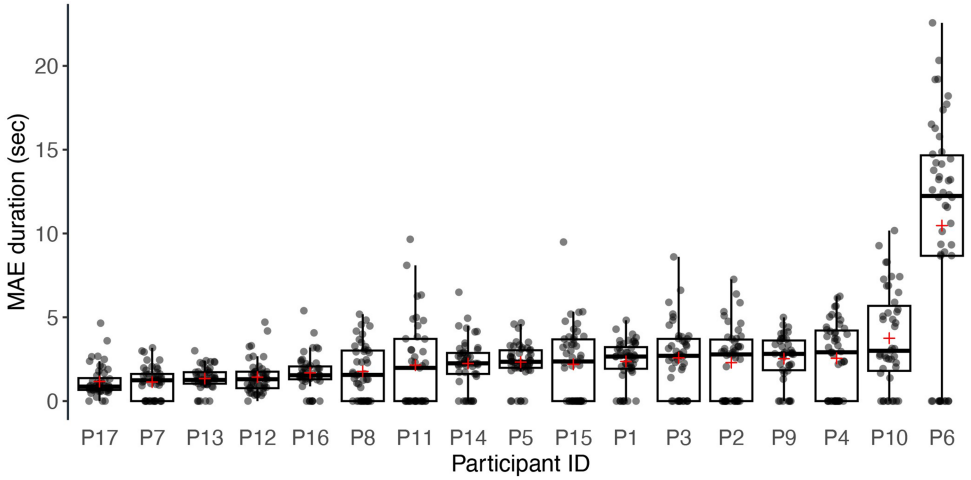

Figure 4. Boxplots summarizing for each participant the combined MAE durations for the CON, INC, and NS conditions, ordered by increasing values of the individual median for each individual (indicated by the dark horizontal bar within the box). Crosshairs indicate the mean for each individual. The durations comprising these plots are from the dataset with zeros included. The same pattern of results can be found in the plots created using the smaller dataset with zeros excluded (see Supplementary Fig. 2).

Tukey 1977), and we pondered whether to remove that person's data from the analysis set. But again, to be on the safe side, we opted to report results with data from P6 included.

At first glance, MAE mavens may be puzzled by the brevity of these MAE durations. But keep in mind that we intentionally set the adaptation conditions—moderate contrast grating and brief duration—to promote robust suppression of the adapting stimulus. Indeed, relatively brief MAE durations were anticipated, and they align with the results from the pilot experiment summarized in Fig. 2b.

For any given individual, durations among the three AV conditions were similar as evidenced by the high correlations for pairwise calculations among conditions derived by Spearman's rho (r_s CON/INC = 0.85; r_s CON/NS = 0.67; r_s INC/NS = 0.73, all highly significant with $P < .01$). The durations averaged over all participants for each of the three AV conditions differ from one another by only fractions of a second: the means for CON, INC, and NS are 2.8 s, 2.4 s, and 2.6 s, respectively, when computed with zero trials included; means without including zero-duration trials are 3.7, 3.4, and 3.4. Removal of the data of P6 (the outlier in Fig. 4) does not alter the pattern of results, just the absolute values. These very small differences between the means of these AV conditions foretell that any impact of auditory sound must be subtle. That said, we performed analyses to evaluate those differences from two complementary perspectives.

To capture a global overview of the relation between MAEs obtained on CON and INC trials, we compared the individual trial results for these two conditions using the shift-plot format described by Rousselet et al. (2017). With this graphic procedure, MAE durations for CON and INC conditions were grouped into separate distributions, and within each distribution, individual MAE durations were grouped into bins each containing the same number of durations for a given duration bin (within the constraint of rounding errors). Those shift plots are shown in Fig. 5a. In creating these plots, we included zero-duration trials (conspicuously shown at the far left of both shift plots). The solid dark lines within each dot cloud denote the median durations for each of the 10 bins (but apparently >10 bins as several bins have median durations of 0 owing to the large incidence of zero-duration trials), and the gray lines connect pairs of median values for corresponding bins associated with the CON and INC datasets. The rightward offset tendency of the CON decile medians relative to the INC medians reveals that CON durations tend to be slightly longer than INC durations. This tendency no doubt undergirds the small differences in average MAE durations between CON and INC noted earlier.

Our second analysis asks to what extent differences in MAE durations between CON and INC are statistically significant when considered from the standpoint of individual participants. For this analysis, we construe the NS condition as providing a baseline estimate of the strength of the MAE for each individual. Thus, an impact of sound on visual adaptation strength would be revealed by a reliable deviation from that person's NS baseline. Figure 5b shows the change in MAE for the CON and the INC conditions, expressed as a percentage change in MAE strength relative to that person's MAE under the NS condition [i.e. (CON – NS)/NS and (INC – NS)/NS]. Based on the Shapiro–Wilk test as implemented in JASP, the distributions of the CON and INC index values do not depart significantly from normality (for CON, W = 0.90, $P = .07$; for INC, W = 0.97, $P = .89$). We thus employed the parametric t-test to evaluate the difference between the two AV conditions. The results revealed that the percent-change index was significantly larger for CON than for INC [$t(16) = 2.129$, $P = .02$, Cohen's $d = 0.52$, $BF_{10} = 2.89$]. The average percent-change score for CON is positive (+10.30%), and the average percent-change score for INC is negative (–6.30%).

While statistically significant, by conventional standards, these differences can only be characterized as "modest" (Cohen's d) or as "anecdotal evidence" for an effect of AV (Bayes factor). Moreover, those interpretations are not mitigated by analyses based on nonparametric statistics (Wilcoxon signed-rank test) or by analyses with zero durations removed from the data arrays nor can we ignore that a few of the participants produced CON difference scores that were substantially "negative", implying that MAEs for those people were "weaker" following adaptation accompanied by congruent sound relative to adaptation unaccompanied by sound. Additionally, 5 out of 17 participants reported an opposing pattern, showing shorter MAE durations in CON than in INC, as illustrated in Fig. 5(b). This subgroup might have contributed to the statistically significant but small-sized effects. While there was no

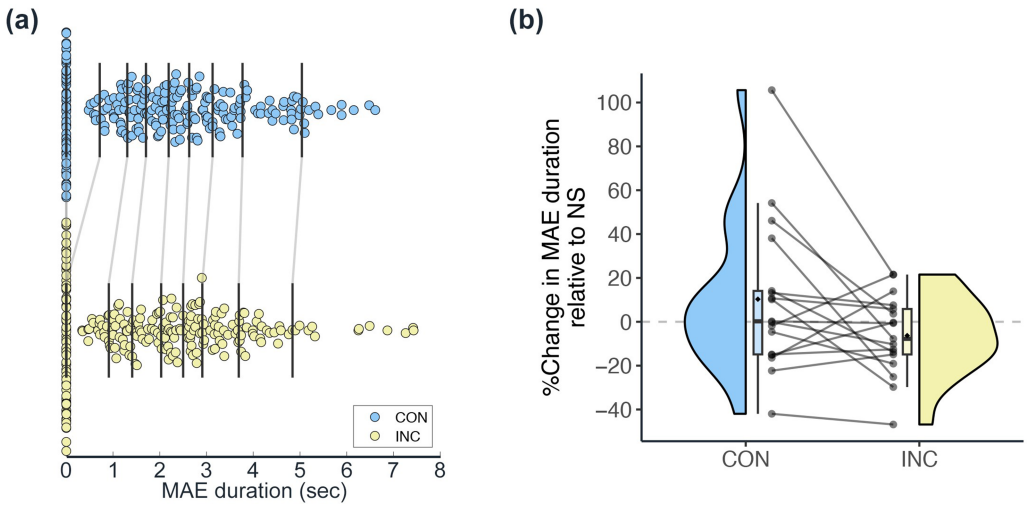

Figure 5. The results of the main experiment. (a) Shift plots summarizing and comparing individual MAE durations for CON and INC conditions (P6 the outlier was excluded for graphical clearance). (b) Cloud plots of values of percentage change in MAE duration for CON and INC conditions, relative to the NS condition. The circles are average MAE durations for each of the 17 participants.

conspicuous distinctiveness in the data from these participants in terms of breakthroughs and incorrect reports, they did, on average, exhibit more zero-duration trials in the CON condition (29%) than they did in the INC condition (16%). For the other 12 participants, the opposite pattern of results was observed, i.e. more zero-duration trials on INC trials than on CON trials. We remain bemused by these individual differences.

Discussion

Using the CFS procedure to render monocularly viewed visual motion perceptually invisible, we found that exposure to motion outside of awareness during periods of adaptation can nonetheless induce measurable MAEs on nearly all trials. This aspect of our finding replicates earlier studies that used binocular rivalry (Lehmkuhle and Fox 1975, Van Der Zwan et al. 1993, Blake et al. 2006) or CFS (Maruya et al. 2008, Kaunitz et al. 2011, Khuu et al. 2014) to induce MAEs to visual motion suppressed from awareness. The novel twist in the present study is the incorporation of auditory motion during visual adaptation to invisible motion. This resulted in the discovery that most, but not all, participants experience stronger MAEs when invisible adaptation motion is accompanied by translational auditory motion congruent in velocity (i.e. direction and speed) to the visual motion. As a reminder, during adaptation trials, participants continuously monitored what they were seeing and indicated whether they saw hints of moving contours within any portion of the dynamic CFS mask. On those infrequent trials, when a patch of visual motion within the CFS mask achieved awareness, those trials were aborted. In other words, the trials contributing to the measurement of MAE durations were only trials on which participants were unaware whether the adapting grating was moving leftward or rightward or was stationary. They did know, of course, what they were hearing, but that audible stimulus was uncorrelated in direction with the visual motion being viewed but not seen. Moreover, results on trials producing false alarm reports of an MAE confirm that the direction of auditory motion did not bias participants' reported directions of MAE motion. We believe that the impact of congruent auditory motion on the visual MAE implicates AV motion interaction during adaptation despite the absence of explicit knowledge about the congruence of those signals.

Strength of audiovisual interactions

The AV congruency effect measured in our study, while statistically significant, is small in magnitude—the differences in MAE durations between CON and INC conditions average ~0.4 s. Moreover, the majority—but not all—of the 17 participants exhibited longer MAE durations for the CON condition. Given those two features of the results, it is important that we consider possible factors besides congruence that might have influenced participants' performance of the task. One such factor is attentional allocation, which is known to influence the magnitude of the MAE under some conditions (Nishida and Ashida 2000, Mukai and Watanabe 2001, Rezec et al. 2004, Taya et al. 2009, Bartlett et al. 2018) but not others (Wohlgemuth 1911, Morgan 2011, 2012, 2013, Pavan et al. 2015). In those previous studies, attention was explicitly manipulated to create testing conditions promoting different amounts of attentional allocation. The design of our study, however, required sustained, focused visual attention on the CFS mask throughout the short periods of visual motion adaptation. To whatever extent the accompanying sound distracted attention during adaptation, that distraction would perforce be unrelated to the direction of visual motion, because auditory and visual motion were uncorrelated in direction over trials.

Turning to another possible factor we considered, could the perceived direction of the audible sound have engendered pursuit eye movements that altered the effectiveness of the visual motion being imaged on the retina? That seems unlikely for several reasons. For one thing, the bulk of prior evidence indicates that visually evoked pursuit eye movements are not crucially involved in the induction of MAEs (Mack et al. 1987, 1989, Morgan et al. 1976, Swanston and Wade 1992, but see Anstis and Gregory 1965). For another, in our previous MAE study (Park et al. 2019), we measured eye stability during periods of adaptation to visual motion accompanied by moving sound and found no differential patterns of fixational eye movements dependent on AV congruency. The AV conditions in the current study closely approximated those from our previous study, and the CFS mask itself portrayed dynamic, unstructured changes in masking elements that should not bias eye movements in a given direction. Based on these considerations, we seriously doubt that differential eye movements played a crucial role in the present findings.

A third factor that could possibly influence the magnitude of AV interaction in our study concerns the nature of the task used to measure the MAE. Participants pressed a key to indicate when illusory motion had slowed to a halt following adaptation to invisible motion, a judgment requiring establishment and maintenance of a sense of what constitutes "cessation" of motion. As Morgan (2012) has opined, this judgment risks being influenced by "an unconscious wish to give the experimenter the desired result" (p. 47). Unless they were blatantly ignoring instructions, that risk is mitigated in our study by the participant's lack of knowledge about the direction of visual motion relative to sound motion on any given trial. It is possible, of course, that the "range" of individual differences in overall MAE durations within our dataset could arise, at least in part, from differences in people's criterion for what constitutes cessation of illusory motion. For what it is worth, individual differences in MAE strength are evident in other studies that have indexed MAE with dependent variable other than duration (e.g. Kaunitz et al. 2011, Morgan 2013).

A fourth possibility for the small differences between CON and INC trials concerns the actual stimulus conditions in our study. We utilized drifting gratings that were of sufficient contrast to generate non-asymptotic MAEs when suppressed by CFS yet sufficiently weak to remain visually suppressed by CFS for the entire duration of visual adaptation. Paired with those relatively low-contrast visual patterns were binaurally presented noise bursts that were themselves sufficiently audible to yield a clear sense of lateral sound motion. In the multisensory integration literature, there is a well-established principle known as inverse effectiveness (e.g. Meredith and Stein 1983, Holmes 2007, Spence 2018): multisensory integration is most effective when the response produced by the more robust unisensory stimulus is relatively weak. Neither of the stimuli comprising the bisensory events in our experiment would qualify as weak when considered in terms of sound amplitude or luminance contrast. It is true that the monocular visual motion stimulus was weakened in its effective contrast by the influence of the CFS mask viewed by the other eye. But it remained effective nonetheless as evidenced by its ability to produce a measurable MAE (recall Fig. 2b). Inverse effectiveness is typically conceptualized within the context of underlying neural activity, a perspective we will discuss shortly.

Finally, the weakness of the AV congruency effect measured in our study might have something to do with our choice of translational motion to induce the MAE. That decision was made for several reasons. First, it is straightforward to create and present an auditory analog of lateral motion, and our previous study confirmed that this conventional MAE was sufficient to reveal reliable direction-selective interactions between auditory and visual motion. Second, neural processing of visual motion transpires within hierarchically organized stages, starting with the primary visual cortex and proceeding to extra-striate cortical areas comprising the dorsal stream (e.g. Andersen 1997). The initial stages of motion processing are responsive to local, translational motion, while higher stages register more complex, global forms of motion including rotation, expansion/contraction, and pendular trajectories characteristic of animate motion of the arms and legs. For our purposes, translational motion was necessary because interocular suppression of translational motion reduces but does not abolish adaptation to translational motion (Blake et al. 2006, Lehmkuhle and Fox 1975; Fig. 2b), whereas more complex patterns of motion are quashed by interocular suppression (Wiesenfelder and Blake 1990, Van Der Zwan et al. 1993). This limits the types of motion that afford a clean dissociation between awareness and motion adaptation.

Alternative conceptualizations of sound's influence on MAE

This brings us to our final point of discussion, i.e. two alternative ways to conceptualize the nature of the interactions between audible sound motion and invisible adaptation motion implicated in our study. One frames the issue within the context of causal inference, and the other focuses on possible underlying neural substrates. It should be stressed that those alternatives are not mutually exclusive, and for our purposes, both are worth considering.

Causal inference

One approach, dubbed causal inference, emerges from the perspective that multisensory integration—like other aspects of perception and cognition—involves causal inference (Shams and Beierholm 2010). To paraphrase from our Introduction section, the neuro-sensory signals that spark perception require concomitant interpretative processing based on prior knowledge, current needs, spatial and temporal context, and, importantly, the relative strength of evidence favoring alternative interpretations (Knill and Richards 1996). According to Shams and Beierholm (2010), causal inference can be especially challenging when it comes to sensory integration across multiple modalities:

> At any given moment, an individual typically receives multiple sensory signals from the different modalities and needs to determine which of these signals originated from the same object and should be combined and which originated from different objects and should not be bound together….This problem can be challenging even in the simplest scenario with only one visual stimulus and one auditory stimulus (p. 426).

In the context of causal inference, there are two questions that arise when considering our results. The first concerns whether the stimulus events in our experiment are sufficient to promote the integration of the auditory and visual components experienced during adaptation. Auditory motion was being heard over headphones, while visual motion was being presented on a computer screen located directly in front of the observer. Does that affect the likelihood of integration of the two sources of information? We think not. Our conditions of stimulus presentation are not unorthodox: they have been employed in numerous earlier studies demonstrating AV binding (Baumann and Greenlee 2007, Deas et al. 2008, Hidaka et al. 2011, 2017, Rosemann et al. 2017) including our own earlier study (Park et al. 2019). Moreover, the present testing setup encouraged a sense of union between sound and vision by phase-locking modulations of sound amplitude with modulations of contrast of the moving grating: both stimuli waxed and waned in synchrony during the adaptation period. Observers were exposed to those conditions without suppression of visual awareness during the introduction to the study prior to the actual experiment.

During the main experiment, of course, visual motion, while being imaged on the retina, was blocked from conscious awareness by the presentation of the CFS mask seen by the other eye. This introduces the second question: can the computational operations putatively subserving causal inference transpire at a level of processing prior to the emergence of awareness? In other words, can information embodied in a suprathreshold auditory stimulus interact with information associated with a suppressed visual stimulus that nonetheless remains sufficiently effective to induce

an adaptation aftereffect following adaptation? A number of published studies using visual priming (Faivre et al. 2014), flash suppression (Palmer and Ramsey 2012, Aller et al. 2015), and binocular rivalry (Chen et al. 2011a, Alsius and Munhall 2013) have arrived at an affirmative answer to that question. Similarly, prior learning of associations between auditory cues and visual colors subsequently empowers colors to emerge from suppression during motion-induced blindness when accompanied by congruent—but not incongruent—auditory cues (Chang et al. 2015). What remains unresolved, however, is the extent to which prior conscious experience of integrated AV items (e.g. mouth movements and speech sounds) is necessary for subsequent binding outside of awareness. As mentioned in the previous paragraph, observers had some exposure to the temporally modulated visual grating and accompanying sound. In a related vein, it remains to be learned at what level of the perceptual/cognitive processing hierarchy integration occurs (Shams and Beierholm 2010, Noel et al. 2015). This latter question brings us to the second framework within which our results can be considered, namely the neural substrates of the MAE.

Neural concomitants of the MAE

How might the presence of audible sound motion weakly but significantly bolster the potency of adaptation produced by invisible visual motion? As is well known, a subset of neurons in early visual areas, most notably the primary visual cortex (Movshon and Newsome 1996) and the medial temporal cortex (Mikami et al. 1986), exhibit direction selectivity meaning that they respond strongly only to translational motion within a narrow range of directions, with preferred direction of motion varying among neurons. It is widely believed that this category of neurons provides the neural substrate for the visual MAE (e.g. Stocker and Simoncelli 2009), a belief reinforced by human brain imaging studies (Petersen et al. 1985, Tootell et al. 1995, Huk et al. 2001). Human brain imaging also reveals that these visual areas can be activated by auditory information including sound motion (for V1, see Watkins et al. 2007, 2006; for V5/MT+, see Alink et al. 2008, Lewis et al. 2000, Poirier et al. 2006, Saenz et al. 2008, Wolbers et al. 2011). Those auditory activations in V1 and V5/MT+ could arise from long-range feedback originating from higher-level multisensory areas and even from the auditory cortex—such feedback connections have been identified in nonhuman primates using retrograde tracing techniques (Falchier et al. 2002, Clavagnier et al. 2004). Of direct relevance to feedback in the human brain studied using functional magnetic resonance imaging, Rezk et al. (2020) were able to successfully decode the direction of auditory motion from voxel-wise activation patterns in visual area V5/MT+ of the human brain. They interpreted their results to imply that

> shared motion-direction information between vision and audition in hMT + /V5 provide evidence that hMT +/V5 may play a crucial role in providing a common representational structure between the two modalities to link auditory and visual motion-direction information. The presence of a common brain code for directional motion in vision and audition might potentially relate to psychophysical studies showing cross-modal adaptation effects for motion directions (p. 2296).

Considered together, these converging lines of neurophysiological evidence suggest that the bolstered MAE associated with congruent auditory and visual motion during adaptation could arise from neural integration within those visual areas that putatively mediate the MAE, i.e. V1 and MT+/V5. Inspired by the findings summarized earlier, we have considered theoretical models of visual MAE as a framework for thinking about how sound might be impacting visual motion portrayed by moving patterns rendered invisible by potent interocular suppression.

The consensus theory of the MAE (Mather 1980, Grunewald and Lankheet 1996, Anstis et al. 1998) posits that motion information is initially registered within local regions of the retinal image; this first stage is thought to be embodied in the direction-selective neurons in V1. The second stage implements the integration of local motion signals, an operation carried out by neurons such as those identified in MT+/V5. Versions of the model further assume that neurons generate narrowly tuned excitatory responses while at the same time exerting broadly tuned inhibitory interactions among neighboring neurons preferring similar directions of visual motion. It is further assumed that this network of neurons exhibits decreased neural activity (i.e. they adapt) over the course of exposure to a given direction of visual motion, with the magnitude of adaptation governed by the level of activation within subsets of those neurons tuned to different directions of motion. The aftereffect of this differential magnitude of responsiveness during adaptation temporarily distorts patterns of activation within this network following adaptation, producing several putatively related visual aftereffects including the classic MAE we have deployed (e.g. Stocker and Simoncelli 2009).

One could surmise that feedback signals triggered by auditory stimulation activate early visual cortical areas thus promoting AV interactions impacting visual motion adaptation. A visual motion stimulus, although suppressed from awareness, does retain effectiveness in activating direction-selective neurons in the early visual cortex (e.g. Yuval-Greenberg and Heeger 2013). Perhaps this activation is potentiated by feedback signals initiated by suprathreshold sound motion, causing a stronger MAE whose decay time is lengthened. To capture the element of AV congruence, those feedback signals would need to be selective for the direction of motion along the lines described earlier in MT+/V5 (Rezk et al. 2020).

Conclusion

We demonstrated the small yet reliable AV direction congruency effect on the post-adaptation MAE duration under conditions where observers would not be influenced by cognitive knowledge about the congruence of auditory and visual motion. Thus, these results suggest that AV interactions can transpire within the early stages of sensory information processing outside of visual awareness.

Author contributions

Minsun Park (Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing—original draft, Writing—review & editing, Visualization), Randolph Blake (Conceptualization, Methodology, Formal analysis, Writing—review & editing, Visualization), and Chai-Youn Kim (Conceptualization, Methodology, Writing—review & editing, Supervision, Funding acquisition).

Supplementary data

Supplementary data is available at *Neuroscience of Consciousness* online.

Conflict of interest.

We declare we have no competing interests.

Funding

This work was supported by the National Research Foundation Grant (NRF-2017M3C7A1029659) funded by the Korean Government to C.-Y.K. and by Centennial Research Funds awarded by Vanderbilt University to R.B.

Data availability

The data supporting this study's findings are openly available on the Open Science Framework at https://osf.io/zgy7k/.

References

Alink A, Singer W, Muckli L. Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. *J Neurosci* 2008;**28**:2690–7.

Aller M, Giani A, Conrad V *et al*. A spatially collocated sound thrusts a flash into awareness. *Front Integr Neurosci* 2015;**9**:16.

Alsius A, Munhall KG. Detection of audiovisual speech correspondences without visual awareness. *Psychol Sci* 2013;**24**:423–31.

Andersen RA. Neural mechanisms of visual motion perception in primates. *Neuron* 1997;**18**:865–72.

Angelaki DE, Gu Y, DeAngelis GC. Multisensory integration: psychophysics, neurophysiology, and computation. *Curr Opin Neurobiol* 2009;**19**:452–8.

Anstis SM, Gregory RL. The after-effect of seen motion: The role of retinal stimulation and of eye movements. *Q J Exp Psychol* 1965;**17**:173–4.

Anstis S, Verstraten FA, Mather G. The motion aftereffect. *Trends Cogn Sci* 1998;**2**:111–7.

Arrighi R, Marini F, Burr D. Meaningful auditory information enhances perception of visual biological motion. *J Vis* 2009;**9**:1–7.

Bartlett LK, Graf EW, Adams WJ. The effects of attention and adaptation duration on the motion aftereffect. *J Exp Psychol Hum Percept Perform* 2018;**44**:1805–14.

Baumann O, Greenlee MW. Neural correlates of coherent audiovisual motion perception. *Cereb Cortex* 2007;**17**:1433–43.

Berger CC, Ehrsson HH. Auditory motion elicits a visual motion aftereffect. *Front Neurosci* 2016;**10**:559.

Blake R, Tadin D, Sobel KV *et al*. Strength of early visual adaptation depends on visual awareness. *Proc Natl Acad Sci USA* 2006;**103**:4783–8.

Brainard DH. The Psychophysics Toolbox. *Spatial Vision* 1997;**10**:433–6.

Chang AYC, Kanai R, Seth AK. Cross-modal prediction changes the timing of conscious access during the motion-induced blindness. *Conscious Cogn* 2015;**31**:139–47.

Chen YC, Huang PC, Yeh SL *et al*. Synchronous sounds enhance visual sensitivity without reducing target uncertainty. *Seeing Perceiving* 2011a;**24**:623–38.

Chen YC, Yeh SL, Spence C. Crossmodal constraints on human perceptual awareness: auditory semantic modulation of binocular rivalry. *Front Psychol* 2011b;**2**:212.

Clavagnier S, Falchier A, Kennedy H. Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. *Cogn Affect Behav Neurosci* 2004;**4**:117–26.

Cox D, Hong SW. Semantic-based crossmodal processing during visual suppression. *Front Psychol* 2015;**6**:722.

Deas RW, Roach NW, McGraw PV. Distortions of perceived auditory and visual space following adaptation to motion. *Exp Brain Res* 2008;**191**:473–85.

Dwyer P, Takarae Y, Zadeh I *et al*. Multisensory integration and interactions across vision, hearing, and somatosensation in autism spectrum development and typical development. *Neuropsychologia* 2022;**175**:108340.

Faivre N, Mudrik L, Schwartz N *et al*. Multisensory integration in complete unawareness: evidence from audiovisual congruency priming. *Psychol Sci* 2014;**25**:2006–16.

Falchier A, Clavagnier S, Barone P *et al*. Anatomical evidence of multimodal integration in primate striate cortex. *J Neurosci* 2002;**22**:5749–59.

Fang F, He S. Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nat Neurosci* 2005;**8**:1380–5.

Faul F, Erdfelder E, Buchner A *et al*. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009;**41**:1149–60.

Fife D. *The Visual Modeling Module*. JASP. 2020. https://jasp-stats.org/2020/04/21/the-visual-modeling-module/ (13 March 2023, date last accessed).

Georgiades MS, Harris JP. Attentional diversion during adaptation affects the velocity as well as the duration of motion after-effects. *Proc Biol Sci* 2000;**267**:2559–65.

Grunewald A, Lankheet MJ. Orthogonal motion after-effect illusion predicted by a model of cortical motion processing. *Nature* 1996;**384**:358–60.

Hedger SC, Nusbaum HC, Lescop O *et al*. Music can elicit a visual motion aftereffect. *Atten Percept Psychophys* 2013;**75**:1039–47.

Heyman T, Maerten A-S, Vankrunkelsven H *et al*. Sound-symbolism effects in the absence of awareness: a replication study. *Psychol Sci* 2019;**30**:1638–47.

Hidaka S, Higuchi S, Teramoto W *et al*. Neural mechanisms underlying sound-induced visual motion perception: an fMRI study. *Acta Psychol* 2017;**178**:66–72.

Hidaka S, Teramoto W, Sugita Y *et al*. Auditory motion information drives visual motion perception. *PLoS One* 2011;**6**:e17499.

Holmes NP. The law of inverse effectiveness in neurons and behaviour: multisensory integration versus normal variability. *Neuropsychologia* 2007;**45**:3340–5.

Huk AC, Ress D, Heeger DJ. Neuronal basis of the motion aftereffect reconsidered. *Neuron* 2001;**32**:161–72.

Jain A, Sally SL, Papathomas TV. Audiovisual short-term influences and aftereffects in motion: examination across three sets of directional pairings. *J Vis* 2008;**8**:7.1–13.

Kanai R, Tsuchiya N, Verstraten FA. The scope and limits of top-down attention in unconscious visual processing. *Curr Biol* 2006;**16**:2332–6.

Kaunitz L, Fracasso A, Melcher D. Unseen complex motion is modulated by attention and generates a visible aftereffect. *J Vis* 2011;**11**:10.

Keck MJ, Palella TD, Pantle A. Motion aftereffect as a function of the contrast of sinusoidal gratings. *Vision Res* 1976;**16**:187–91.

Khuu SK, Gordon J, Balcomb K *et al*. The perception of three-dimensional cast-shadow structure is dependent on visual awareness. *J Vis* 2014;**14**:25.

Kleiner M, Brainard D, Pelli D (2007). What's new in Psychtoolbox-3? Perception 36 ECVP Abstract Supplement

Knill DC and Richards W (eds.), *Perception as Bayesian Inference*. Cambridge, UK: Cambridge University Press, 1996.

Lehmkuhle SW, Fox R. Effect of binocular rivalry suppression on the motion aftereffect. *Vision Res* 1975;**15**:855–9.

Lewis JW, Beauchamp MS, DeYoe EA. A comparison of visual and auditory motion processing in human cerebral cortex. *Cereb Cortex* 2000;**10**:873–88.

Lewis R, Noppeney U. Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas. *J Neurosci* 2010;**30**:12329–39.

Macaluso E, Driver J. Multisensory spatial interactions: a window onto functional integration in the human brain. *Trends Neurosci* 2005;**28**:264–71.

Mack A, Goodwin J, Thordarsen H *et al*. Motion aftereffects associated with pursuit eye movements. *Vision Res* 1987;**27**:529–36.

Mack A, Hill J, Kahn S. Motion aftereffects and retinal motion. *Perception* 1989;**18**:649–55.

Maruya K, Watanabe H, Watanabe M. Adaptation to invisible motion results in low-level but not high-level aftereffects. *J Vis* 2008;**8**:7.1–11.

Mather G. The movement aftereffect and a distribution-shift model for coding the direction of visual movement. *Perception* 1980;**9**:379–92.

Mather G, Pavan A, Campana G *et al*. The motion aftereffect reloaded. *Trends Cogn Sci* 2008;**12**:481–7.

Meredith MA, Stein BE. Interactions among converging sensory inputs in the superior colliculus. *Science* 1983;**221**: 389–91.

Mikami A, Newsome WT, Wurtz RH. Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. *J Neurophysiol* 1986;**55**: 1308–27.

Morgan MJ. Wohlgemuth was right: distracting attention from the adapting stimulus does not decrease the motion after-effect. *Vision Res* 2011;**51**:2169–75.

Morgan MJ. Motion adaptation does not depend on attention to the adaptor. *Vision Res* 2012;**55**:47–51.

Morgan M. Sustained attention is not necessary for velocity adaptation. *J Vis* 2013;**13**:26 .

Morgan MJ, Ward RM, Brussell EM. The aftereffect of tracking eye movements. *Perception* 1976;**5**:309–17.

Movshon JA, Newsome WT. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *J Neurosci* 1996;**16**:7733–41.

Mukai I, Watanabe T. Differential effect of attention to translation and expansion on motion aftereffects (MAE). *Vision Res* 2001;**41**:1107–17.

Naka KI, Rushton WA S-potentials from luminosity units in the retina of fish (Cyprinidae). *The Journal of physiology* 1966;**185**, 587–599.

Nishida SY, Ashida H. A hierarchical structure of motion system revealed by interocular transfer of flicker motion aftereffects. *Vision Res* 2000;**40**:265–78.

Nishida SY, Ashida H, Sato T. Contrast dependencies of two types of motion aftereffect. *Vision Res* 1997;**37**:553–63.

Noel JP, Wallace M, Blake R. Cognitive neuroscience: integration of sight and sound outside of awareness? *Curr Biol* 2015;**25**: R157–9.

Palmer TD, Ramsey AK. The function of consciousness in multisensory integration. *Cognition* 2012;**125**:353–64.

Park M, Blake R, Kim Y *et al*. Congruent audio-visual stimulation during adaptation modulates the subsequently experienced visual motion aftereffect. *Sci Rep* 2019;**9**:19391.

Pavan A, Greenlee MW, Antal A. Effects of crowding and attention on high-levels of motion processing and motion adaptation. *PLoS One* 2015;**10**:e0117233.

Petersen SE, Baker JF, Allman JM. Direction-specific adaptation in area MT of the owl monkey. *Brain Res* 1985;**346**:146–50.

Plass J, Guzman-Martinez E, Ortega L *et al*. Lip reading without awareness. *Psychol Sci* 2014;**25**:1835–7.

Poirier C, Collignon O, Scheiber C *et al*. Auditory motion perception activates visual motion areas in early blind subjects. *Neuroimage* 2006;**31**:279–85.

Rezec A, Krekelberg B, Dobkins KR. Attention enhances adaptability: evidence from motion adaptation experiments. *Vision Res* 2004;**44**:3035–44.

Rezk M, Cattoir S, Battal C *et al*. Shared representation of visual and auditory motion directions in the human middle-temporal cortex. *Curr Biol* 2020;**30**:2289–99.

Rosemann S, Wefel IM, Elis V *et al*. Audio–visual interaction in visual motion detection: synchrony versus asynchrony. *J Optom* 2017;**10**:242–51.

Rousselet GA, Pernet CR, Wilcox RR. Beyond differences in means: robust graphical methods to compare two groups in neuroscience. *Eur J Neurosci* 2017;**46**:1738–48.

Sadaghiani S, Maier JX, Noppeney U. Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *J Neurosci* 2009;**29**:6490–9.

Saenz M, Lewis LB, Huth AG *et al*. Visual motion area MT+/V5 responds to auditory motion in human sight-recovery subjects. *J Neurosci* 2008;**28**:5141–8.

Sathian K and Ramachandran VS (eds.), *Multisensory Perception: From Laboratory to Clinic*. Cambridge, Massachusetts, USA: Academic Press, 2019.

Sekuler R, Sekuler AB, Lau R. Sound alters visual motion perception. *Nature* 1997;**385**:308.

Shams L, Beierholm UR. Causal inference in perception. *Trends Cogn Sci* 2010;**14**:425–32.

Spence C. Multisensory perception. In: Wixted JT and Serences J (eds.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, Vol. 2. New York, USA: John Wiley & Sons, 2018, 1–56.

Spence C, Sathian K. Chapter 11 - Audiovisual crossmodal correspondences: Behavioral consequences and neural underpinnings. In: Sathian K and Ramachandran V.S. (eds.), *Multisensory Perception: From Laboratory to Clinic*. Cambridge, Massachusetts, USA: Academic Press, 2020, 239–58.

Stein BE, London N, Wilkinson LK *et al*. Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. *J Cogn Neurosci* 1996;**8**:497–506.

Stocker AA, Simoncelli EP. Visual motion aftereffects arise from a cascade of two isomorphic adaptation mechanisms. *J Vis* 2009;**9**:1–14.

Swanston MT, Wade NJ. Motion over the retina and the motion aftereffect. *Perception* 1992;**21**:569–82.

Taya S, Adams WJ, Graf EW *et al*. The fate of task-irrelevant visual motion: perceptual load versus feature-based attention. *J Vis* 2009;**9**:12.1–10.

Tootell RB, Reppas JB, Dale AM *et al*. Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature* 1995;**375**:139–41.

Tsuchiya N, Koch C. Continuous flash suppression reduces negative afterimages. *Nat Neurosci* 2005;**8**:1096–101.

Tukey JW. *Exploratory Data Analysis*, Vol. **2**, London, UK: Pearson, 1977, 131–60.

Van Der Zwan R, Wenderoth P, Alais D. Reduction of a pattern-induced motion aftereffect by binocular rivalry suggests the involvement of extrastriate mechanisms. *Vis Neurosci* 1993;**10**:703–9.

Vinke LN, Bloem IM, Ling S. Saturating nonlinearities of contrast response in human visual cortex. *J Neurosci* 2022;**42**:1292–302.

Wade NJ. A selective history of the study of visual motion aftereffects. *Perception* 1994;**23**:1111–34.

Wallace MT, Woynaroski TG, Stevenson RA. Multisensory integration as a window into orderly and disrupted cognition and communication. *Annu Rev Psychol* 2020;**71**: 193–219.

Watkins S, Shams L, Josephs O *et al*. Activity in human V1 follows multisensory perception. *Neuroimage* 2007;**37**:572–8.

Watkins S, Shams L, Tanaka S *et al*. Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage* 2006;**31**:1247–56.

Wiesenfelder H, Blake R. The neural site of binocular rivalry relative to the analysis of motion in the human visual system. *J Neurosci* 1990;**10**:3880–8.

Wohlgemuth A. On the aftereffect of seen movement. *British Journal of Psychology (Monograph Suppl.)*, 1911;**1**:1–117.

Wolbers T, Zahorik P, Giudice NA. Decoding the direction of auditory motion in blind humans. *Neuroimage* 2011;**56**:681–7.

Yuval-Greenberg S, Heeger DJ. Continuous flash suppression modulates cortical activity in early visual cortex. *J Neurosci* 2013;**33**:9635–43.

Neuroscience of Consciousness, 2024, **2024(1)**, 1–14
DOI: https://doi.org/10.1093/nc/niad027

Research Article

Received 11 August 2023; Revised 5 December 2023; Accepted 27 December 2023