

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Hemispheric Asymmetry and Multisensory Learning Impact Visual Perceptual Selection

Permalink

<https://escholarship.org/uc/item/18782599>

Author

Piazza, Elise Ann

Publication Date

2015

Peer reviewed|Thesis/dissertation

Hemispheric Asymmetry and Multisensory Learning Impact Visual Perceptual Selection

By

Elise Ann Piazza

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Vision Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michael A. Silver, Co-Chair

Professor Martin S. Banks, Co-Chair

Professor David Whitney

Professor Frederic Theunissen

Spring 2015

Abstract

Hemispheric Asymmetry and Multisensory Learning Impact Visual Perceptual Selection

by

Elise Ann Piazza

Doctor of Philosophy in Vision Science

University of California, Berkeley

Professors Michael Silver and Martin Banks, Co-Chairs

When visual input is consistent with multiple perceptual interpretations (e.g., the Necker cube), these interpretations compete for conscious awareness. The process of determining which interpretation will be dominant at a given time is known as perceptual selection. We study this process using binocular rivalry, a bistable phenomenon in which incompatible images presented separately to the two eyes result in perceptual alternation between the two images over time.

In one study, we showed that a well-established asymmetry in spatial frequency processing between the brain's two hemispheres applies to perceptual selection. Specifically, a lower spatial frequency grating was more likely to be selected when it was presented in the left visual field (right hemisphere) than in the right visual field (left hemisphere), while a higher spatial frequency grating showed the opposite pattern of results. Surprisingly, this asymmetry persisted for the entire duration of a continuously presented pair of static rivaling stimuli (30 seconds), which is the first demonstration that hemispheric differences in spatial frequency processing continue long after stimulus onset.

In another study, we found that very recently formed audio-visual associations influence perceptual selection. Here, we used a brief (8-minute) crossmodal statistical learning paradigm to expose subjects to arbitrary, consistent pairings of images and sounds. In a subsequent binocular rivalry test, we found that a given image was more likely to be perceived when it was presented with a sound that had been consistently paired with it during exposure than when presented with previously unpaired sounds. Our results indicate that the audio-visual associations formed during the brief exposure period influenced subsequent visual competition and that this effect of learning was largely implicit, or unconscious.

Together, these results demonstrate an unconscious influence of complex perceptual mechanisms on perceptual selection, the process by which we form conscious representations of ambiguous sensory input.

Table of Contents

1. General Introduction.....	1
Persistent hemispheric differences in the perceptual selection of spatial frequencies	
2.1 Introduction.....	2
2.2 Materials and Methods.....	4
2.3 Results.....	5
2.4 Discussion.....	10
Rapid crossmodal learning biases perceptual selection	
3.1 Introduction.....	11
3.2 Materials and Methods.....	12
3.3 Results.....	17
3.4 Discussion.....	19
4. Conclusions and Future Directions.....	21
References.....	23

Acknowledgments

During my time at Berkeley, I have had the great privilege of receiving guidance from, collaborating with, and sharing the friendship of many people. Instead of sitting at one desk for six years, I have had the unique pleasure of absorbing the expertise of multiple labs. First, I would like to thank my primary advisors, Michael Silver and Marty Banks, for shaping me into a scientist and granting me exactly as much freedom and independence as I desired; I am extremely grateful for your generosity and guidance. I would also like to thank David Whitney for his creativity, enthusiasm, and willingness to share his time and resources with a student outside of his lab. I am lucky to have known and worked with the late David Wessel, who shared his joy, musical insights, signal processing knowledge, and the resources of CNMAT with me. Finally, I am grateful for the contributions of Frédéric Theunissen, both as a member of my thesis committee and a collaborator on my ongoing auditory perception work.

As a first-year rotation student, I benefited greatly from the generosity and wisdom of several graduate student mentors and collaborators: Allison Yamanashi-Leib, Emily Cooper, and Rachel Denison. In particular, I attribute several programming-related epiphanies to Rachel's infinite patience. I thank Dennis Levi and Lynn Robertson for serving on my qualifying committee and reassuring me throughout my exam preparations. Lynn's work also inspired my project on hemispheric asymmetries and I am grateful for her suggestions throughout the evolution of that study. I would also like to thank collaborators Sunah Kim, Timothy Sweeny and Valerie Morash for their contributions to my multisensory and auditory perception research.

Several invaluable research assistants and interns—Jacob Sheynin, Karen Wong, Aaron Bloch, Vyoma Shah, Aditya Challa, and Maxwell Schram—helped with stimulus design and contributed hundreds of hours of data collection. My many labmates from the Silver and Banks labs provided support, statistics and programming advice, and critical feedback on countless presentations. I would also like to thank lab manager Peter Illes and the administrative staff of the Vision Science and Psychology programs—Inez Bailey, Olga Lepilina, and Harumi Quinones—for their energy and help. The intellectual freedom I have enjoyed in graduate school is in part due to a generous fellowship from the Department of Defense. I am also grateful to Jennifer Hudin, Frédéric Theunissen, and my advisors for supporting my interest in teaching.

Many friends from the Vision Science program and beyond have surrounded me throughout graduate school. I will miss our weekends at Lake Tahoe and lunches on the music department lawn; our game nights, camping trips, modular origami attempts, and Friday drinks at Freehouse. I owe a good portion of my mental health to Valerie Morash and Miriam Bowring, the other members of the “A-List” trio; thanks for the music making and commiserating.

Finally, I would like to thank my family for constantly reminding me that the significance of one's life is not measured by a p-value.

Dedicated to my husband, Marius Cătălin Iordan, whose humor, encouragement, and scientific feedback propel me ever forward. Talking shop with you at the dinner table is truly a privilege.

Chapter 1

1. GENERAL INTRODUCTION

When visual input is consistent with multiple perceptual interpretations, the brain constructs a coherent representation of this ambiguous sensory information in order to make sense of the surrounding environment. For instance, bistable figures such as the Necker cube and Rubin's face/vase illusion generate competing perceptual interpretations that alternate over time. The study of perceptual selection—the process of determining which of multiple possible percepts will be dominant at a given time—provides important insights into the bases of conscious awareness. Binocular rivalry is a particularly intriguing bistable phenomenon that occurs when two incompatible images are presented separately to the two eyes at overlapping retinal locations, resulting in perceptual alternation between the images, even though the visual stimuli remain constant (reviewed in Blake & Wilson, 2011; Alais & Blake, 2005). This dissociation between unchanging stimuli and fluctuating perception of those stimuli provides a powerful tool for measuring the influence of various factors on conscious awareness.

Traditionally, low-level stimulus properties, such as luminance, contrast, and spatial frequency, have been primarily thought to regulate perceptual selection during binocular rivalry (Blake, 1989; Levelt, 1965). Within this framework, rivalry is often assumed to arise from competition between pools of early visual neurons, driven by these stimulus properties (Lehky, 1988; Blake, 1989). However, more recent work has illuminated higher-level influences on perceptual selection, such as stimulus configuration, emotional valence, interocular grouping, temporal context, attention, pharmacological interventions, and genetic factors (for a review, see Bressler, Denison, & Silver, 2013). These influences stem from neural and psychological phenomena tied to the observer, rather than resulting simply from low-level stimulus properties. Here, we tested the influence of two pervasive and complex phenomena on visual perceptual selection: differences in the perceptual specialization of the brain's two hemispheres (which results in asymmetries in spatial vision between the left and right visual hemifields) and probabilistic learning of associations between sounds and images.

Previous research showed that the right hemisphere processes low spatial frequencies more efficiently than the left hemisphere, which preferentially processes high spatial frequencies (see Ivry & Robertson, 1998, for a review). These studies typically measured response times to single, briefly flashed gratings and/or directed observers to attend to a particular spatial frequency immediately before making a judgment about a subsequently presented stimulus. Thus, it was previously unclear whether the hemispheres differ in perceptual selection from multiple spatial frequencies that are simultaneously present in the environment, without bias from selective attention. Moreover, the time course of hemispheric asymmetry in spatial frequency processing was unknown. In the first study, we addressed both of these questions using binocular rivalry. Subjects viewed a pair of rivalrous orthogonal gratings with different spatial frequencies, presented either to the left or right of central fixation, and continuously reported which grating they perceived. At the beginning of a trial, the low spatial frequency grating was perceptually selected more often when presented in the left hemifield (right hemisphere) than the right hemifield (left hemisphere), while the high spatial frequency grating showed the opposite pattern of results. This hemispheric asymmetry in perceptual selection persisted for the entire 30-second stimulus presentation, continuing long after stimulus onset.

These results indicate stable differences in the resolution of ambiguity across spatial locations and demonstrate the importance of considering sustained differences in perceptual selection across space when characterizing our conscious representations of complex scenes.

It is known that information from one sensory modality can help to resolve ambiguity in another modality. In the second study, we investigated the probabilistic mechanisms by which multisensory associations come to influence perception in a single sensory modality. We exposed subjects to consistent, arbitrary pairings of sounds and images and then measured the impact of this recent statistical learning on subjects' perceptual selection during binocular rivalry. On each trial of the rivalry test, subjects were presented with a pair of rivalrous images (one of which had been consistently paired with a sound during exposure, while the other had not) and an accompanying sound. We found that an image was significantly more likely to be perceived at the onset of rivalry when it was presented with its paired sound than when presented with other sounds. Across subjects, this effect of learning on rivalry did not significantly correlate with recognition memory for the individual audio-visual pairings, as assessed in a separate test, suggesting that the effects of exposure on rivalry were likely due to implicit learning. Our results indicate that recently acquired knowledge about multisensory relationships helps resolve sensory ambiguity, and they demonstrate that statistical learning is a fast, flexible mechanism that facilitates this process.

Chapter 2

2.1 INTRODUCTION

Hemispheric asymmetries in spatial frequency processing result from differences in perceptual specialization between the two hemispheres (Ivry & Robertson, 1998; Kitterle, Christman, & Hellige, 1990; Sergent, 1982). Within a stimulus set, identification and discrimination of low spatial frequencies (LSFs) tend to be faster and more accurate for stimuli presented in the left visual field (LVF), whereas high spatial frequencies (HSFs) are more quickly and accurately processed in the right visual field (RVF). This asymmetry has been observed for both sinusoidal gratings (Christman, 1997; Hellige, 1993; Christman, Kitterle, & Hellige, 1991) and spatial frequency-filtered natural scenes (Peyrin, Chauvin, Chokron, & Marendaz, 2003). Moreover, fMRI studies indicate that areas in the left hemisphere respond preferentially to HSF compared with LSF stimuli, whereas the opposite pattern was observed in the right hemisphere (Musel et al., 2013; Peyrin, Baciú, Segebarth, & Marendaz, 2004), and EEG responses are larger in the left compared with the right hemisphere for HSF stimuli and larger in the right than in the left hemisphere for LSF stimuli (Martínez, Di Russo, Anllo-Vento, & Hillyard, 2001).

Hemispheric asymmetry of spatial frequency processing is also known to be task dependent. For example, there are clear interactions between spatial frequency and hemisphere for spatial frequency discrimination (Proverbio, Zani, & Avella, 1997; Kitterle & Selig, 1991) but not for simple detection (Kitterle et al., 1990). However, the effects of hemispheric asymmetry on perceptual selection and conscious representations are unknown.

Ivry and Robertson (1998) introduced the Double Filtering by Frequency (DFF) theory to account for hemispheric asymmetries in spatial frequency processing. The DFF theory is based

on a two-stage model of spatial frequency filtering. In the first stage, a range of task-relevant frequencies is selected from the environment. In the second stage, frequencies within the selected range are asymmetrically filtered, with the right hemisphere processing relatively lower frequencies more efficiently and the left hemisphere preferentially processing relatively higher frequencies.

Previous behavioral studies of hemispheric asymmetries in spatial frequency processing have relied primarily upon measures of response times (RTs) to single stimuli briefly flashed in either the LVF or RVF. For example, Kitterle and Selig (1991) found that RTs for spatial frequency discrimination of two successively presented sinusoidal gratings were faster for lower spatial frequency gratings (1–2 cpd) in the LVF and for higher spatial frequency gratings (4–12 cpd) in the RVF. Although one of the primary predictions of the DFF theory involves selection from multiple spatial frequencies that are simultaneously present in the environment, there is little direct evidence for this aspect of the theory. One study (Kitterle, Hellige, & Christman, 1992) used gratings with multiple spatial frequency components (a low fundamental and higher harmonics) to compare selective processing of these components in the LVF versus RVF. However, in this study, participants were required to make a single perceptual judgment (either “Are the bars wide or narrow?” or “Are the bars sharp or fuzzy?”) based on a particular frequency component, so simultaneous perceptual processing of both low- and high-frequency components was never assessed within a given trial. Similarly, directing attention to one of two spatial frequency components in a grating while preparing to perform a local- or global-level discrimination of a subsequently presented Navon stimulus differentially modulated the amplitude of alpha band EEG signals in the two hemispheres (Flevaris, Bentin, & Robertson, 2011). In contrast to this previous work, our binocular rivalry study of competition between spatial frequencies provides a direct measure of perceptual selection without bias from selective attention to one of the spatial frequencies.

The DFF model (Ivry & Robertson, 1998) proposes two sequential stages of spatial frequency processing, but little is known regarding the temporal properties of the asymmetric processing that comprises the second stage. For example, it is unclear whether these hemispheric differences persist throughout the entire duration of a continuously presented stimulus or occur only during initial selection of that stimulus from the environment. Peyrin, Mermillod, Chokron, and Marendaz (2006) varied the presentation time of spatial frequency-filtered natural images (preceded by unfiltered natural images) and found the typical pattern of hemispheric asymmetry for brief (30 msec) stimulus durations as well as a right hemisphere advantage for longer (150 msec) exposure durations that was independent of spatial frequency. Although this study suggests that stimulus duration can influence asymmetric processing, it did not investigate the temporal dynamics of hemispheric processing throughout continuous stimulus presentation or for multiple simultaneously presented stimuli containing different spatial frequencies. In this study, we used binocular rivalry to continuously track perceptual selection of multiple spatial frequencies over time, thereby providing a direct measure of the time course of asymmetric processing in the two hemispheres. Our approach allows us to address two novel questions: first, does hemispheric asymmetry in spatial frequency filtering apply to perceptual selection, and second, how long does hemispheric asymmetry in spatial frequency processing persist after stimulus onset?

2.2 MATERIALS AND METHODS

Participants

Fourteen right-handed participants (aged 18–30 years, 11 women), including one of the authors (E.P.), completed this study. All participants provided informed consent, and all experimental protocols were approved by the Committee for the Protection of Human Subjects at the University of California, Berkeley. Each participant completed a 1-hr session composed of two blocks. We collected data from eighteen participants but excluded four participants' data sets from analysis. Of the excluded participants, one initially claimed to be right-handed but later notified the experimenter that he was born left-handed, one was unable to align the stereoscope to position the two monocular stimuli at corresponding retinal locations, and two were missing data from an entire condition because of incorrect response key mapping.

Stimuli

Binocular rivalry displays were generated on a Macintosh PowerPC using MATLAB and Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and were displayed on a gamma-corrected NEC Multi-Sync FE992 CRT monitor with a refresh rate of 60 Hz at a viewing distance of 100 cm. Participants viewed all stimuli through a mirror stereoscope with their heads stabilized by a chin rest. Stimuli were monochromatic circular patches of sine-wave grating 1.8° in diameter that were surrounded by a black annulus with a diameter of 2.6° and a thickness of 0.2° (Figure 1). Binocular presentation of this annulus allowed it to serve as a vergence cue to stabilize eye position. The stimuli were presented on the horizontal meridian, centered at 3.5° eccentricity either to the left or right of a black central fixation cross. Because the fixation crosses were in the same location on the screen in both hemifield conditions (Figure 1, top), participants' eye position, relative to the head, was the same in both conditions. All gratings were presented at 100% contrast and had the same mean luminance as the neutral gray background (59 cd/m^2). In each trial, one of the two sinusoidal gratings had a spatial frequency of 1 cycle per degree (LSF), and the other had a spatial frequency of 3 cycles per degree (HSF; Figure 1). The two gratings were orthogonal, with $\pm 45^\circ$ orientations. The spatial frequency and orientation of the grating presented to each eye were fully counterbalanced and randomly selected across trials.

Procedure

Before starting the experiment, each participant adjusted the stereoscope by rotating its mirrors until the two eyes' images (Figure 1, bottom, with the orthogonal gratings replaced by identical figures in both eyes for this adjustment phase) were fused and the participant could see only one cross and one annulus with binocular viewing. All participants completed five practice trials in each hemifield condition before starting the experiment to ensure that they were using the correct response keys and that the stereoscope was properly aligned.

In each trial, the static gratings, fixation cross, and annuli (Figure 1) were presented continuously for 30 sec with a 1500-msec blank interval (consisting of only the fixation cross and annuli) between trials. A brief (250 msec) pure tone auditory cue was presented immediately before the onset of the grating stimuli to signal the beginning of each trial. Throughout each trial, participants used two keys to indicate their percept: grating tilted to the left or grating tilted to the right. Participants were instructed to continuously press a key with their right hand for as long as the corresponding percept was dominant and to not press any key for ambiguous percepts. The experiment was separated into two blocked conditions: left hemifield and right

hemifield, the order of which was counterbalanced across participants. Twelve participants completed 28 trials per condition, and two participants completed 32 trials per condition.

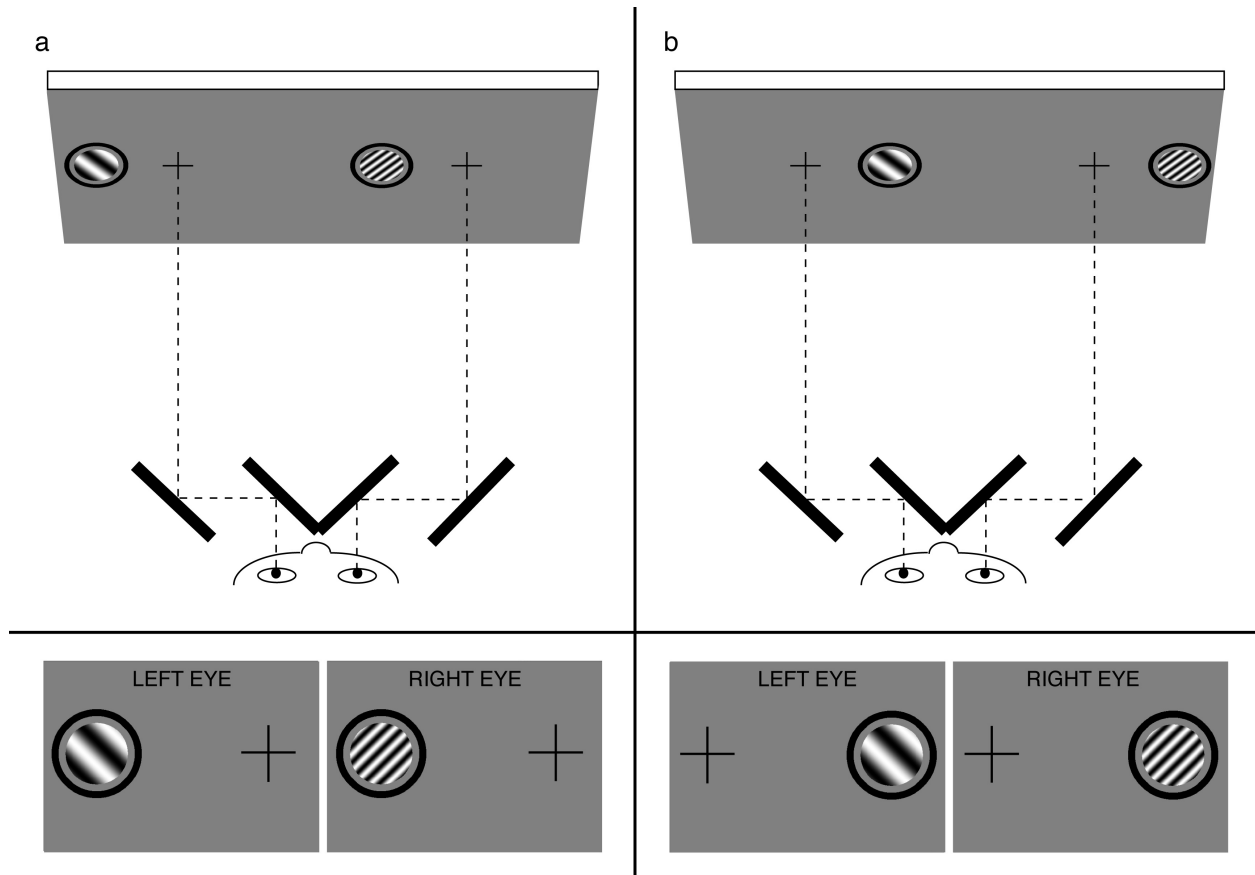


Figure 1. Top panels: schematic of an example visual display and mirror stereoscope in the (a) left visual hemifield condition and (b) right visual hemifield condition. Bottom panels: example images presented dichoptically to the two eyes. Subjects maintained fixation on a cross while continuously reporting their perception of rivalrous gratings that were presented to either the left (a) or the right (b) of the fixation cross.

2.3 RESULTS

Initial Response

At the beginning of a binocular rivalry trial, there is often a period in which participants experience an ambiguous percept (consisting of a patchwork or mixture of the two images), followed by a perceptual alternation between two distinct images. We defined the initial response as the first key press of a trial, as this indicates the participant's first percept that clearly corresponded to one of the two gratings. In our study, the average initial response latency (the time between stimulus presentation and the initial response) was 925 msec. Only 15% of the total trials across participants contained any response in the first 500 msec, indicating that participants typically did not have an unambiguous percept until several hundred milliseconds after stimulus onset.

We measured the proportion of initial responses on each trial corresponding to either the

LSF or HSF grating and found a highly significant Hemifield \times Spatial Frequency interaction across participants (ANOVA, $F(1, 13) = 8.23, p < .02, \eta^2_p = .39$; Figure 2A). More specifically, simple contrasts revealed that participants were more likely to initially perceive the lower spatial frequency in the LVF-RH (left visual field/right hemisphere condition) than in the RVF-LH, $F(1, 13) = 8.23, p < .02$, and they were more likely to initially see the higher spatial frequency in the RVF-LH than the LVF-RH, $F(1, 13) = 8.23, p < .02$ (Figure 2A). The results of these two contrasts are identical because the comparisons are complementary; all responses were either LSF or HSF, so the proportion of initial LSF responses = (1 – proportion of initial HSF responses).

There was also a highly significant main effect of Spatial Frequency, such that the LSF (1 cpd) grating was initially selected more often than the HSF (3 cpd) grating in both visual hemifields, $F(1, 13) = 66.11, p < .001, \eta^2_p = .84$ (Figure 2A). This main effect is not surprising, given existing knowledge regarding visual processing of different spatial frequencies. Specifically, LSF channels have shorter latencies and integration times than HSF channels (Breitmeyer, 1975), and LSF gratings interfere with orientation discrimination of HSF gratings more than HSF gratings interfere with discrimination of LSF gratings (Hughes, 1986). In addition, LSF stimuli evoke larger neural responses than HSF stimuli (Peyrin et al., 2004). These behavioral and physiological findings are consistent with our observation that initial perceptual selection was generally biased toward the LSF grating. However, this main effect of spatial frequency is orthogonal to our finding of an interaction of hemisphere and spatial frequency in perceptual selection.

Initial Response Duration

Because initial perceptual selection and maintenance of a binocular rivalry percept have been shown to result from separate mechanisms (de Weert, Snoeren, & Koning, 2005), we also measured the duration of the initial response in each trial and again observed a significant Hemifield \times Spatial Frequency interaction, $F(1, 13) = 7.24, p < .02, \eta^2_p = .36$ (Figure 2B). Specifically, simple contrasts showed that initial responses corresponding to the HSF grating were significantly longer in the RVF-LH than in the LVF-RH, $F(1, 13) = 5.33, p < .05$. However, initial response durations for responses corresponding to the LSF were not significantly different between the two hemifield conditions, $F(1, 13) = 1.67, p = .22$. In addition, initial responses in the LVF-RH condition were significantly longer for LSF than HSF gratings, $F(1, 13) = 15.02, p < .01$, and there was a trend for initial responses in the RVF-LH condition to be longer for the LSF than HSF, $F(1, 13) = 4.48, p = .054$. As we found for the initial response type, there was a significant main effect of Spatial Frequency on initial response duration, $F(1, 13) = 13.92, p < .01, \eta^2_p = .52$, with longer durations for LSF than HSF percepts. There was also a trend toward a main effect of Hemifield on initial response duration, $F(1, 13) = 3.82, p = .07, \eta^2_p = .23$.

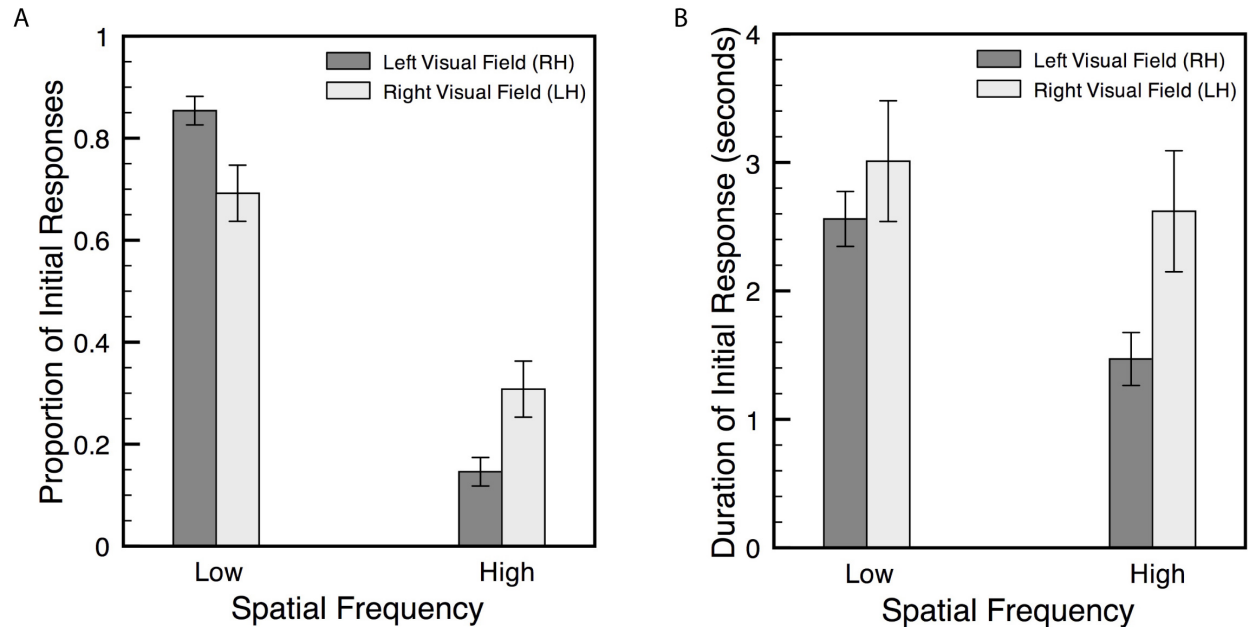


Figure 2. Interaction of spatial frequency and visual hemifield in the (a) proportion and (b) duration of initial responses to rivalrous stimuli. $N = 14$. Error bars are s.e.m. across subjects.

Time Course of Asymmetric Perceptual Selection

To characterize the persistence of hemispheric asymmetry in perceptual selection following stimulus onset, data from the 30-sec trials were divided into 60 time bins of 500 msec each (1–500 msec, 501–1000 msec, etc.). For each trial, we recorded the total amount of time in each bin that a participant responded either LSF or HSF. Each bin was then classified as “LSF” or “HSF” for that trial, based on a winner-take-all procedure. For each bin, we then computed the proportion of “LSF” trials (of those trials in which there was a response to at least one of the two gratings within that bin) for each participant and then averaged these proportion values across participants (Figure 3). Because all responses were either 1 cpd or 3 cpd, the proportion of “LSF” trials and “HSF” trials for each bin always added to 1, explaining the symmetric pattern of results within each hemifield condition (Figure 3).

The strong hemispheric asymmetry present in the first few time bins (Figure 3) (i.e., the relatively stronger bias toward the LSF in the right hemisphere than in the left hemisphere) was primarily due to the initial response (Figure 2A). On average, the termination of the initial response occurred in Bin 7 (3300 msec after stimulus onset). We therefore examined the persistence of hemispheric asymmetry in perceptual selection by separately analyzing the data for initial selection (proportion values collapsed across Bins 1–7; to the left of and including the dotted line in Figure 3) and sustained selection (collapsed across Bins 8–60; to the right of the dotted line in Figure 3).

For Bins 1–7, we found a significant Hemifield \times Spatial Frequency interaction, $F(1, 13) = 9.60$, $p < .01$, $\eta_p^2 = .43$, and a significant main effect of Spatial Frequency, $F(1, 13) = 62.43$, $p < .001$, $\eta_p^2 = .83$, confirming the pattern of results found for the initial response (Figure 2A). More specifically, simple contrasts again showed that participants were more likely to perceive the lower spatial frequency in the LVF-RH than in the RVF-LH, $F(1, 13) = 9.60$, $p < .01$, and they were more likely to perceive the higher spatial frequency in the RVF-LH than in the LVF-

RH, $F(1, 13) = 9.60, p < .01$. Again, the results of these two contrasts are identical because the two proportion values for each bin always add to 1.

For Bins 8–60, we also found a significant Hemifield \times Spatial Frequency interaction, $F(1, 13) = 5.47, p < .05, \eta^2_p = .30$, demonstrating hemispheric asymmetry in the perceptual selection of spatial frequencies that persisted well beyond the initial response. Specifically, simple contrasts revealed that participants were more likely to perceive the lower spatial frequency in the LVF-RH than in the RVF-LH, $F(1, 13) = 5.47, p < .05$, and they were more likely to perceive the higher spatial frequency in the RVF-LH than in the LVF-RH, $F(1, 13) = 5.47, p < .05$. There was also a significant main effect of Spatial Frequency in Bins 8–60, $F(1, 13) = 6.66, p < .05, \eta^2_p = .34$, but this was once again orthogonal to the Hemifield \times Spatial Frequency interaction. These results demonstrate that spatial frequency selection differs between the two hemispheres both during the initial response and throughout the remainder of the stimulus presentation.

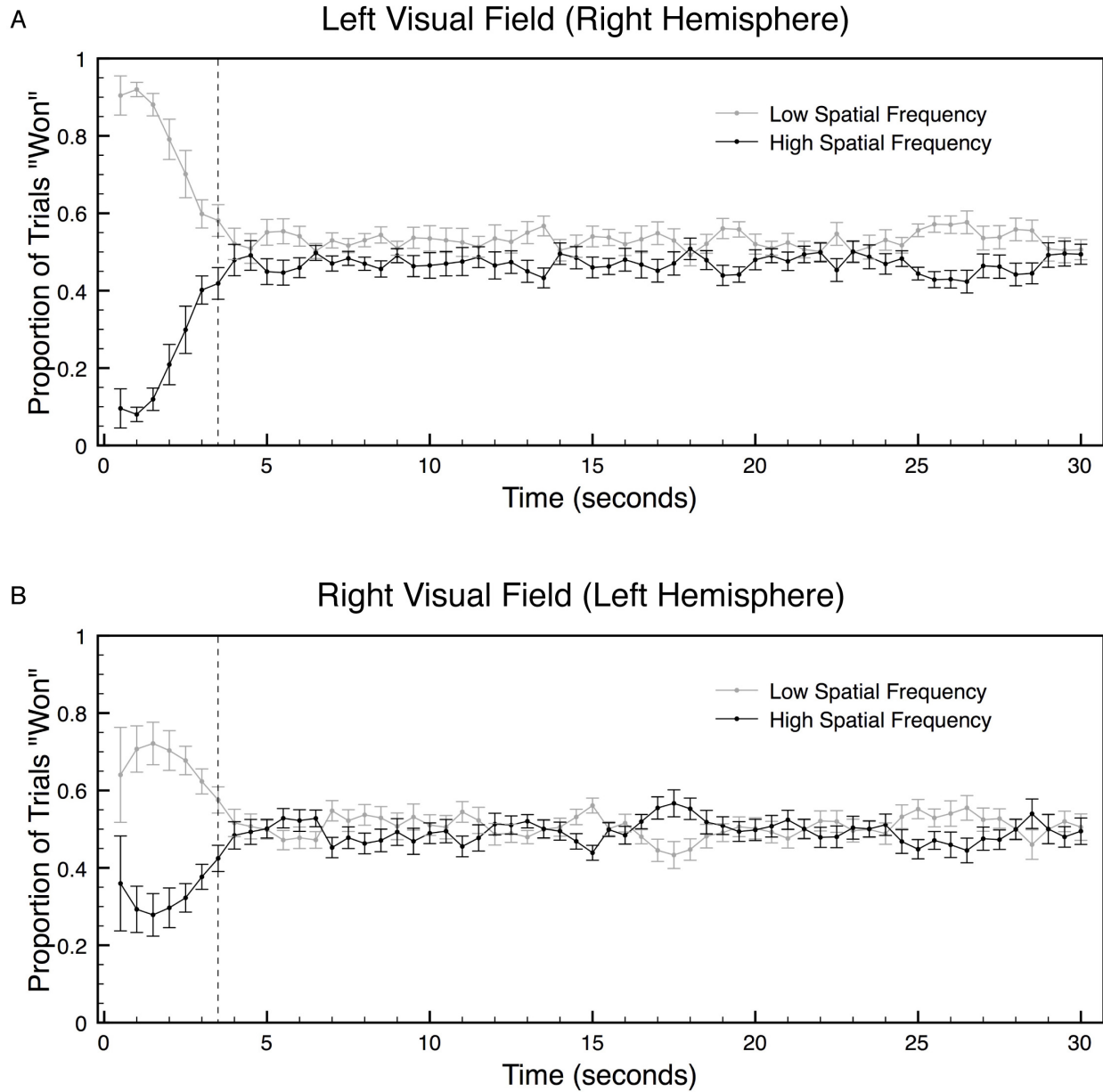


Figure 3. Time course of asymmetric selection of spatial frequencies. Each trial was divided into 500-ms bins; abscissa values correspond to the time at the end of each bin. Using a winner-take-all procedure (see Results), we determined which grating (LSF or HSF) was more dominant in each bin for each trial and plotted the proportion of “won” trials for each spatial frequency in each bin in the (a) left visual field and (b) right visual field conditions. Because all responses were either 1 cpd or 3 cpd, the two proportion values for each bin always add to 1, thereby explaining the symmetric pattern of results within each hemifield condition. $N = 14$. The dotted lines indicate the bin in which the initial response terminated, on average. Error bars are s.e.m. across subjects.

2.4 DISCUSSION

Our study provides the first evidence of hemispheric differences in perceptual selection of spatial frequencies. Specifically, we found a significant interaction between hemifield and spatial frequency that was in the direction predicted by the DFF theory, such that participants initially perceived the LSF more often in the LVF (right hemisphere) than the RVF (left hemisphere) and initially perceived the HSF more often in the RVF (left hemisphere) than the LVF (right hemisphere). In addition, we have shown that visual hemifield differences in perceptual selection persist throughout the entire duration of stimulus presentation.

An interaction between spatial frequency and visual hemifield has previously been reported for spatial frequency discrimination (Proverbio et al., 1997; Kitterle & Selig, 1991) but was not found for contrast sensitivity or visible persistence of gratings (Peterzell, Harvey, & Hardyck, 1989) or for spatial frequency-dependent detection of gratings (Kitterle et al., 1990). Our findings indicate that the two hemispheres differ in their perceptual selection of spatial frequencies from ambiguous stimuli, each exhibiting a bias toward selecting a visual interpretation that is most consistent with its relative perceptual specialization. This result has important implications regarding the factors that influence conscious perception, as it suggests that distinct types of information are selected at different locations within a visual scene.

Our finding that hemispheric differences persist throughout the entire 30-sec duration of stimulus presentation is the first demonstration that hemispheric asymmetry in spatial frequency processing is not simply a transient mechanism for filtering spatial frequency information. On the basis of previous findings using briefly presented gratings, it was unknown whether differences in preferential processing of frequencies between the two hemispheres occurred only for initial exposure to an image or whether they persisted for the entire duration of presentation. Our results clearly demonstrate that the right hemisphere has a significantly stronger bias toward LSF stimuli than the left hemisphere does and that this asymmetry persists throughout the entire 30-sec trial. This suggests that the right hemisphere plays an important role in continually selecting LSFs (which provide crucial information about global structure and coarse features) from the visual environment, whereas the left hemisphere divides its resources relatively more equally among multiple spatial frequencies.

Robertson (1996) reported spatial frequency-based, sequential priming effects, in the context of an “attentional print” model, that may be related to the persistent biases we report for perceptual selection of spatial frequency. In this previous work, participants judged whether one of two letter targets was present in a Navon stimulus, with the target appearing randomly at either the global (lower spatial frequencies) or local (higher spatial frequencies) level. Level-specific priming (i.e., faster RTs when the local/global level of the target was the same as in the previous trial) occurred regardless of changes across consecutive trials in features that were independent of spatial frequency (e.g., target letter identity, stimulus location, color) and persisted up to the longest intertrial interval that was tested (3 sec). However, when spatial frequency information was altered from one trial to the next, level-based priming was eliminated. This pattern of results was explained by an attentional print that contains information about recently relevant spatial frequencies (i.e., the previous trial). In our study, the initial period of dominance could reflect intrinsic hemispheric biases in spatial frequency selection, resulting in the generation of an attentional print that is maintained throughout the remainder of the trial and influences ongoing perceptual selection.

The hemispheric asymmetry in perceptual selection that we have characterized using

binocular rivalry should be distinguished from the interhemispheric switch framework (Pettigrew, 2001; Miller et al., 2000), which postulates that perceptual alternation during binocular rivalry is controlled by midbrain structures. This model is based on results from experimental manipulation of activity in one hemisphere, either through caloric vestibular stimulation or TMS, and proposes that perceptual switches in rivalry are driven by a bistable oscillator circuit that alternately activates each hemisphere. In contrast, our results reflect stable hemispheric differences in processing LSFs versus HSFs.

Chapter 3

3.1 INTRODUCTION

To accurately perceive the natural world, we must learn that cues from different sensory modalities point to the same object: a friend's face and her voice, a rose and its fragrance, a blackberry and its flavor. Through repeated exposure to consistent couplings between sensory features, we learn to associate information from various senses. Once these associations are well established, they can help resolve information that is ambiguous or impoverished for one of the senses (Conrad, Bartels, Kleiner, & Noppeney, 2010; Lunghi, Binda, & Morrone, 2010). However, the mechanisms for forming multisensory associations that then influence the contents of conscious perception are unclear. One possibility is that long-term exposure to multisensory couplings is required, potentially over extended periods of development. Alternatively, the brain may have more flexible mechanisms for constraining perceptual interpretations that make use of newly encountered multisensory couplings.

The stimulus context in which the rivalrous images are presented can bias perceptual selection (the brain's choice of the image that is perceived at any given time), often resulting in increased dominance of rivalrous images that are compatible or congruent with the established context (Bressler, Denison, & Silver, 2013). Of particular relevance to the current study, recent work has shown that perceptual selection during rivalry is affected both by predictive visual context and by input from other sensory modalities. For instance, exposure to a sequence of rotating gratings just before presentation of a rivalrous pair of static orthogonal gratings enhances initial perceptual selection of the grating that is consistent with the expected motion trajectory established by the rotating sequence (Denison, Piazza, & Silver, 2011). Concurrent sounds also affect binocular rivalry: a sound that is congruent with one of two rivalrous images in flicker frequency (Kang & Blake, 2005), motion direction (Conrad, Bartels, Kleiner, & Noppeney, 2010), or semantic content (Chen, Yeh, & Spence, 2011) enhances predominance of the rivalrous congruent image over the incongruent image. Similarly, smelling an odor that is semantically congruent with one of two competing images (Zhou, Jiang, He, & Chen, 2010), or haptically exploring a grooved stimulus of the same orientation as one of the rivaling images (Lunghi, Binda, & Morrone, 2010), increases that image's dominance. In these examples, the congruencies that influence predominance during binocular rivalry are explicit and learned through a lifetime of experience with rotating motion trajectories and crossmodal associations. However, it is unknown whether very recently learned crossmodal associations can also influence visual perceptual selection.

It is clear that associations between stored representations of sensory cues can be established quickly through probabilistic inference (Aslin & Newport, 2012). Both adult (Fiser & Aslin, 2001) and infant (Saffran, Aslin, & Newport, 1996) observers can rapidly learn sequential or spatial patterns of sensory stimuli (e.g., abstract shapes, natural images, or sounds) through passive exposure. This phenomenon, known as statistical learning, is thought to be crucial for detecting and forming internal representations of regularities in the environment. In the exposure phase of a typical statistical learning experiment, subjects are passively exposed to long streams of stimuli that contain sequences of two or more items that always appear in the same temporal order (Saffran, Johnson, Aslin, & Newport, 1999; Aslin & Newport, 2012). Afterwards, when asked to judge which of two sequences is more familiar, subjects are more likely to choose sequences that were presented during the exposure phase than random sequences of stimuli that had not previously been presented in that order. Statistical learning is often thought to be implicit, since subjects typically report being unaware of the identity of particular sequences from the exposure phase, even though they perform above chance on forced choice familiarity judgments.

Statistical learning also occurs for multisensory sequences: following exposure to streams of bimodal quartets (each containing two consecutive audio-visual, or AV, pairs), subjects performed above chance on familiarity judgments for crossmodal (AV) as well as unimodal (AA and VV) associations (Seitz, Kim, van Wassenhove, & Shams, 2007). Statistical learning is therefore a fast, flexible associative mechanism that could constrain perceptual interpretations in one sensory modality based on information in another modality. However, whether multisensory statistical learning can influence perception, as opposed to only recognition memory, has not been investigated.

To address this question, we examined the influence of statistical learning of arbitrary auditory-visual associations on subsequent visual perceptual selection. We asked whether perceptual interpretations during binocular rivalry could be rapidly updated based on recent multisensory experience, or whether learners might instead require days or even years of exposure to joint probabilities between sounds and images before these congruencies begin to influence perceptual selection. Specifically, we tested whether formation of associations between particular sounds and images during an 8-minute exposure phase would cause the presentation of a sound to alter perceptual dominance of its associated image during subsequent binocular rivalry. Discovering this type of impact of crossmodal statistical learning on binocular rivalry would indicate that recently acquired probabilistic information about the conjunctions of sounds and images, likely represented in multisensory areas and/or higher-order cortical areas, influences the resolution of conflicting visual information about simple stimuli that are likely represented in low-level visual areas.

3.2 MATERIALS AND METHODS

Participants

Twenty participants (ages 18-39, 14 female) completed this study. This sample size is comparable to those of previous studies on statistical learning (Fiser & Aslin, 2001) and on the influence of predictive information on binocular rivalry (Denison, Piazza, & Silver, 2011). Regarding statistical power, the latter study (specifically, Experiment 3) is most relevant to the present study and reported an effect size of Cohen's $d = 0.76$ for the effect of a predictive image sequence on the proportion of initial responses in binocular rivalry. The sample size of twenty

employed in the present study results in an estimated power of 0.88 to detect an effect size equivalent to that found in Experiment 3 of Denison, Piazza, & Silver (2011). All subjects provided informed consent, and all experimental protocols were approved by the Committee for the Protection of Human Subjects at the University of California, Berkeley. We originally collected data from 30 participants but excluded 10 subject data sets from analysis. Of the excluded subjects, one was unable to align the stereoscope to position the two monocular stimuli at corresponding retinal locations, two were missing data due to incorrect response key mapping, and seven were excluded for having incorrect responses on more than 1/3 of the catch trials (see Procedure).

Stimuli

All visual stimulus displays were generated on a Macintosh PowerPC using MATLAB and Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and were displayed on a gamma-corrected NEC MultiSync FE992 CRT monitor with a refresh rate of 60 Hz at a viewing distance of 100 cm. All images were presented at the fovea and had 100% contrast and the same mean luminance as the neutral gray screen background (59 cd/m²). Each participant's head position was fixed with a chin rest throughout the entire experimental session.

The stimulus set included six images and six sounds (Figure 1A). The stimuli were designed to be simple, easily discriminable from one another within a given modality, and unfamiliar to the subjects. The images were selected so that they could be grouped into three pairs (orthogonal sinewave gratings with +/- 45° orientations; two hyperbolic gratings with 0 and 45° orientations; a polar radial grating and a polar concentric grating), such that the members of each pair would rival well with each other (Figure 2). We refer to these image pairs as "rivalry pairs". During the rivalry test (see Procedure), the two images making up a given rivalry pair were always presented together, one in each eye. The sounds included two sinewave pure tones (D5, B6) and four chords composed of sinewave pure tones (two distinct dissonant clusters, an A-flat major chord, and an F-minor chord). The sounds were presented through headphones at a comfortable level.

Procedure

Each subject completed a 1.5-hour session composed of three phases: exposure, recognition test, and rivalry test.

Exposure phase: Each participant passively viewed and heard an 8-minute stream of sounds and images. All images were presented in the center of the screen during this phase. Participants were instructed to attend to the stimuli and fixate the images but were not required to perform any task, and the experimenters did not disclose the existence of any patterns in the AV (audio-visual) streams. Within this continuous stream (Figure 1B), each sound was presented for 500 ms and then continued playing while an image was presented for another 500 ms, after which the sound and image presentations ended simultaneously (Figure 1C). This was followed by a 500-ms blank interval before the onset of the next sound. Each participant was exposed to a total of 180 of these AV presentations.

For every participant, three images (one from each rivalry pair) and three sounds were randomly chosen to be consistently paired during the exposure phase (Figure 1A). Each of these three selected sounds was always presented with its paired image, corresponding to a total of three "AV pairs". Ninety of the 180 AV presentations (frames with dotted borders, Figure 1B)

during the exposure phase corresponded to one of these consistent AV pairs, for a total of 30 identical presentations of each pair. In the other 90 AV presentations (frames with solid borders, Figure 1B), random combinations of the remaining, unpaired, images and sounds (three of each) were presented, so there was no consistent mapping between any of these images and any sound. Selection of the images and sounds for pairing was counterbalanced across subjects to eliminate possible bias due to any inherent congruence between the stimuli; thus, across the group, AV pairings were fully randomized and arbitrary.

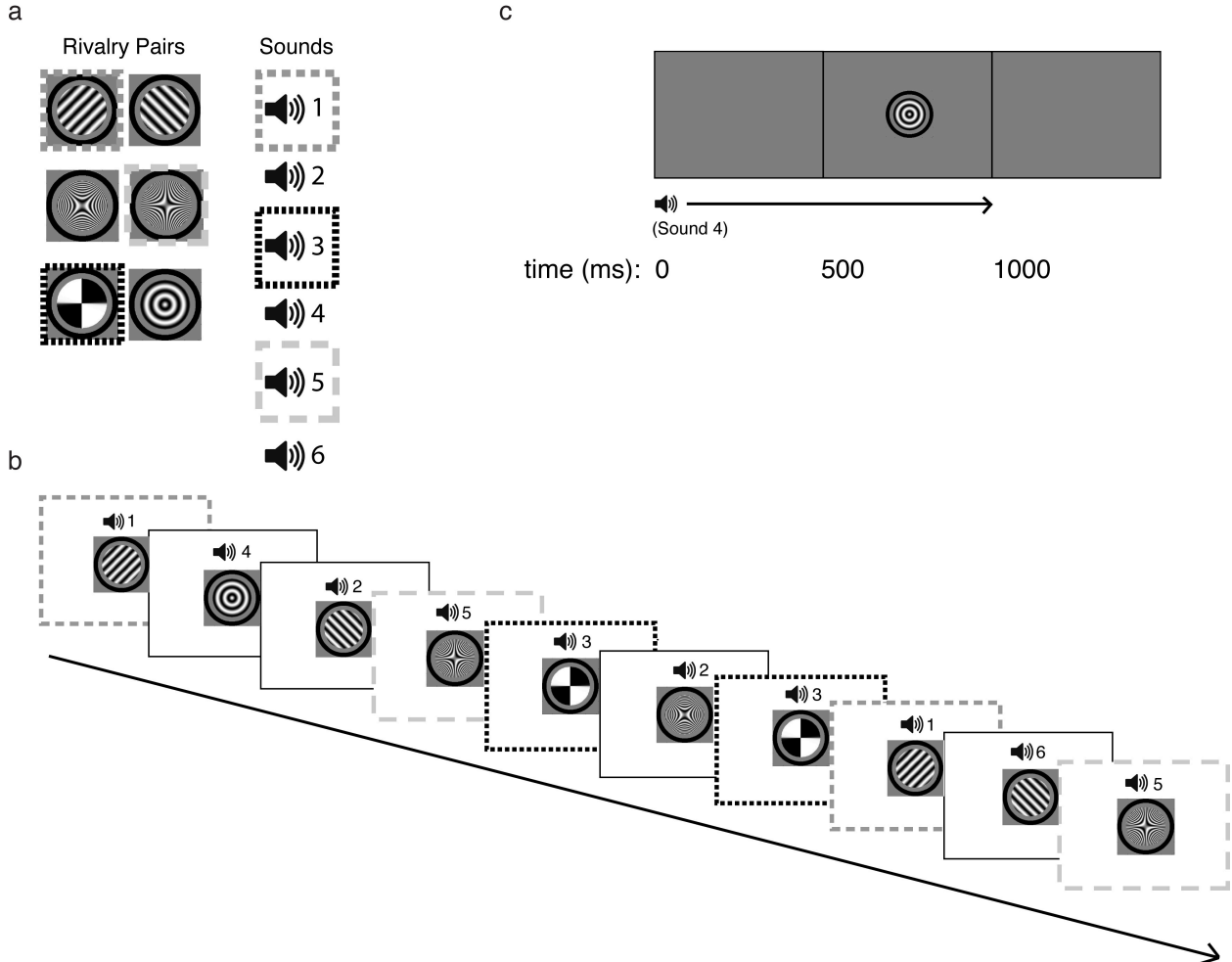


Figure 1. Exposure phase. (a) The full set of images and sounds. Borders indicate paired stimuli for an example subject. For each subject, one image from each of the three rivalry pairs was randomly chosen to be a “paired image”, which means that it was consistently paired with a particular sound during the exposure phase (an “AV pair”). For each paired image, the associated sound was randomly chosen from the set of six sounds. The remaining three sounds were left unpaired, meaning they could be presented with any of the non-paired images on different individual AV presentations. (b) Example AV presentations during the exposure phase. (c) The time course of a single example AV presentation.

Recognition test: To test participants’ knowledge of the AV pairs that were presented during exposure, one of the six sounds was presented, followed immediately by all six images, randomly arranged in a row on the screen. Participants were asked to choose the image that had

typically been presented with that sound during the exposure phase. Each participant completed 36 trials, with six repetitions of each of the six sounds, for a total of approximately 5 minutes.

Rivalry test: Subjects viewed pairs of images (the rivalry pairs described above, presented separately to the two eyes) through a mirror stereoscope, with their heads stabilized by a chin rest. Each image was a circular patch 1.8° in diameter, surrounded by a black annulus with a diameter of 2.6° and a thickness of 0.2° . Binocular presentation of this annulus allowed it to serve as a vergence cue to stabilize eye position. The two images in a rivalry pair were tinted red and blue during the rivalry test, thereby allowing participants to report their perception during binocular rivalry using only two (“red” or “blue”) instead of six (each of the images) response categories. The color and member of the rivalry pair presented to each eye were fully counterbalanced and randomly intermixed across trials. The use of colored image tints also served to increase the exclusivity of rivalry (decrease piecemeal percepts), and employing color as a response variable is standard in binocular rivalry research (Alais & Blake, 2005).

Before starting the rivalry test, each subject adjusted the stereoscope mirrors until the two eyes’ images (with gratings replaced by identical figures in both eyes for this adjustment phase) were fused and the subject could only see one annulus with binocular viewing. All subjects completed ten practice trials before starting the test to ensure that they were using the correct response keys and that the stereoscope was properly aligned.

In each trial of the rivalry test, one of the six (randomly selected) sounds was presented, followed by the static visual images (Figure 2). The timing of the onset of stimulus presentation was the same as in the exposure phase, but here the images were presented continuously for 10 seconds instead of 500 ms (persisting for 9.5 seconds beyond the termination of the sound). There was a 1000-ms blank interval (consisting of only the binocular annulus) between trials. Throughout each trial, subjects could press one of two keys to indicate their percept: either the red- or blue-tinted image. Subjects were instructed to continuously press a key for as long as the corresponding percept was dominant and not to press any key for ambiguous percepts.

Our analysis of the rivalry data focused on the initial response (the first reported percept in a given trial) because the sounds were presented at the beginning of the trial and effects of prediction on the initial response have been previously demonstrated (Denison, Piazza, & Silver, 2011). For each subject, we computed the proportion of trials for each rivalry pair in which the subject’s initial response (Figure 2) corresponded to the paired image (as opposed to the other rivaling image) for two types of trials: 1) when the paired sound was presented, and 2) when one of the other five sounds was presented (Figure 3A). We then calculated the within-subject difference between these proportion values to measure the effect of a concurrently presented paired sound on perception during rivalry. This normalization controlled for possible baseline differences in dominance between the images in a rivalry pair due to physical stimulus factors. A difference score above zero indicates that when a given image was accompanied by its paired sound from the exposure phase, it was more likely to be initially perceived during binocular rivalry, compared to when it was accompanied by other sounds. Each participant contributed one difference score, which quantifies the effect of crossmodal learning on rivalry, for each of the three rivalry pairs (Figure 3A).

We collected a total of 216 rivalry trials, divided across three blocks and lasting 45 minutes, from each participant. Because the rivalry test exposed subjects to many combinations of stimuli that violated the AV pairings established during the 8-minute exposure phase, we expected this test to interfere with statistical learning, causing the strength of the effect of

crossmodal learning on rivalry to decline over the course of the rivalry blocks. Indeed, this effect (for combined data across the three rivalry pairs) was only significant in the first of the three blocks (two-tailed t-tests; first block, $t(19) = 3.18, p < .01$; second block, $t(19) = -.27, p = .79$; third block, $t(19) = 1.07, p = .30$). Therefore, all analyses of rivalry test data were conducted on only the first block, which contained 72 rivalry trials (12 trials per sound) and 12 randomly interleaved catch trials for which the images were identical in both eyes, for a total of approximately 15 minutes of testing. Catch trials were considered to be incorrect if they contained any responses that did not correspond to the presented image. These trials were included to ensure that participants were attending to the stimuli and could distinguish the images based on color tint.

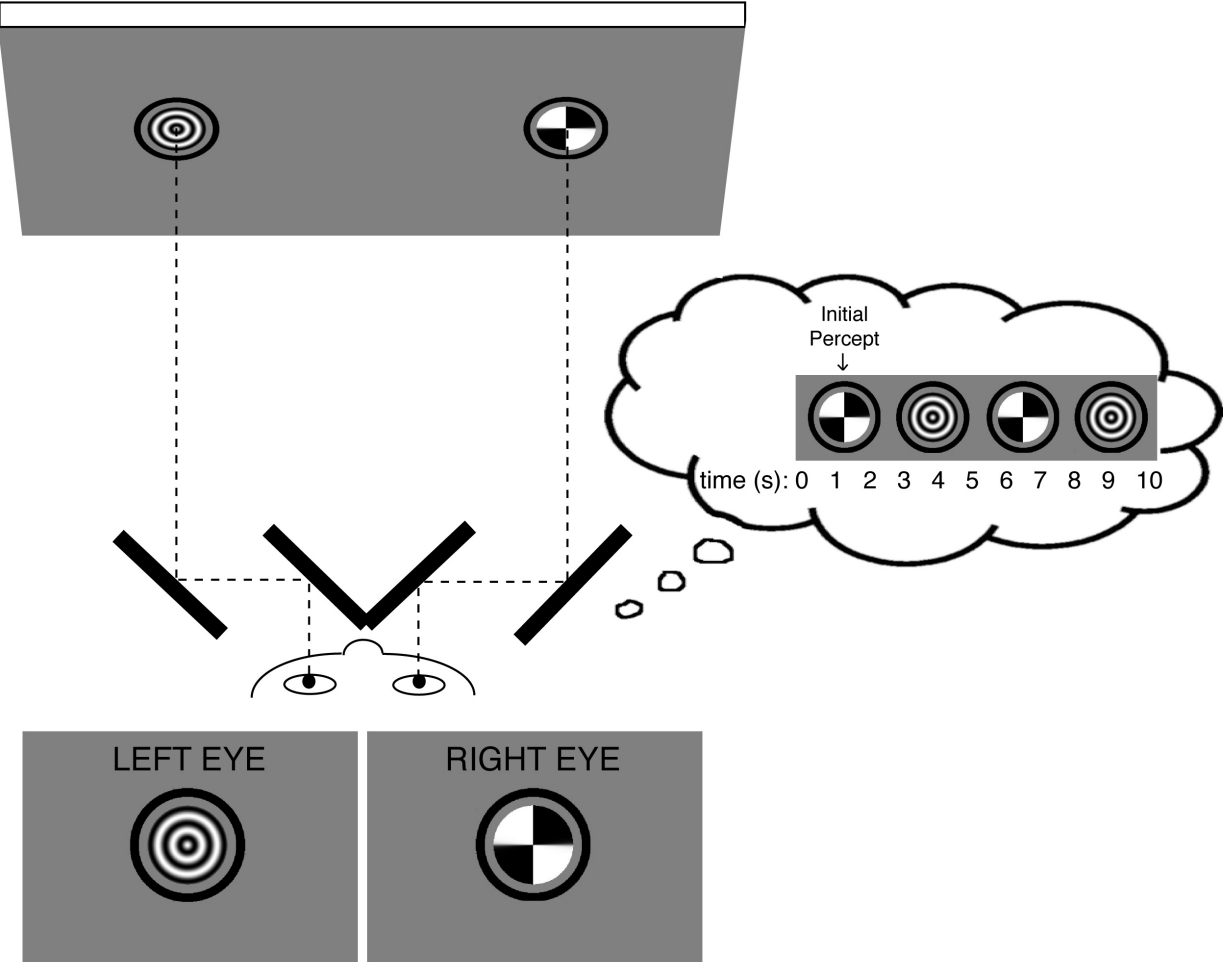


Figure 2. Example images presented separately to the two eyes (at corresponding retinal locations) through a mirror stereoscope and schematic sequence of possible percepts in a given trial of the rivalry test. Although the presented images did not change during a given trial, subjects typically perceived a continuous alternation between the two images throughout the 10-second stimulus presentation.

3.3 RESULTS

After exposing participants to streams of sounds and images (with half of the sounds and images consistently paired with each other), we measured the impact of exposure to the associated AV pairs on perceptual selection during binocular rivalry. We also measured the amount of explicit learning of the pairings in a separate recognition test. All AV pairings were randomly chosen; therefore, any effects of crossmodal associations on rivalry were due to learning that occurred during the 8-minute exposure phase.

Rivalry Test

Each rivalry pair had two images: the paired image and the unpaired image (see Methods and Figure 1). During the exposure phase, the paired image was consistently presented with a particular sound, while the unpaired image could be presented with any of the three non-paired sounds on different individual AV presentations. In the rivalry test, we presented all possible combinations of sounds (six total) and rivalry pairs (three total).

We computed the effect of crossmodal learning on rivalry (see Methods and Figure 3a), a measure of how much the initial dominance of an image is enhanced when it is presented with its paired sound, relative to when it is presented with an unpaired sound. The effect of crossmodal learning on rivalry responses was significantly greater than zero (two-tailed t-test; $t(19) = 3.18$, $p < .01$, Cohen's $d = .71$, 95% CI [0.03, 0.11]) (Figure 3a), indicating that exposure to arbitrary AV pairs influenced subsequent perceptual selection during binocular rivalry. We did not find a significant effect of the type of rivalry pair (i.e., sinewave gratings vs. hyperbolic gratings vs. bullseye/radial) on the size of the AV learning effect on rivalry (ANOVA, $F(2, 18) = .75$, $p = .47$, $\eta^2_p = .04$).

Recognition Memory Test

To test participants' recognition memory of the AV pairings presented during exposure, we measured the number of trials (out of 6 for each of the three AV pairs) in which the subject correctly identified the image that was paired with the presented sound during the exposure phase. The recognition test was conducted immediately following the end of the exposure phase in order to maximize the probability of detecting explicit learning, before any possible contamination of the learned associations by the rivalry test (in which all possible combinations of sounds and images were presented for an extended period of time). This was a conservative choice, as the resulting time delay and additional stimulus presentations between the exposure phase and the rivalry test may have weakened the effects of learning on rivalry. Recognition performance was significantly above chance (two-tailed t-test; $t(19) = 9.03$, $p < .0001$, Cohen's $d = 2.02$, 95% CI [.41, .66]) (Figure 3B). This indicates that, on average, participants learned the AV pairings during the exposure phase. There was no significant effect of the type of rivalry pair (ANOVA, $F(2, 18) = 2.80$, $p = .08$, $\eta^2_p = .13$) on recognition performance (i.e., no significant difference between the three rivalry pair categories).

However, many subjects' recognition test scores were numerically above chance (more than one correct response out of six, or 16.7%) but not perfect (six out of six) for a given AV pair (Figure 4). Specifically, out of the sixty total scores on the recognition test (twenty subjects, each contributing one score per AV pair), 21/60 (about one-third) were above chance but not perfect. Another one-third of the total responses (21/60) were at or below chance (zero or one correct response out of six). Slightly less than one-third (18/60) were perfect (six out of six). Overall,

this pattern indicates that for a given AV pair, some subjects had not explicitly learned the pairing but were still able to guess the correct answer well enough to exceed chance performance on our test, suggesting that this test may have been measuring a combination of implicit and explicit learning.

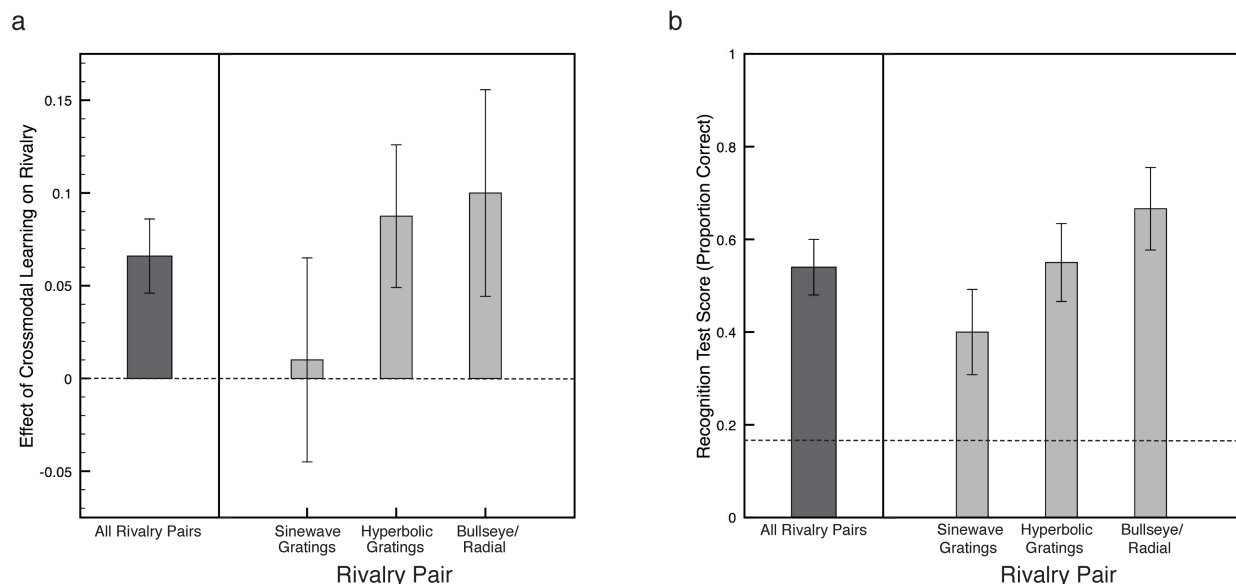


Figure 3. Effects of statistical learning of AV pairs on subsequent binocular rivalry and forced-choice recognition. (a) For each subject, the proportion of trials in which a given image was initially selected following the onset of binocular rivalry was measured in two conditions: in trials in which that image was presented with its paired sound, and in trials in which it was presented with an unpaired sound (averaged across all five unpaired sounds). The difference between these two scores quantifies the influence of crossmodal statistical learning on initial perceptual selection in binocular rivalry. The mean difference value across the three image pairs was calculated for each subject, and the resulting average across subjects is shown in “All Rivalry Pairs”. (b) On each trial of the recognition test, a single sound was presented, followed by a simultaneous display of all six images. Subjects were instructed to select the image that they thought had been paired with the presented sound during the exposure phase. Dotted line indicates chance performance (16.7%). N = 20. Error bars are s.e.m. across subjects.

Relationship Between Explicit Learning and Rivalry

If explicit knowledge of the individual AV pairings were driving the effect of crossmodal learning on rivalry, those subjects with the most explicit learning would be expected to have the largest effects of statistical learning on binocular rivalry. We used the recognition test scores to examine this possibility, even though it is likely that our recognition memory test measured some combination of explicit and implicit learning. To assess the relationship between recognition memory and rivalry, we computed a correlation across individual subjects between the recognition test score and the effect of crossmodal learning on rivalry for each rivalry pair (Figure 4). None of the three correlations was significant (sinewave gratings, $r(18) = .24, p = .31$, 95% CI [-.23, .62]; hyperbolic gratings, $r(18) = .25, p = .29$, 95% CI [-.22, .62]; bullseye/radial, $r(18) = .22, p = .35$, 95% CI [-.25, .60]). The correlation across subjects between recognition test

score and the effect of learning on rivalry for combined data from the three rivalry pairs was also not significant ($r(18) = .37, p = .11, 95\% \text{ CI } [-.08, .70]$). This lack of a clear relationship between recognition test and rivalry test scores across subjects suggests that there was little or no impact of explicit knowledge on the effects of statistical learning on binocular rivalry. It is thus likely that implicit statistical learning—rather than, for example, a response bias due to explicit knowledge of the AV pairs—was the dominant source of the effect of crossmodal learning on rivalry.

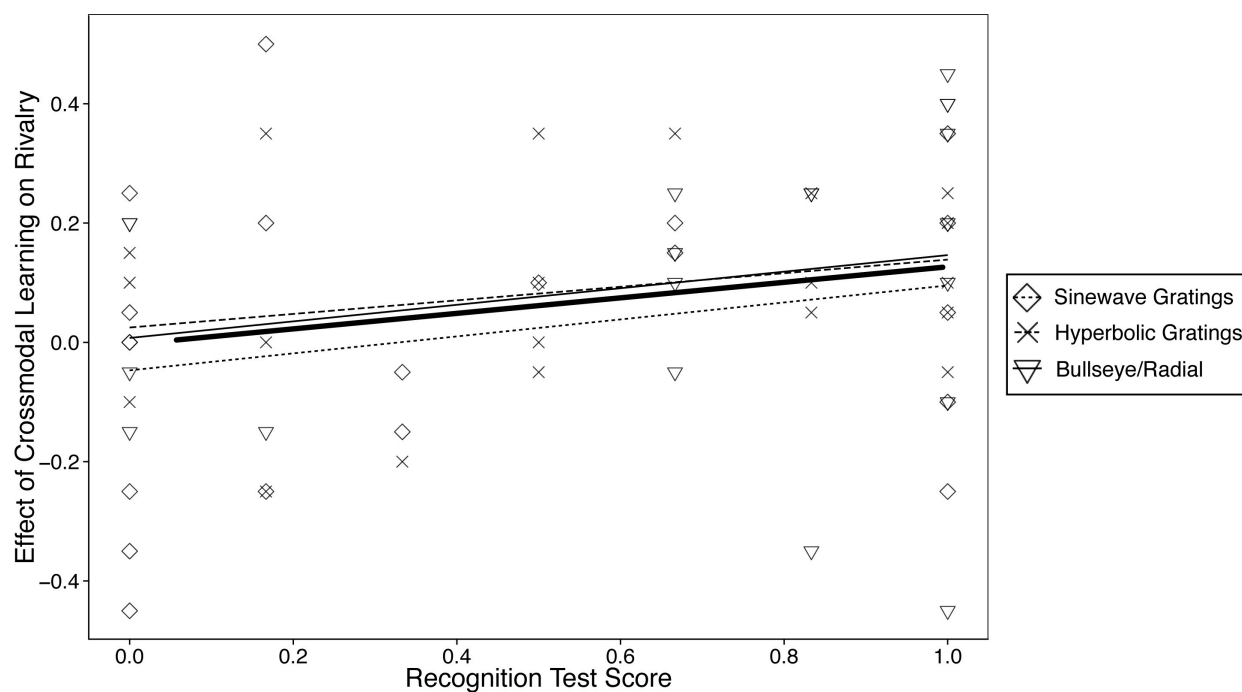


Figure 4. Relationship between recognition test score and the effect of crossmodal learning on binocular rivalry. Each data point is from a single subject for one of the three rivalry pairs. Lines indicate linear regression for one of the rivalry pairs (thin lines) or combined data from the three rivalry pairs (thick line). No significant correlations were observed. $N = 20$.

3.2 DISCUSSION

Here, we provide the first demonstration that recently formed associations between sounds and images impact what we see when the visual environment is ambiguous. Specifically, we found that a given image was more likely to be initially perceptually selected during binocular rivalry when it was preceded (and accompanied) by its paired sound than by non-paired sounds, indicating that crossmodal statistical learning influenced which of two competing images first reached conscious awareness. There was no significant correlation between the strength of this rivalry effect and subjects’ explicit knowledge of the specific AV pairs, suggesting that the effects of exposure to the AV pairs on subsequent binocular rivalry responses were primarily due to implicit crossmodal learning.

The facilitative effect of statistical learning on binocular rivalry we report here is consistent with previous evidence for influences of predictive information on perceptual

selection. For example, brief presentation of a low-contrast image can increase its likelihood of being subsequently perceived during rivalry (Brascamp, Knapen, Kanai, van Ee, & van den Berg, 2007), as can merely imagining an image prior to rivalry (Pearson, Clifford, & Tong, 2008) or maintaining a “perceptual memory” trace of a dominant image across temporal gaps in stimulus presentation (Leopold, Wilke, Maier, & Logothetis, 2002; Chen & He, 2004). In addition, a grating is more likely to be initially selected during rivalry when the rivalrous pair is immediately preceded by a stream of rotating gratings whose motion trajectory predicts that grating (Denison, Piazza, & Silver, 2011). Relatedly, when rivalry trials are interleaved with trials containing a visual search task, observers are more likely to initially select the image from a rivalrous pair that results in more efficient completion of the search task (Chopin & Mamassian, 2010).

The effect of crossmodal statistical learning on rivalry that we have found is consistent with a wide range of multisensory phenomena in which information from one modality biases the perceptual interpretation of an ambiguous stimulus (Sekuler, Sekuler, & Lau, 1997; Kang & Blake, 2005; Conrad et al., 2010; Chen, Yeh, & Spence, 2011). Importantly, however, unlike these forms of congruence, which are explicit and well-learned from extensive experience (e.g., the image of a bird coupled with its song), the congruence in our study is based on probabilistic information derived from novel, arbitrary AV pairings during a brief (8-minute) and recent exposure period. Our study is therefore the first to demonstrate that visual experiences are continually updated by new statistical information regarding the relationships between images and sounds in our environment.

Our findings conflict with previous studies in which associative learning of arbitrary AV mappings—between a particular direction of 3D rotation of a Necker cube and either auditory pitch (Haijiang, Saunders, Stone, & Backus, 2006) or mechanical sound identity (Jain, Fuller, & Backus, 2010)—did not bias visual perceptual interpretation of the Necker cube. If the influence of crossmodal associations on perception depends on processes related to object identification, one possible resolution of the discrepancies between these studies and ours is that the rivaling images in our study may have been distinct enough to be represented as separate objects, each of which could be bound to a particular paired sound, whereas perceptual selection of the Necker cube is based on competing perspectives on the same object.

Because our study used simple images (sinewave, hyperbolic, and polar gratings) composed of features that are processed relatively early in the visual pathways, the effect of statistical learning on rivalry that we report here is likely to reflect selection in early stages of visual processing. Perceptual alternations during viewing of rivalrous pairs of similarly simple stimuli correlate with physiological responses in the LGN (Haynes, Deichmann, & Rees, 2005; Wilke, Mueller, & Leopold, 2009; Wunderlich, Schneider, & Kastner, 2005) and V1 (Tong & Engel, 2001; Polonsky et al., 2000; Wilke, Logothetis, & Leopold, 2006). Thus, the learned crossmodal associations in our study are likely to impact competition at early stages of visual processing, a hypothesis that is consistent with other reports of multisensory interactions in early sensory areas (Driver & Noesselt, 2008; Lakatos, Chen, O’Connell, Mills, & Schroeder, 2007; Noesselt et al., 2007). Moreover, our study provides the first evidence for very recently established multisensory associations influencing visual competition.

As organisms gather information about the statistics of the natural world through sensory experience, they form and hone associations between sounds and images. The consistency of relationships between particular sounds and images in the environment modulates these associations, influencing the likelihood of predicting the presence of one stimulus based on the

occurrence of another. For example, when we see a brown, furry animal far in the distance, as soon as we hear it bark we know it is more likely to be a dog than a bear. Our results show that rapid probabilistic learning can transform arbitrarily linked object features into automatic associations that can in turn shape perception and help resolve visual ambiguity. Our experimental paradigm is useful for characterizing the time course and mechanisms of updating of these crossmodal associations and the resulting effects on perceptual selection of ambiguous sensory inputs.

Chapter 4

4. CONCLUSIONS AND FUTURE DIRECTIONS

In the first study, we found that hemispheric differences in spatial frequency processing impact perceptual selection, indicating that our conscious awareness of spatial information differs between the left and right visual hemifields. Some questions that naturally arise from this result are: at what level of the neural hierarchy does this asymmetry emerge, and what is the biological basis of its impact on binocular rivalry? The neural substrates of rivalry have been the subject of much debate (Tong, Meng, & Blake, 2006; Blake & Logothetis, 2002; Logothetis, Leopold, & Sheinberg, 1996). Physiological correlates of perceptual alternations in binocular rivalry have been observed in areas as early as V1 (Tong & Engel, 2001; Polonsky, Blake, Braun, & Heeger, 2000) and the LGN (Haynes, Deichmann, & Rees, 2005; Wunderlich, Schneider, & Kastner, 2005). However, perceptual (interocular) grouping (Kovács, Pappathomas, Yang, & Fehér, 1996) and selective attention (Chong, Tadin, & Blake, 2005) have been shown to influence rivalry, suggesting a contribution of top-down feedback from higher visual areas.

Relatedly, hemispheric asymmetries in spatial frequency processing can reflect relative, not absolute, differences in selected spatial frequencies (Hellige, 1993; Christman et al., 1991). In particular, the same spatial frequency can be preferentially processed by either the left or right hemisphere, depending on the available range of task-relevant spatial frequencies (Ivry & Robertson, 1998). This relative nature of hemispheric asymmetries in spatial frequency processing indicates critical roles for context and top-down processing. In an ongoing study employing a broader range of spatial frequencies, we have confirmed that relative—not absolute—spatial frequency processing is responsible for the asymmetry we have reported here. Specifically, a medium SF grating, when rivaling with a high SF grating, was more likely to be perceptually selected when the rivalry pair was presented in the left, compared to the right, visual hemifield. However, this same medium SF grating, when it was paired with a low SF grating, was more likely to be perceptually selected in the right hemifield. Thus, the visual system's classification of a given SF as “low” or “high” (and therefore, which hemisphere preferentially processes that SF) depends on the other SFs that are present in the environment at any given time, demonstrating an influence of top-down, contextual processing on hemispheric differences in visual perceptual selection and conscious representations of space. In the future, these studies could be combined with neurophysiological measures to elucidate the neural substrates of these contextual influences on hemispheric asymmetries in spatial vision.

In the second study, we demonstrated that very recently learned and arbitrary associations

between sounds and images bias subsequent visual perceptual selection during binocular rivalry. These results suggest that statistical information about recently experienced patterns of sounds and images helps resolve ambiguities in sensory information by influencing low-level competitive interactions between visual neurons. Future work could elucidate whether recently learned associations can also influence perceptual selection of ambiguous sounds. For instance, the tritone paradox, an ambiguous stimulus composed of consecutive Shepard tones that is approximately equally likely to be perceived as increasing or decreasing in pitch, may be influenced by statistical learning of patterns that encourage one interpretation or the other (Deutsch, 1991). Another promising future research direction is the impact of stimulus complexity: for example, does statistical learning of associations between natural images and sounds (e.g., faces, speech sounds) have a stronger influence on perceptual selection during binocular rivalry compared to statistical learning of pairings of simple artificial visual and auditory stimuli? In addition, the factors distinguishing facilitative effects (as in the current study, in which a given sound increases the dominance of its associated image) from “surprise” or novelty effects (as in Mudrik, Deouell, & Lamy, 2011) remain unknown; future studies manipulating the length, statistical properties, or attentional demands of the exposure period could help reconcile these results.

REFERENCES

- Alais, D., & Blake, R. (Eds.). (2005). *Binocular Rivalry*. MIT Press.
- Aslin, R. N., & Newport, E. L. (2012). Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science*, *21*, 170-176.
- Blake, R. (1989). A neural theory of binocular rivalry. *Psychological Review*, *96*(1), 145-167.
- Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*, 13–21.
- Blake, R., & Wilson, H. (2011). Binocular vision. *Vision Research*, *51*, 754–770.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Brascamp, J. W., Knapen, T. H., Kanai, R., van Ee, R., & van den Berg, A. V. (2007). Flash suppression and flash facilitation in binocular rivalry. *Journal of Vision*, *7*(12), 1-12.
- Breitmeyer, B. G. (1975). Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Research*, *15*, 1411–1412.
- Bressler, D. W., Denison, R. N., & Silver, M. A. (2013). High-level modulations of binocular rivalry: Effects of stimulus configuration, spatial and temporal context, and observer state. In S. M. Miller (Ed.), *The constitution of visual consciousness: Lessons from binocular rivalry* (pp. 253–280). Amsterdam: John Benjamins.
- Chen, X., and He, S. (2004). Local factors determine the stabilization of monocular ambiguous and binocular rivalry stimuli. *Current Biology*, *14*, 1013-1017.
- Chen, Y-C., Yeh, S-L., & Spence, C. (2011). Crossmodal constraints on human perceptual awareness: Auditory semantic modulation of binocular rivalry. *Frontiers in Psychology*, *2*, 212.
- Chong, S. C., Tadin, D., & Blake, R. (2005). Endogenous attention prolongs dominance durations in binocular rivalry. *Journal of Vision*, *5*, 1004–1012.
- Chopin, A., and Mamassian, P. (2010). Task usefulness affects perception of rivalrous images. *Psychological Science*, *21*, 1886-1893.
- Christman, S. (1997). Hemispheric asymmetry in the processing of spatial frequency: Experiments using gratings and bandpass filtering. In S. Christman (Ed.), *Cerebral asymmetries in sensory and perceptual processing* (pp. 3–30). New York: Elsevier.
- Christman, S., Kitterle, F. L., & Hellige, J. (1991). Hemispheric asymmetry in the processing of absolute versus relative spatial frequency. *Brain and Cognition*, *16*, 62–73.
- Conrad, V., Bartels, A., Kleiner, M., & Noppeney, U. (2010). Audiovisual interactions in binocular rivalry. *Journal of Vision*, *10*, 1-15.
- de Weert, C. M. M., Snoeren, P. R., & Koning, A. (2005). Interactions between binocular rivalry and Gestalt formation. *Vision Research*, *45*, 2571–2579.
- Denison, R. N., Piazza, E. A., & Silver, M. A. (2011). Predictive context influences perceptual selection during binocular rivalry. *Frontiers in Human Neuroscience*, *5*, 166.
- Deutsch, D. (1991). The tritone paradox: An influence of language on music perception. *Music Perception*, *8*, 335–347.
- Driver, J. & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on ‘sensory-specific’ brain regions, neural responses, and judgments. *Neuron*, *57*, 11-23.
- Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*, 499–504.
- Flevaris, A. V., Bentin, S., & Robertson, L. C. (2011). Attentional selection of relative SF mediates global versus local processing: Evidence from EEG. *Journal of Vision*,

- 11(7):11, 1–12.
- Haijiang, Q., Saunders, J. A., Stone, R. W., & Backus, B. T. (2006). Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning. *Proceedings of the National Academy of Sciences USA*, 103, 483–488.
- Haynes, J.-D., Deichmann, R., & Rees, G. (2005). Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. *Nature*, 438, 496–499.
- Hellige, J. B. (1993). *Hemispheric asymmetry: What's right and what's left*. Cambridge, MA: Harvard University Press.
- Hughes, H. C. (1986). Asymmetric interference between components of suprathreshold compound gratings. *Perception & Psychophysics*, 40, 241–250.
- Ivry, R., & Robertson, L. (1998). *The two sides of perception*. Cambridge, MA: MIT Press.
- Jain, A., Fuller, S., & Backus, B. T. (2010). Absence of cue-recruitment for extrinsic signals: Sounds, spots, and swirling dots fail to influence perceived 3D rotation direction after training. *PLoS ONE*, 5, e13295.
- Kang, M.-S., & Blake, R. (2005). Perceptual synergy between seeing and hearing revealed during binocular rivalry. *Psychologija*, 32, 7–15.
- Kitterle, F. L., Christman, S., & Hellige, J. B. (1990). Hemispheric differences are found in the identification, but not the detection, of low versus high spatial frequencies. *Perception & Psychophysics*, 48, 297–306.
- Kitterle, F. L., Hellige, J. B., & Christman, S. (1992). Visual hemispheric asymmetries depend on which spatial frequencies are task relevant. *Brain and Cognition*, 20, 308–314.
- Kitterle, F. L., & Selig, L. M. (1991). Visual field effects in the discrimination of sine-wave gratings. *Perception & Psychophysics*, 50, 15–18.
- Kovács, I., Papathomas, T. V., Yang, M., & Fehér, Á. (1996). When the brain changes its mind: Interocular grouping during binocular rivalry. *Proceedings of the National Academy of Sciences, U.S.A.*, 93, 15508–15511.
- Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, 53, 279–292.
- Lehky, S. R. (1988). An astable multivibrator model of binocular rivalry. *Perception*, 17(2), 215–228.
- Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature Neuroscience*, 5, 605–609.
- Levelt, W. J. M. (1965). *On binocular rivalry*. Soesterberg, The Netherlands: Institute for Perception RVO-TNO.
- Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? *Nature*, 380, 621–624.
- Lunghi, C., Binda, P., & Morrone, M. C. (2010). Touch disambiguates rivalrous perception at early stages of visual analysis. *Current Biology*, 20, 143–144.
- Martínez, A., Di Russo, F., Anllo-Vento, L., & Hillyard, S. A. (2001). Electrophysiological analysis of cortical mechanisms of selective attention to high and low spatial frequencies. *Clinical Neurophysiology*, 112, 1980–1998.
- Miller, S. M., Liu, G. B., Ngo, T. T., Hooper, G., Riek, S., Carson, R. G., et al. (2000). Interhemispheric switching mediates perceptual rivalry. *Current Biology*, 10, 383–392.
- Mudrik, L., Deouell, L. Y., & Lamy, D. (2011). Scene congruency biases binocular rivalry. *Consciousness and Cognition*, 20, 756–767.

- Musel, B., Bordier, C., Dojat, M., Pichat, C., Chokron, S., Le Bas, J.-F., et al. (2013). Retinotopic and lateralized processing of spatial frequencies in human visual cortex during scene categorization. *Journal of Cognitive Neuroscience*, *25*, 1315–1331.
- Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H., & Driver, J. (2007). Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *Journal of Neuroscience*, *27*, 11431–11441.
- Pearson, J., Clifford, C. W., & Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Current Biology*, *18*, 982–986.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Peterzell, D. H., Harvey, L. O., & Hardyck, C. D. (1989). Spatial frequencies and the cerebral hemispheres: Contrast sensitivity, visible persistence, and letter classification. *Perception & Psychophysics*, *46*, 443–455.
- Pettigrew, J. D. (2001). Searching for the switch: Neural bases for perceptual rivalry alternations. *Brain and Mind*, *2*, 85–118.
- Peyrin, C., Baciú, M., Segebarth, C., & Marendaz, C. (2004). Cerebral regions and hemispheric specialization for processing spatial frequencies during natural scene recognition. An event-related fMRI study. *Neuroimage*, *23*, 698–707.
- Peyrin, C., Chauvin, A., Chokron, S., & Marendaz, C. (2003). Hemispheric specialization for spatial frequency processing in the analysis of natural scenes. *Brain and Cognition*, *53*, 278–282.
- Peyrin, C., Mermillod, M., Chokron, S., & Marendaz, C. (2006). Effect of temporal constraints on hemispheric asymmetries during spatial frequency processing. *Brain and Cognition*, *62*, 214–220.
- Piazza, E. A. & Silver, M. A. (2014). Persistent hemispheric differences in the perceptual selection of spatial frequencies. *Journal of Cognitive Neuroscience*, *26*(9), 2021–2027.
- Polonsky, A., Blake, R., Braun, J., & Heeger, D. J. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neuroscience*, *3*, 1153–1159.
- Proverbio, A. M., Zani, A., & Avella, C. (1997). Hemispheric asymmetries for spatial frequency discrimination in a selective attention task. *Brain and Cognition*, *34*, 311–320.
- Robertson, L. C. (1996). Attentional persistence for features of hierarchical patterns. *Journal of Experimental Psychology: General*, *125*, 227–249.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52.
- Seitz, A. R., Kim, R., van Wassenhove, V., & Shams, L. (2007). Simultaneous and independent acquisition of multisensory and unisensory associations. *Perception*, *36*, 1445–1453.
- Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*: 308.
- Sergent, J. (1982). The cerebral balance of power: Confrontation or cooperation? *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 253–272.
- Tong, F., & Engel, S. A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature*, *411*, 195–199.

- Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends in Cognitive Sciences*, *10*, 502–511.
- Wilke, M., Logothetis, N. K., & Leopold, D. A. (2006). Local field potential reflects perceptual suppression in monkey visual cortex. *Proceedings of the National Academy of Sciences USA*, *103*, 17507-17512.
- Wilke, M., Mueller, K.-M., & Leopold, D. A. (2009). Neural activity in the visual thalamus reflects perceptual suppression. *Proceedings of the National Academy of Sciences USA*, *106*, 9465-9470.
- Wunderlich, K., Schneider, K. A., & Kastner, S. (2005). Neural correlates of binocular rivalry in the human lateral geniculate nucleus. *Nature Neuroscience*, *8*, 1595–1602.
- Zhou, W., Jiang, Y., He, S., & Chen, D. (2010). Olfaction modulates visual perception in binocular rivalry. *Current Biology*, *20*, 1356-1358.