Associative Learning as a Source of Visual Bias

by

Davood Ghara Gozli

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy

Department of Psychology

University of Toronto

© Copyright by Davood Ghara Gozli 2015
Abstract

The visual system does not encounter isolated visual features, but sets of features that bind together and constitute complex perceptual events. Visual features are also typically accompanied by the observer's own movements, and it is believed that these movements are also included in the representation of events. In the present thesis, argue that the units of attentional selection are such multi-featured events. This means that learning the regularities in the relation among features can guide visual attention toward multiple features that belong to the selected event. Through a series of experiments, I demonstrate how visual bias for colour is influenced by the selection associated features, such as a shape or a manual movement. In all instances, I argue, selecting or inhibiting one feature impacts the treatment of the associated colour feature. In short, feature association can extend the impact of feature-based attention across multiple dimensions, which supports the idea that events, not features, are units of attentional selection.
Acknowledgments

I am greatly indebted to my advisor, Jay Pratt, for his wisdom, support, and patience during the past five years. His mentorship has been a privilege and I look forward to our future collaborations and friendship. I am also thankful to Susanne Ferber, Alison Chasteen, Ulrich Ansorge, Tim Welsh, Matthias Niemeier, James Brockmole, Liana Brown, Michael Reynolds, and Hugo Lehmann for their guidance in various stages of my research career. My colleagues Jason Rajsic and Greg Huffman played an important part in improving my thinking about this project. My friends Peter Limberg, Siavash Kazemian, Nicole Kim, Kevin McCormick, Christoph Wrzosek, Austin Sauco, Eric Taylor, Marcos Nadal, David Chan, Tarek Amer, Sabrina Thai, Jessica Maxwell, and Rannie Xu – each in their own unique way – helped me stay inspired and resilient. I should, of course, also acknowledge the financial support of the Natural Sciences & Engineering Research Council of Canada. Finally, I thank my family for their love and encouragement.
# Table of Contents

Abstract.................................................................................................................................ii
Acknowledgements................................................................................................................iv
Table of Contents..................................................................................................................v
List of Figures............................................................................................................................vii
Chapter 1: Introduction.........................................................................................................1
Chapter 2: Perceptual Associative Learning ................................................................. 13
Chapter 3: Sensorimotor Associative Learning (I): Response Execution ................. 36
Chapter 4: Sensorimotor Associative Learning (II): Response Selection .............. 47
Chapter 5: Sensorimotor Associative Learning (III): Response Priming ............. 59
Chapter 6: Visual Primes with Multiple Features....................................................... 75
Chapter 7: Conclusion....................................................................................................... 82
References..............................................................................................................................87
List of Figures

Figure 1. Sequence of events on a sample trial of the acquisition phase (top row) and a sample trial of the test phase (bottom row) in Experiments 1-2.

Figure 2. Response time data from Experiment 1, graphed as a function of the colour and location of the spatial cue. Cue colour, in relation to the primed shape, could be congruent, incongruent, or neutral. Cue location, in relation to target location, could be valid or invalid. Error bars represents 95% within-subjects confidence intervals (Cousineau, 2007).

Figure 3. Response time data from Experiment 1, divided based on the strength of the effect of the shape cue. The top panel represents the group with a strong shape cueing effect. The bottom panel represents the group with a weak shape cueing effect. Error bars represents 95% within-subjects confidence intervals.

Figure 4. Post-hoc statistical power for the critical two-way interaction in Experiment 1 (congruency x validity). The graph is generated using G*Power software, version 3.1.9.2 (Faul et al., 2007; 2009).

Figure 5. Response time data from Experiment 2, graphed as a function of the colour and location of the spatial cue. Cue colour, in relation to the primed target shape, could be congruent, incongruent, or neutral. Cue location, in relation to the target location, could be valid or invalid. Error bars represents 95% within-subjects confidence intervals.

Figure 6. Response time data from Experiment 2, divided based on the strength of the effect of the shape cue. The upper panel represents the group with a strong shape cueing effect, whereas the lower panel represents the group with a weak shape cueing effect. Error bars represents 95% within-subjects confidence intervals.

Figure 7. Response time data from Experiment 3, graphed as a function of the colour and location of the spatial cue. Cue colour, in relation to the primed target shape, could be congruent, incongruent, or neutral. Cue location, in relation to the target location, could be valid or invalid. Error bars represents 95% within-subjects confidence intervals.
**Figure 8.** Panel (a): according to the inhibition hypothesis, correct prediction of the cue feature reduces signal strength caused by the cue, which should reduce the effect of the both valid and invalid cues in our paradigm. Panel (b): according to the preactivation hypothesis, internal prediction increases the baseline activity for the predicted feature, which should strengthen the effect of valid cues. Figure reprinted from Roussel et al. (2013) with the publisher's permission.

**Figure 9.** Sequence of events in a sample trial of Experiment 4. The sample trial is from the test phase of the "invalid cue" condition.

**Figure 10.** Response time data from Experiment 4, graphed as a function of cue colour (absent, action-congruent, or action-incongruent) and cue validity (valid vs. invalid). Percentage of errors for each condition is presented at the base of the corresponding bar graph. Error bars represent 95% within-subjects confidence intervals.

**Figure 11.** Sequence of events on a sample trial of the acquisition phase (top row) and a sample trial of the test phase (bottom row) of Experiment 5-6.

**Figure 12.** Response time data from the acquisition phase of Experiments 5-6, graphed as the function of the spatial mapping between the response cue (which is also the response-outcome location) and the response key and the delay between the response cue onset and target onset (SOA). Error bars represents 95% between-subjects confidence intervals.

**Figure 13.** Response time data from the test phase of Experiments 5-6, graphed as the target location (congruent vs. incongruent colour in relation to the learned action-outcome colour) and the delay between the response cue onset and target onset (SOA). The top graph represents data from Experiment 5 (direct S-R spatial mapping in the acquisition phase), and the bottom graph represents data from Experiment 6 (inverse S-R spatial mapping in the acquisition phase). The test phase of the two experiments were identical. Error bars represents 95% within-subjects confidence intervals.

**Figure 14.** Sequence of events in one trial of the acquisition phase (above sequence) and the test phase (lower sequences) of Experiments 7-8. During the acquisition phase, the colour effect of each response was equally likely to appear inside the left or right circle. The colour prime, in test phase of Experiment 7, consisted of a change in both
placeholders above and below fixation, whereas in Experiment 8, it consisted of a change in a single placeholder and, therefore, also functioned as a visual-attentional cue (this cue was equally likely to appear above or below fixation and it did not predict the target location, which was itself equally likely to be above or below).

**Figure 15.** Response time data from Experiment 7, graphed as a function the first (spatial) and the second (colour) primes. Error rates are provided at the base of each bar. Error bars represent within-subjects 95% confidence intervals.

**Figure 16.** Response time data from Experiment 8, graphed as a function of the relationship between the first (spatial) response cue and the performed response, the second (colour) response cue and the performed response, and the validity of the colour cue in signaling the location of the imperative stimulus. Error bars represent within-subjects 95% confidence intervals.

**Figure 17.** Sequence of events on a sample trial of Experiment 9.

**Figure 18.** Response time data for the detection task of Experiment 9, graphed as the function of the relationship between the cue direction and the first target (valid vs. invalid arrow direction) and the relationship between the cue colour and the second target (repeat vs. switch). Error bars represent 95% within-subject confidence intervals.

**Figure 19.** Response time data for the colour discrimination task of Experiment 9, graphed as the function of the relationship between the cue direction and the first target (valid vs. invalid arrow direction) and the relationship between the cue colour and the second target (repeat vs. switch). Error bars represent 95% within-subject confidence intervals.
Introduction

The world of perception is abundant with regularities. Recognizing an environment as a classroom or an office guides our expectation about the kind of perceptual events we might encounter; similarly, recognizing an object as a desk shapes our expectation about other objects that we might see nearby, e.g., a chair. Once such regularities are learned, they shape how we interact with a given environment (Bar, 2007; Dayan, Kakade, & Montague, 2000; Chun & Turk-Browne, 2008). In the present set of studies, I am concerned with the process of attending to visual items and how attention might be guided by previously acquired regularities. For the present purpose, I define attention to be a selective bias that facilitates responding to one item over other available items (Bundesen & Habekost, 2008; Knudsen, 2007; Chun, Golomb, & Turk-Browne, 2011). As such, attention reflects relatively early stages of perception following the encounter with the stimulus. Modulations in attention, therefore, would be reflected in the time it takes to execute a correct response following the onset of a stimulus (Posner, 1978; Sternberg, 1969).

If we assume that attention reflects relatively early phases of perceptual processes, it is conceivable that it might be insensitive to various learned associations between objects and events, allowing the initial selection process to best serve the current demands of the environment and the observer's intentions. Not being biased by past regularities might also optimize our ability to learn new regularities. If, however, attention is sensitive to learned regularities, then each encounter with a visual display is both biased by what is previously perceived and will itself bias future perception (Dayan et al., 2000). I begin the present investigation with a review of research on visual attention, highlighting two major categories of debates surrounding the function of attention. In light of these debates, I argue for the benefit of considering associative learning as a source of attentional bias.

Traditional cognitive psychology tended to examine attention largely in separation from learning (Neisser, 1967). Separating attention from learning would be consistent with the computer metaphor of the mind and the corresponding assumption of modularity (Fernandez-Duque & Johnson, 1999; Lakoff & Johnson, 1999). Within the modularity framework, attention is regarded as a filter that selects the input for perception, independently of modules that are devoted to later stages of processing such as memory or action (Fodor, 1983). Characterizing
attention as an early filter in the information processing stream motivated much research, generating debates with regard to the locus of the purported filter, the number of items that could pass the filter at any given time, whether the unattended stimuli are completely filtered out, and so forth (e.g., Broadbent, 1958; Deutsch & Deutsch, 1963; Treisman, 1964; Sperling, 1960; 1967). Although recent perspectives characterize attention, instead of a single stage, as a property of multiple stages of information processing (e.g., Chun et al., 2011; Desimone & Duncan, 1995; Knudsen, 2007), several of the old debates regarding attention continue to motivate empirical work. I will review two broad categories of such debates, and will argue that an emphasis on associative learning as a guide for attention informs both issues. 

Most debates concerning attention fit within two general categories. The first category concerns the basis of selection (e.g., Broadbent, 1954; Deutsch and Deutsch, 1963; Theeuwes, 1993) and the second concerns the units of selection (e.g., Duncan, 1984; Kahneman, Treisman, & Gibbs, 1992). In examining the basis of selection, attention is treated as an effect and other physical, psychological, or neurological factors are examined as potential causes of attention (e.g., Corbetta & Shulman, 2002; Folk, Remington, & Johnston, 1992; Theeuwes, 1993). When examining the units of selection, attention is treated as a cause to other behavioural, psychological, or neurological effects (e.g., Briand & Klein, 1987; Carrasco, 2010; Hommel, 1998; Remington & Folk, 2001; Treue & Trujillo, 1999). The arguments and evidence that I present here are connected to both lines of questioning. When considering the basis of selection, my argument is that associative memory constitutes an important basis for selection. Perceptual bias toward a feature will cause a bias toward other features that are associated with it. This, in turn, has an implication about the units of selection. By regarding associative learning as a source of perceptual bias, it follows that the beneficiary of attentional bias are not isolated features, but a set of associated features. In other words, the units of attentional selection are multi-featured events and not single features (Hommel, 1998; 2004; Hommel, Müsseler, Aschersleben, & Prinz, 2001). Before elaborating on the argument, it would be useful to elaborate further on the two categories of debates regarding the basis and units of selection.

Causes of Selection

One of the first debates regarding the basis of selection gained popularity with Broadbent (1958), who argued that selection is determined based on crude, physical properties of the
perceptual input. The meaning of percepts, Broadbent argued, is recognized only after selection. As such, the semantic categories of items were not thought to play a role in the selection process. Broadbent's argument was heavily influenced by Cherry (1953), whose work demonstrated that an auditory signal can be effectively selected on the basis of location or pitch, but not meaning. In opposition to this view, Deutsch and Deutsch (1963) argued that the meaning of objects in the environment are recognized prior to their selection. The proposal of pre-attentive recognition was difficult to falsify for two reasons. First, pre-attentive recognition, by definition, was not required to correlate either with conscious experience or a change in behaviour (Deutsch & Deutsch, 1963). Second, pre-attentive recognition of meaning was empirically indistinguishable from the less interesting alternative proposal that the unattended stimuli are sometimes selected, due to lapses in performance (Johnston & Dark, 1986; Smallwood & Schooler, 2006).

For the most part, the issue of whether attentional selection precedes or follows categorical recognition has largely been resolved in favor of a variant of Broadbent's proposal. At present, the most influential accounts of visual attention assume that the allocation of attention is largely determined by basic physical features that lack categorical meaning (e.g., Wolfe, 2007; Wolfe & Horowitz, 2004). At the same time, there is some empirical evidence in support of the sensitivity of the visual system, prior to attentional selection, to certain object categories including fearful facial expressions (e.g., West, Anderson, Ferber, & Pratt, 2010; West, Anderson, & Pratt, 2009). Additionally, substantial periods of practice with item categories has been shown to give the practiced category the ability to guide attention in a manner that resembles basic features (e.g., Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). In the present set of studies, I only used the basic feature of colour as the target of attentional selection, thus avoiding the issue of pre-attentive categorical recognition altogether.

Another issue regarding the basis of selection is about whether the physical characteristics of the display can fully determine the allocation of attention, regardless of the internal state of the observer. Although the physical properties of the visual scene constitute a robust source of perceptual bias (e.g., Itti & Koch, 2001; Pomerantz & Portillo, 2011; Theeuwes, 2010), it is also recognized that the goals and intentions of the observer can also modulate attention (e.g., Folk, Remington, & Johnston, 1992, Yantis & Egeth, 1999; Remington & Folk, 2001; see also, Awh, Belopolsky, & Theeuwes, 2011; Theeuwes, 2010). For instance, looking for a red item increases the cost of red distractors, due to their feature resemblance to the target
(Kaptein, Theeuwes, & van der Heijden, 1995; Moore & Egeth, 1998; see also, Gozli & Pratt, 2012). This issue is relevant to the present investigation. Specifically, I will demonstrate that the effect of the observer's goal is not restricted to the primary target of the goal. Instead, if an item or a feature is selected, the selection will not only have consequences for how that item is processed, but it will also affect the processing of items that are closely associated with it due to repeated, consistent co-occurrence (Kumar, Manjaly, & Sunny, 2015). In other words, associated features can generate bias toward each other.

The present argument regarding the transfer of feature-based attention from one feature to an associated feature rests on the assumption that early feature-based attention is possible. Strong support for feature-based attention comes from single-cell recording studies of the macaque visual brain (McAdams & Maunsell, 2000; Treue & Martinez-Trujillo, 1999; for a review, see Maunsell & Treue, 2006). McAdams and Maunsell (2000) reported activity in V4 cells in response to an attended colour feature that appeared in the unattended hemifield, when monkeys attended to the feature. Imaging studies with human subjects have also confirmed the idea of feature-based attention, by finding increased neural response to features in spatially non-specific manner (Saenz, Buracas, & Boynton, 2002; Zhang & Luck, 2009). A recent study by Leonard, Balestreri, & Luck (2015) showed that feature-based attention is not uniform across the visual field, but declines as items appear further from the attended area. Nonetheless, we can assume that features other than location could serve as the basis of attentional selection.

Assuming the possibility of feature-based attention, it should be noted that attention to a feature is not always triggered by voluntary processes. Indeed, an important line of research has demonstrated how the observer’s mode of action, regardless of task goals, can bias visual processing of basic features. That is, the same visual display can be processed differently depending on what action the observer is performing or planning to perform (e.g., Bekkering & Neggers, 2002; Brown, Moore, & Rosenbaum, 2002; Hommel & Schneider, 2002; Müßeler & Hommel, 1997; for a review, see Pratt, Taylor, & Gozli, 2015). For instance, grasp movements facilitate processing shape, while pointing movements facilitate processing of location (Fagioli et al., 2007; Wykowska et al., 2009). The logic of associative learning can be applied to motor responses and their sensory consequences (Hommel et al., 2001; Hommel, 2004). If a movement is repeatedly and consistently accompanied by visual selection of a feature, then selection of the movement alone might impact visual bias toward the associated feature. In Chapters 3-5, I will
demonstrate that processes involved in selection and execution of a motor response can generate a visual bias, due to associative learning of responses and their corresponding visual outcomes.

To summarize, the initial target of attentional selection largely consists of the physical properties of stimuli. The selection of such physical properties is guided not merely by the physical salience of stimuli (Theeuwes, 2010), but also by the perceptual goal of the observer (Folk & Remington, 2006), and observer's actions (Bekkering & Neggers, 2002). An additional factor proposed in the present set of studies is associative learning. That is, given what the observer has already selected, associative learning can guide attention toward physical features that are associated with the already-selected item. This type of associative learning differs from learning to associate a feature with reward, i.e., operant conditioning, or the repeated exposure to a feature, i.e., repetition priming (Awh et al., 2012; Anderson, Laurent, & Yantis, 2011). Instead, this form of associative learning is based on repeated co-occurrence of features (e.g., Chun & Jiang, 1998; 1999; see also, Brockmole, Castelhano, & Henderson, 2006; Brockmole & Henderson, 2006).

**Effects of Selection**

The second category of questions concerns the units of selection. That is, once attention is allocated based on a physical feature, what is then selected? This issue is important because depending on what is selected, and what is not selected, attention will have different impact on subsequent behaviour. Research on the units of attention was also influenced by the seminal work of Broadbent (1958) and the particular language he used in framing the function of attention. Specifically, he described attention as selecting an "input channel", defining an input channel as a set of stimuli that share a common physical feature (e.g., location). The concept of a channel, although clear in communication engineering, is rather vague when applied to psychology (Fernandez-Duque & Johnson, 1999). When listening to a friend, for instance, what should be regarded as the input channel? Is it the friend's voice, the friend's voice plus other facial expressions and gestures, or the friend's expressions plus the shared conversational context? (Bartlett, 1958; Sperber & Wilson, 1987) Regardless, when considering the units of selection, Broadbent's work served as an important ground for future work.

In the domain of visual attention, the works of Michael Posner and Anne Treisman during the early 1980s had considerable impact in upholding that location is the unit of selection.
Posner employed experimental designs in which participants could attend to one out of multiple possible locations in space (1980; Posner & Cohen, 1984; Posner, Snyder, & Davidson, 1980; for a review, see Wright & Ward, 2008). The methodological character of Posner's paradigm promotes a spatial characterization of visual attention (Fernandez-Duque & Johnson, 1999). Treisman also characterized attention as operating on an internal map of space. According to both frameworks, attention was thought to be allocated spatially (Treisman, 1988; Treisman & Gelade, 1982; Treisman, Kahneman, & Burkel, 1983). An item could not be selected without first selecting a location. Spatially non-specific feature inhibition was considered possible, but attending to a location seemed precede selection of other features (e.g., Nissen, 1985; Treisman, 1988; Treisman & Gelade, 1982; Treisman & Sato, 1990; see also, Bundesen, 1991).

Despite the importance of location, evidence soon amassed against the hypothesis that location is the unit of selection. Duncan (1984) reported that identifying two features of a single visual object is easier compared to reporting two features from two different objects, suggesting that identifying features from different objects requires a disengagement of attention, even when the two objects occupy the same location. By contrast, identifying two features of the same object does not require such a disengagement. Kahneman and Treisman (1984) similarly proposed that the outcome of attentional selection is one representational unit, "object file", that includes all features of the selected object, e.g., location, shape, size. The benefit of such a unified representation includes efficient tracking and updating of object features, when the object changes some of its features (Hollingworth &Franconeri, 2009).

Strong support for the notion of object file came from the study of Kahneman, Treisman, and Gibbs (1992), who used a sequence of three displays on each trial. The first display was the preview, consisting of two letters each inside a square placeholder. The letters then disappeared and the two squares moved to new locations. Finally, one letter appeared in one of the squares. Participants reported whether they had seen the letter in the preview display, regardless of its location. The experiment consisted of three types of trials; same-object repetition (a letter reappeared in its original placeholder), different-object repetition (a letter reappeared but in the alternative placeholder), and switch (an altogether new letter appeared in a placeholder). Performance was consistently faster in the same-object repetition compared to the different-object repetition. These object-specific preview effects were taken as evidence for binding the placeholder and the letter into an object file. Whereas with same-object repetition observers
could rely on the object files they had already formed during preview, different-object repetition violated the object files and required updating those representations, causing slower and less accurate performance.

The preview paradigm of Kahneman et al. (1992) provided an important contribution to the issue of units of attentional selection. Attention does not seem to select single features, but multi-featured objects. Research by Hollingworth and Franconeri (2009) supports the notion of object files as consisting of both surface features (shape and colour) and location. When performing object recognition, performance seems to rely both on the continuity of objects across time and space, and the preservation of surface features (Hollingworth & Franconeri, 2009). Furthermore, feature selection is thought to precede the formation of object files. For this reason, feature similarity between visual items can interfere with creating unique object files that correspond to separate objects in the display (Goodhew, Edwards, Boal, & Bell, 2015). Indeed, feature similarity not only interferes with finding a target among feature-similar distractors (Duncan & Humphreys, 1989), it can even interfere with detecting two similar feature-singletons that appear against a background of uniform distractors (West, Pun, Pratt, & Ferber, 2010).

Given that feature selection is followed by creation of object files, could we then presume that all features of a selected object are favored equally? Chen (2009) offered an important contribution to this question. In a set of experiment, participants were presented with a sequence of displays each containing a single task-relevant item. The first item, the cue, determined whether or not participants were required to respond to the target on that trial. The second item, the target, required a discrimination response based on one feature (e.g., shape). The findings showed that the non-spatial feature of the cue was processed only if it was task-relevant. For instance, the similarity between cue colour (task-irrelevant) and target colour only facilitated performance if participants were discriminating target colours (Chen, 2009; see also, Chen, 2005; Hommel, 1998; Memelink & Hommel, 2013; Pratt & Hommel, 2003; Rajsic & Wilson, 2014). These findings show that the task-relevant dimension is processed to a larger degree at the expense of task-irrelevant dimension (Found & Muller, 1996; Muller & Krummenacher, 2006; Muller, Reimann, & Krummenacher, 2003).

In light of the observation that task-relevant features of visual events receive more extensive processing than the irrelevant features (e.g., Chen, 2009), in the present thesis, I argue against a strong version of feature-specific attention. It should be noted that previous studies on
the selective bias against task-irrelevant feature involved no systematic association between features across different dimension. For instance, in Chen's (2009) experiments, it is unclear whether the cue's irrelevant feature would have remained unattended if its value had been predictive of the cue's relevant feature (see, e.g., Bruhn & Bundesen, 2012). Therefore, when features across different dimensions (e.g., shape and colour) are predictive of each other, then selection of a feature across one dimension might modulate attention to the second dimension. This proposal challenges the strong version of feature-specific (or, dimension-specific) attention.

Selection of Events

Following the studies of Kahneman et al. (1992; Treisman, 1992), Hommel (1998) extended the concept of a object representation to include not only the perceptual features, but also features of the observer's own action. He introduced the concept of "event files" to refer to the unified representation consisting of both perceptual and motor constituents (Hommel, 1996; 1997; 1998; 2004). The inclusion of movements in the event representation is consistent with the notion that perceptual systems evolved in synergy with action systems and, as such, any mechanism that binds features into objects must contain information about responses made toward such stimuli or responses that brought about the stimuli (O'Regan & Noë, 2001).

To provide an empirical demonstration of event files, Hommel (1998) employed a modified version of the preview experiment of Kahneman et al. that included the participants' motor response as an additional feature. The procedure on each trial consisted of a sequence of three stimuli; a response cue (left or right arrows indicating which response to prepare), the first stimulus, to which the prepared response is produced, and the second stimulus, which is the identification target. In this procedure, the response to the first stimulus is a detection response, while the response to the second stimulus is a discrimination response (e.g., left key for ‘O’ and right key for ‘X’). Both stimuli could vary in shape (X or O), colour (red or green) and location (above or below fixation). In relation to the first stimulus (S1), the second stimulus (S2) could repeat or switch along any of the three visual dimensions. In addition, the response feature coinciding with S1 and S2 could also be switched or repeated. Thus, sometimes the two events were exactly the same (full repetition), sometimes entirely different (full switch), and sometimes differed along a subset of features (partial repetition). The findings showed that repeating a single visual feature or response does not have a benefit if other features are switched. That is,
perceptual priming effect is not based on feature repetition. The effect is, instead, driven by repetition of bound features. Furthermore, findings indicated that not all features participate in the construction of an event file. Namely, shape and colour were reliably bound with location and response only when they are task-relevant (Hommel, 1998; cf., Chen, 2005; 2009).

Another important line of work that lends support to the notion of event files comes from investigations on action planning. Theories of action assume that, although action and perception involve distinct underlying neural and cognitive processes, they operate on shared perceptual features (Hommel, 2004; Hommel et al., 2001; Müsseler, 1999; Prinz, 1990; 1997). Specifically, the representation of an action is thought to include the expected perceptual outcomes of that action. For instance, turning on a light involves planning to bring about the proprioceptive sensation of the arm moving toward the light switch, the tactile sensation of the fingers against the switch, the ‘click’ sound from the switch, and the visual sensation of light turning on. The assumption that action representation and planning involves recruiting perceptual codes provides the foundation for the ideomotor theory of action (James, 1890; Shin et al., 2010; Stock & Stock, 2004). This theory proposes that intentional action planning consists of planning to bring about a set of desired perceptual outcomes. At its core, the ideomotor theory is a theory of event files that begins with action selection. While an event-based theory of attention assumes that selection of a single visual feature results in the selection of a multi-featured perceptual event, an event-based theory of action assumes that selection of a movement involves selection of a multi-featured perceptual-motor event (Hommel, 2009; 2010; Hommel & Elsner, 2009).

Event-based theories of action presuppose that the anticipated perceptual outcome of an action is incorporated into the action plan such that its congruency with other features of the action would influence performance efficiency (Ansorge, 2002; Elsner & Hommel, 2001; Hommel, 1993, 1996; Kunde, 2001). For instance, in a task in which a single key should be pressed softly or forcefully, performance is more efficient if the soft keypress produces a soft tone and the forceful keypress produces a loud tone, compared to the opposite response-outcome mapping (Kunde, 2001; 2003; Kunde, Hoffman, & Zellmann, 2002; Kunde, Koch, & Hoffmann, 2004). Similarly, in a task where a left or right keypress should be performed, performance is more efficient if the keys turn on lights on corresponding locations (e.g., left keypress turning on a light on the left-hand side), compared to when they turn on lights on opposite locations.
These findings suggest that the anticipated perceptual outcome of an action is essential in action selection.

Further support for event-based response selection comes from studies in which perceiving a known action-outcome, prior to response selection, biases participants toward selecting the corresponding response (e.g., Hommel, 1996; 2004; Elsner & Hommel, 2001; 2004; Ziessler & Nattkemper, 2011). For instance, participants in a study by Hommel (1996, Experiment 3) learned, in an initial acquisition phase, that each of their two possible responses would produce a tone of a distinct pitch (e.g., left key & low tone, right key & high tone). Next, in a test phase, the tones were presented together with target stimuli. Although the task only required responding to visual targets, the tone could either be congruent or incongruent with the learned effect of correct response. Findings showed faster and more accurate performance when the tone matched the known effect of the correct response. Later experiments by Elsner and Hommel (2001) ruled out the possibility that these findings were due to stimulus-stimulus mapping between target letters and the tones. Therefore, perceiving a sensory action-outcome prior to response selection can bias participants toward selecting the corresponding response. This observation supports an event-based proposal that selecting a set of perceptual features is an essential component of action selection.

In short, the notion of event file has been useful when considering the units of attentional selection. The outcome of selection seems to be a unified event that consist of multiple integrated features, including both perceptual and motor features. Furthermore, event-based theories have been useful in explaining findings from studies of response selection. Selection of a response seems to involve selection of the known perceptual outcomes along with the motor plan. The course of selecting an event, therefore, could begin with selecting a perceptual feature or selecting a motor response. In both cases, the outcome of attention appears to be a unified event (Hommel, 2004). Where event-based theories have had the least impact is when considering the causes of visual attention (Schütz-Bosbach & Prinz, 2007). That is the scope of the present investigation. In the following chapters, I ask whether the binding of two features into an event representation will influence visual attention. More specifically, would selecting or inhibiting one feature of an event influence processing of a second event feature?
Overview of the Chapters

The original notion of an event file refers, specifically, to the short-term post-selection processing phase, although it is considered that repeated exposure to consistent co-occurrence of features could result in relatively long-term event representation (Elsner & Hommel, 2001; Hommel & Elsner, 2009; Logan, 1988; 1990). In order to expand the scope of event file representation to long-term processing, in Chapters 2-5, I will focus on attentional consequences relatively long-term and stable event representations that have been acquired over a pre-test period of learning. In Chapter 6, I will return to the topic of short-term consequences of event representation.

Chapter 2 describes an experiment in which participants were introduced to two arbitrary shape-colour associations in an initial acquisition phase. That is, they consistently perceive one shape (e.g., "S") to appear in one colour (e.g., red), and another shape (e.g., "H") to appear in another colour (e.g., green). Following the acquisition phase, the experiment tested whether expecting to see one target feature (shape) could lead to prioritizing the other task-irrelevant feature (colour). Finding a colour-based bias that begins with selection of shape supports the argument that attention to a colour feature is influenced by prioritizing the event to which it belongs. This chapter, therefore, represents an instance of perceptual associative learning and its possible impacts on feature-based attention.

Chapters 3-5 describe experiments in which participants are introduced to two arbitrary response-colour associations. Following the ideomotor learning principle, in the acquisition phase, each response was immediately followed by a colour response-outcome. Assuming that the colour-outcome was then part of the response event-file, the experiments tested the impact of response selection on visual bias for colour. In Chapter 3, the impact of response execution was examined on visual bias. This chapter begins with a review of previous research on perception of learned action-outcomes and attempts to connect this literature with the broader concept of associative learning as a guide for perception. Chapter 3, therefore, is particularly important in

light of the literature on self-caused sensory stimuli as well as feature-based visual attention (Cardoso-Leite et al., 2010; Kumar et al., 2015; Waszak et al., 2012).

In Chapters 4-5, I discuss the impact of response selection on visual bias. Unlike the experiment in Chapter 3, visual bias is examined prior to response execution. This design enables an examination of the time-course of interaction between response selection and the visual bias in favor of the associated colour. In the experiments reported in Chapter 4, each response is selected based on a response cue that fully determines the correct response. By contrast, in the experiment reported in Chapter 5, response cues are uninformative and cause a response bias in an automatic manner. Critically, the findings suggest that both voluntary and involuntary response activation can modulate visual bias in favor of the associated response-outcome.

In Chapter 6, I explore a novel consequence of event-based attentional selection. While previous studies (including Chapters 2-5) show that selection of a feature results in selection of an entire event, the variety of ways in which a feature could be selected, or inhibited, are not determined. The experiment reported in Chapter 6 examines consequences of event-based selection. Results demonstrate that the colour of an arrow remains selected (i.e., generating a bias in subsequent events) if its direction validly signals a target. By contrast, if the direction of an arrow is invalid in relation to the visual target, the colour of the arrow will also not be retained. The findings of this experiment suggest that selection of an event could take place in terms of higher-order or symbolic features, such as directional validity of a colourful arrow, providing further evidence for an event-based view of visual attention.

To summarize, the following studies will explore the consequences of associative learning for visual attention. When features consistently and repeatedly co-occur, they form relatively stable event representations. Once an event representation is activated, through selection of a feature, other features of the event will be selected more readily. This, in turn, generates a bias toward the visual features that belong to the event, compared to features that do not belong to the event. The conclusions that result from studying the interplay between associative learning and visual bias concern both the basis of selection and the units of selection. Given that visual events do not randomly co-occur in real-world situations, it is reasonable to assume that selection of features typically leads to selection of an event representation, guiding attention across multiple dimensions at the same time.
Chapter 2: Perceptual Associative Learning can Guide Feature-Based Attention

Previous work on visual attention has examined the impact of learned regularities, with perhaps the simplest form of regularity being probability of target appearance in space. If a location has a relatively high probability of containing the visual target, participants can learn this and begin to respond more efficiently to targets at this location, relative to low-probability locations (Shaw & Shaw, 1977). The task, however, could involve multiple target locations and multiple potential targets, with each visual target having a unique spatial probability distribution. Employing such a design, Miller (1988) found that learning about a high-probability location does not confer an equal benefit for all visual targets, but rather benefits only the specific target associated with that location.

A more complex type of regularity has to do with the relationship between a visual item and its surrounding items. In general, exposure to arrangements of visual objects can lead to learning consistent patterns, including where particular shapes are likely to appear in relation to one another (Fiser & Aslin, 2001). Examining the impact of such learning on visual attention, Chun and Jiang (1998) employed a visual search task in which the spatial arrangement of distractors reliably predicted target location, and found faster performance relative to when distractor locations were not predictive. Similarly, when distractor shapes predicted target shape, performance was faster than when distractor shapes were not predictive (Chun & Jiang, 1999).

In addition to the regularities in spatial layout (e.g., Brockmole et al., 2006; Chun & Jiang, 1998; 1999), a recent study by Zhao, Al-Aidroos, & Turk-Browne (2013) examined whether observers are sensitive to temporal regularities of items appearing within one location. The procedure involved maintaining gaze at a central fixation while a sequence of shapes appeared at four peripheral locations. At one of the peripheral locations, the shapes followed a predictable order of appearance, whereas the shapes did not follow a predictable pattern at the other locations contained. Findings showed that observers were faster to respond to visual target that appeared at the predictive sequence of shapes, compared to when it appeared at a non-predictive sequence. The authors argued that attention is biased toward the location (or feature; Zhao et al., 2013; Experiments 2 & 3) that is associated with a predictive sequence of stimuli, although it is also possible that the interruption a predictive sequence captures attention more
effectively than interruption of a non-predictive sequence. Regardless of the exact reason, the findings clearly demonstrate that the visual system can detect temporal regularity in a stimulus sequence.

A consequence of visual learning, pertaining to the present study, is that learning can give a non-target stimulus attentional advantage due to its association to a target stimulus. That is, selection of the initial target feature might extend to the selection of other features associated with the target. William James considered this form of attention and called it derived attention (James, 1890), and prior to James, Sir William Hamilton described a similar phenomenon in memory as redintegration (Wertheimer, 2012). A present-day demonstration of derived attention was reported by Lambert et al. (1999). On each trial, participants responded to a target appearing inside the left or right placeholder. Just prior to target onset, the letters ‘S’ and ‘W’ were presented near the two placeholders, with one letter having a high probability (80%) of appearing near the upcoming target. After some exposure to the task, participants’ performance demonstrated signs of learning this letter-target association. The time-course of this form of bias mirrors stimulus-driven spatial orienting of the cued location (facilitation and inhibition with short and long cue-target delays, respectively; Posner & Cohen, 1984), but further experiments suggested a dissociation between the two (Lambert et al., 2000; also, Geng & Behrmann, 2002; 2005).

Although previous work supports the possibility of derived attention, evidence thus far has been obtained in rather simple conditions. For instance, participants can learn to prioritize the space near 'S' over 'W' when the two letters are laterally presented (Lambert et al., 1999; 2000), the left half of an object over its right half (Kristjánsson et al., 2001), or the green portion of an object over the red portion (Kristjánsson and Nakayama, 2003). What these tasks have in common is the requirement to continually select a single feature over an alternative feature. It is, therefore, unclear whether a more flexible form of attention to non-target stimuli could develop for both possible features. For instance, in the experiments by Lambert et al. (1999; 2000), could participants learn to differently prioritize between the two letters ('S' and 'W') based on other information? There is, indeed, some evidence against such possibility.

Evidence against derived attention was found in an experiment by Kristjánsson et al. (2001), where the target predictably switched from appearing on the left or right portion of a cueing object. Interestingly, this fully predictable pattern (an association between trial order and target location) did not lead to better performance compared to a fully randomized condition. In a
variant of this design, Kristjánsson and Nakayama (2003) tested whether participants could benefit from a relationship between cue feature (e.g., colour) and the segment of the cue that contained the target (left vs. right). When target location was contingent on a cue colour (or cue shape), there was no performance benefit relative to complete randomization. The authors concluded that while repetitive selection of one feature facilitates further selection of that feature (similar to continuously selecting 'S' over 'W' in Lambert et al., 1999), second-order relationships between features (i.e., multiple rules in the form of “if green then left”) are not possible. It is worth noting that the type of learning reported by Kristjánsson et al. emerged quite rapidly and within a few trials (for a review, see Kristjánsson, 2006). This time frame may not be sufficient for higher-order learning or its impact on attention.

Another reason Kristjánsson and Nakayama (2003) did not find any attentional benefit with second-order relationships could be that the cue features (itself a task-irrelevant dimension) predicted an irrelevant target feature (within-cue location) rather than the relevant target feature (i.e., target identity). Given the complex task structure, involving four separate dimensions (target identities, target locations, cue features, and within-cue locations), the irrelevant dimensions may receive reduced processing weight compared to when the irrelevant dimension is set on a constant feature and is selected on every trial (e.g., Kristjánsson et al., 2001; Lambert et al., 1999; 2000). In effect, binding between the two irrelevant features, and the subsequent associative learning, may have been prevented. As demonstrated by Hommel (1998), the chance of feature binding between two irrelevant dimensions is lower than the chance of binding between a relevant dimension and an irrelevant dimension.

In addition to the task-irrelevance of the to-be-associated features, another problem may be with the brief exposure times (~70 ms) used by Kristjánsson and Nakayama (2003). Requiring extraction of information from thin slices of time (i.e., masked, briefly-presented displays) may artificially alter the contribution of the magnocellular pathway (Westheimer & Pettet 1990). Although an increase in magnocellular activity would facilitate perception through higher temporal acuity, it may interfere with feature binding (Gozli, Ardron, & Pratt, 2014; Kelly & Brockmole, 2014) and, therefore, with visual associative learning. Consistent with this possibility, Braet and Humphreys (2009) proposed that feature binding takes place following the first 200 ms after stimulus exposure. This 200-ms delay, according to Braet and Humphreys, reflects the necessity of the posterior parietal feedback in the binding process.
For all of the aforementioned reasons, it remains possible that we can learn second-order associations, in a way that would impact visual attention. Exploring this possibility is the major goal of this Chapter. To assist associative learning between features, we implemented an initial acquisition phase (addressing the issue of sufficient learning time) and used a task in which one of the two features was task-relevant (without briefly-presented, masked visual targets).

Experiment 1

Participants. Twenty University of Toronto undergraduate students gave their informed consent and participated in exchange for course credit. Participants reported normal or corrected-to-normal vision and were unaware of the purpose of the study. All experimental protocols reported in this thesis were approved by University of Toronto Research Ethics Board.

Apparatus. All experiments reported in this thesis were run in Matlab (MathWorks, Natick, MA), using the Psychophysics toolbox (Brainard, 1997; Pelli, 1997; version 3.0.8) on Windows-run PCs. The tasks were performed in dimly lit rooms. Stimuli were presented on 19" CRT monitors set at 1024 x 768 resolution and 85 Hz refresh rate. Using a head-rest, participants' distance from the display was fixed at about 45 cm.

Stimuli. The display structure and the sequence of events are shown in Figure 1. Stimuli appeared in white, red, green, or blue, against a black background. Two horizontally aligned squares (size = 2° x 2°) functioned as placeholders for shape stimuli (targets: 'S' and 'H'; distractors: 'E' and 'U', approximately .5° x 1° in size, presented in Arial font). The center of each placeholder deviated by 5° of visual angle from fixation.

Acquisition procedure. Each trial of the acquisition phase began with the presentation of empty placeholders and the central fixation cross "+" (1° x 1°). Participants were instructed to begin by looking at the fixation cross. After a 1000-ms delay, letter stimuli appeared inside the placeholders, consisting of one target letter ('S' or 'H') and one distractor letter ('U' or 'E'). Participants were instructed to respond as quickly as possible with a right-hand keypress (the '/' button on the keyboard) when identifying 'S' and with a left-hand keypress (the 'Z' button on the keyboard) when identifying 'H'. The letters remained on the screen until one of the two possible responses was recorded. Immediately after a response, all items on the display changed colour (turning red after a right-hand keypress, and green after a left-hand keypress). This colour
outcome remained on the screen for 500 ms, and was produced depending on the performed response and regardless of the correctness of the response. After an incorrect keypress, participants first saw the colour outcome associated with their performed response (500 ms), and then the error feedback (“MISTAKE”, 2000 ms). Participants also received feedback for anticipation errors ("TOO QUICK!") or late responses ("TOO LATE!"), respectively defined as response times faster than 100 ms and slower than 2000 ms.

![Diagram of Acquisition and Test Phases](image)

**Figure 1.** Sequence of events on a sample trial of the acquisition phase (top row) and a sample trial of the test phase (bottom row) in Experiments 1-2.

**Test procedure.** Each trial of the test phase also began with the presentation of the central fixation cross and the empty placeholders. After a 1000-ms delay, the fixation cross was replaced by a shape cue: one of two possible targets ('S' or 'H'), remaining on display for 200 ms. Participants were informed that the shape cue predicted the target with 80% accuracy. The shape cue was then replaced by the fixation cross, remaining for another 500 ms. Next, the colour transient appeared: one of the placeholders changed colour (to red, green, or blue) and after a 200-ms delay changed back to white. Participants were informed that the colour transient was uninformative. Simultaneously with the disappearance of the colour transient the letter stimuli appeared inside the two placeholders. Similar to the acquisition phase, all stimuli changed colour
immediately after a response. The shape-response-colour mappings did not change across the two parts of the experiment for all participants.

**Design and Analysis.** Participants completed one acquisition phase, consisting of 200 trials, followed by a test phase that consisted of 240 trials. Both phases preceded by 10 practice trials. Target letter (‘S’ vs. ‘H’), distractor letter (‘E’ vs. ‘U’), and target location (left vs. right) were randomized and equally probable. In addition, during the test phase the central shape cue was equally likely to indicate 'H' or 'S', predicting target shape with 80% validity. On the other hand, location and colour of the spatial cue were both uninformative, predicting target location and response with chance-level validity. The colour transient coincided with target location on 50% of trials (valid). On a third of the trials, the colour transient had the same colour as the outcome of the primed sequence (congruent). On another third of the trials, the transient had the colour associated with the alternative sequence (incongruent). Finally, on a third of trials the transient was blue, not associated with either sequences (neutral).

Analysis of data from the test phase was performed in two steps. First, we tested the effect of shape cues. In particular, we needed to establish that the central letter cue biased performance without being the sole source of response selection. That is, trials with invalid shape cues (20% of trials) served as catch trials and high accuracy on these trials would suggest that participants processed the target stimuli, rather than solely responding based on the shape cue. In the second step, performance was analyzed as a function of the colour of the spatial cue (congruent, incongruent, or neutral in relation to the activated stimulus-response-outcome sequence) and the location of the spatial cue relative to the target (valid vs. invalid). In this step of the analysis valid and invalid shape-cue trials were collapsed together. If attention is biased toward the colour cue, we expect faster response times (RTs) with valid spatial cues, compared to invalid spatial cues. If the validity effect is sensitive to whether the colour of the cue belongs to the expected stimulus-response-outcome sequence, a stronger validity effect generated by congruent spatial cues compared to incongruent spatial cues should be found.

**Results**

**Acquisition.** After excluding incorrect responses and responses faster than 100 ms and those that fell beyond 2.5SDs of the total mean (2.6% of trials), mean RT was 551 ms (SE = 14).
Mean error rate was 4.0% (SE = 1%). Overall, these data suggest good performance during the acquisition phase.

**Test phase: effect of shape cues.** To ensure that participants processed the shape cue without actually responding to it, we compared RTs and error rates as a function of the validity of the shape cue. With valid shape cues, responses were faster (M ± SE = 515 ± 18 ms) and more accurate (error rate = 3% ± 6%), compared to the RTs (579 ± 17 ms, t[19] = 4.86, SE = 13, p < 0.001) and error rate (7% ± 1.3%, t[19] = 2.53, SE = .012, p = .02) following invalid shape cues. This suggests that the shape cue, on average, biased performance. On the other hand, with a mean accuracy of 93% on trials with invalid shape cues (all participants had accuracies above 77% on these trials), we reasoned that participants made responses to the target stimuli and not to the shape cue.

**Test phase: effect of colour transients.** Mean RTs were submitted to a 3 x 2 repeated-measures ANOVA, with colour (congruent, incongruent, or neutral) and location (valid vs. invalid) as factors (see Figure 2). This ANOVA revealed no main effect of colour, F(2,38) < 1, but there was a main effect of location, F(1,19) = 16.60, p < .001, η² = .466, and a two-way interaction, F(2,38) = 3.68, p = .034, η² = .162. Overall, RTs were faster with valid spatial cues (519 ± 16 ms) than invalid spatial cues (534 ± 18 ms). Most importantly, examining the two-way interaction revealed significant validity effects when the colour of the transient was congruent [t(19) = 3.48, SE = 6.34, p = .003] or neutral [t(19) = 3.04, SE = 6.55, p = .007], but not with incongruent transients [t(19) < 1].

An initial analysis of error rates did not reveal any significant finding. To obtain a more sensitive analysis of identification errors, missed trials were excluded (those including errors of anticipation or late responses; together constituting only 4% of the total trials). Due to the low percentage of the remaining errors (3.8%), we performed the repeated measures ANOVA on arcsine transformation of the mean error rates. This analysis revealed a marginal effect of cue colour (F[2,38] = 2.79, p = .074, η² = .128), but no effect of cue location (F[1,19] = 1.28, p = .271, η² = .063) or a two-way interaction (F < 1). The marginal effect of cue colour was indicative of a higher error rate with congruent colours (4.5%), compared to incongruent (3.2%) and neutral colours (3.6%). It is important to distinguish the interpretation of this effect from that of a two-way interaction. Whereas a two-way interaction would indicate different attentional treatment of colours, a main effect of colour does not indicate such attentional modulation.
Instead, the main effect of colour likely indicates a response bias (e.g., the colour green made subjects more sensitive to selecting the "H" response, regardless of target location and cue-target spatial relationship). This response bias would increase the likelihood of erroneously reporting a letter that is associated with the colour cue, but it will not vary based on cue location. In the absence of a two-way interaction in identification errors, we can rule out the possibility that the main RT findings were due to speed-accuracy trade-off.

Next, we submitted percentage of missed trials (arcsine transformed) to the same ANOVA. This analysis did not reveal any main effect or interaction (F values < 1.2, p values > .3). In spite of the low percentage of missed trials, it stands to reason that anticipatory errors and late responses do not vary systematically with our experimental manipulations.

![Figure 2](image.png)

**Figure 2.** Response time data from Experiment 1, graphed as a function of the colour and location of the spatial cue. The cue colour, in relation to the primed shape, could be congruent, incongruent, or neutral. The location, in relation to target location, could be valid or invalid. Error bars represent 95% within-subjects confidence intervals (Cousineau, 2007).
The findings of the present experiment showed, first of all, that the shape cue successfully biased performance. In addition to faster responses on valid shape-cue trials, shape cues also influenced attention to colour transients. Specifically, a stronger orienting effect by congruent transients relative to incongruent transients suggests a bias toward the colour of the activated sequence and/or a bias against the colour of the alternative sequence.

There are two indications that the modulation of attentional bias for colour transients was, at least partly, due to inhibition of the incongruent colours. The first indication is that the cueing produced by the congruent and neutral colours were statistically indistinguishable. The second indication of inhibition of incongruent colours was revealed in examining individual differences in using the shape cue. We assumed, on the one hand, that any different treatment of colour transients (i.e., the difference between congruent and incongruent colours) would be stronger in participants who utilized the shape cue more reliably. On the other hand, those participants who ignored the shape cue or used this cue less reliably were expected to show similar treatment of congruent and incongruent colour transients. Using the median-split method, I divided the participants based on the effect of the shape cue (invalid RT – valid RT), which resulted in a "strong shape cueing" group (mean cueing effect ± SE = 108 ± 16 ms) and a "weak shape cueing" group (21 ± 10 ms). I then submitted the grouped RT data to a mixed 3 x 2 x 2 ANOVA with congruency and validity of the spatial cue as the within-subjects factors and the strength of shape cueing as the between-subjects factor (see Figure 3). Overall, a main effect of group (F[1,18] = 8.98, p < .01, \(\eta_p^2 = .33\)) revealed that participants who utilized the shape cue strongly on average performed faster (484 ± 20 ms) than participants who did not utilize this cue (569 ± 20 ms). Similar to the original analysis, this ANOVA also revealed no main effect of congruency (F[2,36] < 1), a significant main effect of validity (F[1,18] = 16.44, p < .001, \(\eta_p^2 = .48\)), and a two-way interaction between congruency and validity (F[2,36] = 4.33, p = .020, \(\eta_p^2 = .194\)).

Most importantly, a three-way interaction (F[2,36] = 4.31, p = .021, \(\eta_p^2 = .193\)) pointed to different patterns of findings across the two groups. As seen in the lower panel of Figure 3, participants who used the shape cue robustly showed spatial cueing by congruent (t[9] = 2.24, SE =7.14, p = .05) and neutral (t[9] = 3.26, SE =9.12, p = .01) colour transients, but no such cueing effect by the incongruent colour (t[9] = 1.67, SE = 4.40, p = .13). Indeed, the direction of the effect by incongruent colours was opposite to what is expected, indicative of inhibition. By contrast, participants who did not reliably use the shape cues showed significant orienting effect by congruent (t[9] = 2.67, SE =10.53, p = .026) and incongruent (t[9] = 2.71, SE =5.13, p =
.024) colours, though they did not show such cueing effect by the neutral colour (p > .2). This is consistent with a relationship between using the shape cue to activate one of the possible shape-colour episodes and inhibiting the alternative shape-colour episode.

Figure 3. Response time data from Experiment 1, divided based on the strength of the effect of the shape cue. The top panel represents the group with a strong shape cueing effect. The bottom
Experiment 2

In Experiment 1, the stimulus-response-outcome mappings were completely consistent across the acquisition phase and the test phase. This consistency did not allow us to narrow down the mechanism underlying the bias for colours. In particular, a shape cue may have modulated attentional bias to colours due to the shape-colour associations and/or due to the response-colour associations. Experiment 2 aimed to test the potential role of response-colour associations. Specifically, I asked whether learned shape-colour associations alone can drive the modulated attention to colours. In this case, after the acquisition phase, we switched the responses assigned to the target shape, thus switching both response-shape associations and response-colour associations. This switch, however, preserved the shape-colour association. Consequently, shape and response will be activating two different colours. Congruency of colour transients was defined based on the shape-colour relationships. A similar pattern of results to Experiment 1 would suggest that the altered bias for colours is indeed driven by learned shape-colour associations. Finding the opposite pattern of results would indicate that response-colour associations were the sole underlying mechanisms that modulate attentional bias for colours.

The stimuli, procedure, and design were identical to Experiment 1, with the following exception. After completing the acquisition phase, participants were instructed to switch the response keys assigned to each target letter. That is, in the test phase participants responded with the left-hand key to 'S' and with the right-hand key to 'H'. The colour effects associated with the keys were also switched, which means the shape-colour associations were preserved across the two phases of the experiment. Twenty-one University of Toronto undergraduate students participated in this experiment. They reported normal or corrected-to-normal vision and were unaware of the purpose of the study.

A brief note on selection of sample size is necessary at this point. Based on the effect size observed in Experiment 1, Cohen's $d = .76$ for the two-way interaction between validity and congruency (congruent vs. incongruent), the required sample size has to include over twenty participants for statistical power to reach 90%. The graph for this analysis is displayed in Figure
4. In the following experiments, therefore, I followed this conclusion and collected 20-27 participants before analyzing the data.

**Figure 4.** Post-hoc statistical power for the critical two-way interaction in Experiment 1 (congruency x validity), excluding the neutral condition. The graph is generated using G*Power software, version 3.1.9.2 (Faul et al., 2007; 2009).

**Acquisition phase.** After excluding incorrect responses and responses faster than 100 ms and those that fell beyond 2.5SDs of the total mean (2.3 % of trials), mean response time (RT) was 519 ms (SE = 9). Mean error rate was 3.9 % (SE = .5 %). Overall, these data suggest good performance during the acquisition phase.

**Test phase: effect of shape cues.** Participants had faster (mean RT ± SE = 477 ± 14 ms) and more accurate (mean error rate ± SE = 4% ± 1%) responses with valid shape cues than with invalid shape cues (RT = 536 ± 12 ms, percent error = 7% ± 1%). The effect of shape cue was significant on both RTs [t(20) = 6.04, SE = 9.72, p < .001] and error rates [t(20) = 3.73, SE = .010, p < .01], which suggests the shape cues successfully biased performance. On the other hand, with a mean accuracy of 92% on trials with invalid shape cues (all participants had accuracies above 78% on these trials), I reasoned that responses were not solely made based on shape cues.
**Test phase: effect of colour transients.** RT data were submitted to a 3 x 2 repeated-measures ANOVA, with colour (congruent, incongruent, or neutral) and location (valid vs. invalid) of the transients as factors (see Figure 5). This analysis did not reveal a main effect of colour \( [F(2,40) = 2.27, \ p = .116, \ \eta_p^2 = .102] \), but revealed a marginally significant effect of location \( [F(1,20) = 3.89, \ p = .063, \ \eta_p^2 = .162] \), and a significant two-way interaction \( [F(2,40) = 10.04, \ p < .001, \ \eta_p^2 = .333] \). When the colour was congruent with the shape cue, location validity had a robust effect, i.e., faster responses with valid cues (475 ± 14 ms) compared to invalid cues (507 ± 15 ms, \( t[20] = 4.59, \ SE = 7.06, \ p < .001 \)). By contrast, incongruent (\( p = .43 \)) or neutral (\( p = .18 \)) colour transients did not produce a reliable effect.

In analyzing identification error data, missed trials were excluded (anticipations and late responses; these constitute .3% of the total trials). The remaining error rates (4%) were arcsine transformed and analyzed using the same ANOVA. The main effects of cue colour and cue location did not reach significance (\( F \) values < 1), but a marginally significant interaction was found \( (F[1,40] = 2.69, \ p = .08, \ \eta_p^2 = .108) \). Examining this trend further showed the same pattern as the RTs. When the colour transient was congruent with the shape cue, its validity produced a significant reduction in errors (2.8% and 5% for valid and invalid trials, respectively, \( t[20] = 3.18, \ SE = .007, \ p < .01 \)). No such cueing effect was observed with incongruent (\( p = .98 \)) or neutral (\( p = .27 \)) colour transients. The similar patterns in RTs and errors, rule out the possibility of a speed-accuracy trade-off.

Next, I submitted percentage of missed-trial errors (arcsine transformed) to the same ANOVA. This analysis did not show a main effect of cue colour \( (F[2,40] = 2.15, \ p = .130, \ \eta_p^2 = .097) \), cue location \( (F < 1) \), or a two-way interaction \( (F < 1) \).
Figure 5. Response time data from Experiment 2, graphed as a function of the colour and location of the spatial cue. Cue colour, in relation to the primed target shape, could be congruent, incongruent, or neutral. Cue location, in relation to the target location, could be valid or invalid. Error bars represents 95% within-subjects confidence intervals.

Similar to Experiment 1, I divided participants into two groups of ten (excluding the median) based on the strength of shape cueing effect (invalid RT – valid RT), with a "strong shape cueing" group (M ± SE = 96 ± 8 ms) and a "weak shape cueing" group (21 ± 9 ms). I submitted the RT data to a mixed 3 x 2 x 2 ANOVA with colour congruency and location validity as the within-subjects factors and the strength of shape cueing as the between-subjects factor (see Figure 6). The main effect of group (F[1,18] = 3.10, p < .095, ηₚ² = .147) did not reach significance. Furthermore, this analysis revealed a main effect of location [F(1,18) = 4.73, p = .043, ηₚ² = .210] and a marginally significant group x location interaction [F(1,18) = 3.76, p = .068, ηₚ² = .208]. As Figure 6 shows, these effects are driven by a general orienting effect by all three colour transients in the "weak-shape-cueing" group [mean valid RT ± SE = 515 ± 21 ms; invalid RT = 534 ± 21 ms, t(9) = 3.02, SE = 6.26, p = .014]. By contrast, no such orienting pattern was obtained in the "strong-shape-cueing" group [valid RT = 472 ± 21 ms; invalid RT=
$473 \pm 21$ ms, $p > .8$). I also found a congruency x validity interaction, matching the results of the initial analysis. The three-way interaction did not reach significance [$F(2,36) = 2.30$, $p = .118$, $\eta_p^2 = .113$].

**Figure 6.** Response time data from Experiment 2, divided based on the strength of the effect of the shape cue. The upper panel represents the group with a strong shape cueing effect, whereas the lower panel represents the group with a weak shape cueing effect. Error bars represent 95% within-subjects confidence intervals.
The results of this experiment are consistent with a relationship between using the shape cue to prime one of the possible sequences of stimuli and inhibiting the colour feature that is not identical to the primed sequence. Spatial orienting effects by congruent colours were similar across weak and strong shape cueing groups. By contrast, incongruent colour transients had different effects across weak and strong shape cueing groups. Although this is consistent with the idea that the source of these effects is the inhibition of colours incongruent with the activated stimulus-outcome sequence, the effect of neutral transients does not support an account that is purely based on selective inhibition of incongruent colours. Such an account cannot explain why strong shape cueing was also associated with a disadvantage for neutral colour transients. It is likely that facilitated processing of congruent colours might also play a role in the congruent-incongruent difference.

**Experiment 3**

The findings thus far suggest that the differential treatment of congruent and incongruent colours relies on the learned target-colour associations. We are, however, faced with two possible interpretations of these findings, one of which is that this effect may depend on seeing the shape-colour conjunction. Participants saw the target in the congruent colour after performing a correct response. Does this shape-colour conjunction play a critical role in the associative learning and the derived attention to colours? If not, then it is possible that simply completing the sequence of shape-response-colour is sufficient for altering the attentional treatment of colours. This latter possibility is consistent with previous work on learning sequences of perceptual elements and the ability to predict the next stimulus based on the present (e.g., Remillard, 2003). The former possibility would be consistent with the notion that continually encountering feature conjunction will lead to formation of multi-featured object representations, which then drive the attentional modulation effect. To discriminate between these two accounts, I modified Experiment 1 such that the shapes were now presented only for 100 ms (unmasked). With this modification, the target shapes were still consistently followed by colours (immediately after a response was performed), but the two never occupied the same location at the same time. Although temporal proximity of features can still lead to their inclusion into a single object representation, it is thought that the proximity needs to be shorter than 350 ms (e.g., Zmigrod & Hommel, 2011). Binding would be unlikely to occur because
subjects do not see the colour associated with a target shape appear until after this critical period (given the average RT was around 500 ms). Therefore, replicating the pattern found in the previous experiment would support the sequential-learning account, while not finding the same pattern would support an account based on visual object representation.

The stimuli, procedure, and design were identical to Experiment 1, with the following exception. The letter stimuli (‘S’, ‘H’, ‘E’, and ‘U’) appeared for 100 ms only (instead of remaining until response), during both acquisition and test phases. After performing a response, the remaining items on the display (the two placeholders and fixation cross) changed colour. Participants were twenty-two University of Toronto undergraduate students who reported normal or corrected-to-normal vision and were unaware of the purpose of the study.

**Acquisition phase.** After excluding incorrect responses and responses faster than 100 ms and those that fell beyond 2.5SDs of the total mean (2.7 % of trials), mean response time (RT) was 540 ms (SE = 9). Mean error rate was 6.3 % (SE = .8 %). Overall, these data suggest good performance during the acquisition phase.

**Test phase: effect of shape cues.** Participants had faster and more accurate responses with valid shape cues (mean RT ± SE = 516 ± 13 ms, error rate ± SE = 5% ± 1%) than with invalid shape cues (RT = 534 ± 15 ms, error rate = 4% ± 1%). The effect of shape cue was significant on RTs (t[21] = 4.93, SE = 6.73, p = .01], but not on errors (t[21] = 1.61, SE = .005 , p = .121). RTs were superior with valid shape cues compared to invalid shape cues. On the other hand, with a mean accuracy of 95% on catch trials (all participants had accuracies above 87% on catch trials), I reasoned that responses were not made solely based on the shape cue.

**Test phase: effect of colour transients.** Mean RTs were submitted to a 3 x 2 repeated-measures ANOVA, with colour (congruent, incongruent, or neutral) and location (valid vs. invalid) of the spatial cue as factors (Figure 7). This analysis did not reveal a main effect of colour (F[2,42] = 2.26, p = .12, ηp² = .097), but did reveal a significant effect of location (F[1,21] = 4.28, p = .05, ηp² = .17). Most importantly, there was no interaction (F < 1). Both congruent and incongruent colour transients caused attentional orienting, with faster responses after valid transients (520 ± 14 ms) relative to invalid transients (531 ± 13 ms).

Identification error rates (arcsine transformed) were submitted to the same ANOVA. Neither the main effects nor the interaction reached significance (F values < 1.3, p values >.25).
Thus, identification errors (4.7% ± .7%) did not vary with the independent variables. Next, I analyzed percentage of missed trials (.4% in total). This analysis also revealed no main effect or interaction (F values < 1).

Once again, I divided participants into two equal groups based on the magnitude of shape cueing effect (invalid RT - valid RT), with a "strong shape cueing" group (mean cueing effect ± SE = 42 ± 8 ms) and a "weak shape cueing" group (-4 ± 5 ms). I then submitted the RT data to a mixed 3 x 2 x 2 ANOVA with colour and location as the within-subjects factors and the strength of shape cueing as the between-subjects factor. Unlike previous experiments, the main effect of group did not reach significance (F[1,20]=1.18, p=.29). I also did not find an interaction between group and any other factor (F values < 1). Instead, I only found a marginally significant effect of cue location (F[1,20] = 3.67, p < .07, ηp² = .155), similar to the original analysis.

![Graph](image.png)

Figure 7. Response time data from Experiment 3, graphed as a function of the colour and location of the spatial cue. Cue colour, in relation to the primed target shape, could be congruent, incongruent, or neutral. Cue location, in relation to the target location, could be valid or invalid. Error bars represents 95% within-subjects confidence intervals.

The findings of the present experiment revealed no difference between the effect of congruent and incongruent colour transients in orienting attention. Not only were the effects of
congruent and incongruent colour transients indistinguishable, the findings were also not modulated by the strength of shape cueing (i.e., strong vs. weak shape cueing groups). This leaves us with two conclusions. First, complementary to the findings of Experiment 2, the difference in attention to colours is attributed to shape-colour associative learning (and not, response-colour association). Second, the effect seems to rely on perception of shape-colour conjunction (as an object) instead of disjunctive shape-colour sequences. Of course, at this point we cannot determine whether the learning of shape-colour sequence did not take place or whether this learning did not affect attention to colours. Although participants encountered the shape-colour sequences (Experiment 3) with the same degree of consistency as they encountered the shape-colour conjunctions (Experiments 1-2), it appears that encountering the sequences failed to either a) give rise to shape-colour sequence learning, or b) give rise to a robust-enough associative learning, such that a bias for shape would modulate attentional bias for colours.

Two observations are worth noting. First, the effect of the shape cue was significantly reduced in Experiment 3 (18 ± 7 ms), compared to Experiments 1 (64 ± 13 ms, t[40] = 3.2, p < .01) and Experiment 2 (61 ± 10 ms, t[41] = 3.46, p < .01). Second, the overall effect of colour transients in spatial orienting was significantly reduced in Experiment 3 (12 ± 6 ms) compared to Experiment 1 (43 ± 11 ms, t[40] = 2.73, p < .01), though the effect did not significantly differ across Experiments 2 and 3 (27 ± 14 ms, t[41] = 1.16, p = .25; nor did this effect differ across Experiments 1 and 2, p = .42). Although this could simply be due to individual differences across the experiments, it may also be a consequence of the modified presentation method. Namely, in Experiment 3, targets were characterized by brief exposure to new objects inside placeholders (i.e., rapid onset and offset of new objects). As such, participants' sensitivity to events that are not characterized by brief exposure of new objects might have been reduced (Burnham, 2007; colour transients and shape cues both consisted of transformations in already-present visual objects, as opposed to onset and offset of new objects). Nonetheless, given that the effect of transient cues was not completely eliminated in Experiment 3 we could still draw the conclusion that congruent and incongruent colours were treated in a similar way.

The small effect of the shape cue in Experiment 3 is perhaps more problematic. One could argue that because shape cues were not utilized reliably, the congruent and incongruent colour transients were also not treated differently. If that is true, then some indication of a difference between congruent and incongruent colours should be present in the subset of the participants
who did show a relatively strong effect of the shape cue. The shape-cueing effect in those participants (Cohen's $d = 1.72$) was not weaker than the overall shape-cueing effects obtained in Experiments 1-2 (Cohen's $d$ values = 1.09 and 1.34, respectively). Even if we confine analysis to those eleven participants, the effects of congruent and incongruent colours would not only remain indistinguishable ($t[10] < 1, p = .84$), but also highly correlated ($r = .70, p = .016$). In short, the findings of this experiment suggest the learned associations in the present paradigm require the conjunctive presentation of the shape and colour (i.e., a single object that contains both features).

**Discussion**

The present study examined the effect of associative learning on feature-based visual attention. Orienting of attention toward transient events is an extensively studied phenomenon, although the role of associative learning has not been examined in this paradigm (cf. Kristjánsson & Nakayama, 2003). The majority of previous research on attentional orienting paradigms attributes the effect of a transient to stimulus properties (Theeuwes, 2010), goal-driven processes (Folk & Remington, 2006; Egeth & Yantis, 1997), or repetition effects (Awh et al., 2012; Theeuwes, 2010). Importantly, the benefit of repetition is often restricted to a single feature (Kristjánsson, 2006) or a single dimension (Found & Müller, 1996; Müller et al., 2003). Here, I argue that a learned association between the shape and colour extended the bias for shape to a bias for colour. While expecting a target shape, participants showed reliable attentional orienting to the colour transient associated with the target shape while showing no attentional orienting with the colour associated with the alternative target shape. Experiments 2-3 confirmed that the advantage of congruent colours over incongruent colours was indeed due to learning shape-colour associations and was not driven by a response-colour association. Furthermore, Experiments 2-3 indicated that this bias depended on having viewed the two features in conjunction, as part of the same object, and not as disjunctive events of a sequence.

It is worth noting that these novel attention findings have some precedence in the learning literature from Fiser and Aslin (2001), who demonstrated associative learning between visual items in an implicit learning paradigm. In their experiments, sets of (two) distinct shapes could consistently co-occur in the same spatial arrangement and recognition performance for the
associated sets was not only better than chance, but it could not be solely attributed to recognition of individual items. It would seem reasonable that associative learning of this kind should also influence attention (i.e., attention to one element of the set may bias attention toward the other element). Kristjánsson and Nakayama (2003) did not find such an effect, perhaps because their features of interest were both task-irrelevant and did not receive the same degree of attention as the stimuli in Fiser and Aslin. In the present study, I implemented an initial acquisition phase that was easy to perform and consisted of two hundred trials, in order to maximize the possibility of associative learning. In addition, colour features were not entirely task-irrelevant by virtue of being associated to the visual targets (shapes). These findings also fit nicely with an event-based view of attentional selection (Hommel, 2004). Once a stable representation of an event is formed, activating one feature will activate other features belonging to the same event file.

Given that presenting shape and colour as two temporally separate visual events did not give rise to the same finding, it seems that conjunctive presentation of shape-colour objects is essential in observing modulated bias for colours. I propose that what participants really acquire during the acquisition phase were not two sets of sequences, but two distinct visual objects. Assuming the object representations included both shape and colour, anticipating one object feature would necessarily have predictive consequences regarding other features. One could describe the cause of association as the Gestalt principle of "common fate". Although I would not object to this description, I should hasten to point out that the "common fate" here does not mean belonging to the same sensorimotor sequence (i.e., same trial), but the same visual object (i.e., spatiotemporal coincidence).

In light the findings of this chapter, regarding objects as the units of attention, as opposed to features or locations, may be more appropriate (Duncan, 1984). An object-based framework can offer novel interpretation with regard to some of the previous findings. Namely, Kristjánsson et al. (2001) found that if the visual target is consistently presented, say, in the left portion of the cue, participants rapidly learn to select the target-containing portion of the cue. One could argue that participants were forming representation of a single visual object (consisting of the cue and the target), whose center of mass, due to the target, was shifted leftward. After acquiring this representation, even viewing a portion of the object (the cue alone) would activate the entire
object representation, causing attentional orienting toward the object's center of mass (Kravitz & Behrmann, 2008).

It is generally thought that the neural correlates of multi-featured object representations are situated along the ventral visual pathway (e.g., the perirhinal cortex, Barense et al., 2005; 2007). Assuming that motor interaction with visual objects would reduce the contribution of the ventral pathway to visual processing (increasing the contribution of the action-oriented dorsal pathway), I predict that the shape-colour associative learning and its attentional consequences will be weaker in conditions where participants engage in motor actions directed at the visual objects (Ganel & Goodale, 2003; Goodale, Cant, Króliczak, 2006; Gozli et al., 2014).

The present findings included virtually no evidence of associative learning between responses and colours. If such evidence was available, what would it look like? A strong response-colour association could have eliminated, or reversed, the difference between congruent and incongruent colours, in Experiment 2. That is because the response-colour mapping was reversed after the acquisition phase. In the case of Experiment 3, a strong response-colour association could have resulted in a replication of the pattern found in Experiments 1 and 2.

It is possible that strong response-colour associative learning simply did not take place. Although there is evidence that action-outcome learning can be acquired with exposure to consistent mapping (e.g., Elsner & Hommel, 2001; Hommel, 2004), as argued by Herwig et al. (Herwig, Prinz, & Waszak, 2007; Herwig & Prinz, 2009), such association is weakened when participants focus on the relationship between stimulus-response mapping (i.e., the first two elements in the stimulus-response-outcome sequence). Herwig et al. (2007) proposed a distinction between sensorimotor and ideomotor modes of action. Whereas the former mode is characterized by a focused on the stimulus, to which an appropriate response should be made, the latter mode is focused on planned outcomes, and the appropriate response that would bring about the outcome (see also, Pfister, Kiesel, & Hoffmann, 2011). It is possible that our task encouraged a dominantly sensorimotor mode of acting, resulting in attenuating the action-outcome link. Given that response-outcome associative learning is the focus of the following chapter, the experiments reported in Chapter 3 employ voluntary responses in order to maximize the likelihood of response-outcome associative learning.

It is possible that although the response-colour associations were learned in the present experiments, the shape cues did not induce a bias in action selection despite being an effective
source of perceptual bias. Although the shape cues were 80% valid, participants still had to wait before determining the correct response. Consistent with this possibility, Ziessler and Nattkemper (2011) showed that learned action-outcome association impact behaviour after an action has been selected. Prior to action selection, on the other hand, the learned association between action and outcome did not seem to exert an impact. I will return to this issue in Chapters 4 and 5. At present, it should be noted in the present experiment there is no evidence that shape cues induced an action bias. For this reason, even if colours and responses were to some degree associated, this association did not impact performance.

In designing of the test phase, I also implemented "neutral" (blue) transient cues in order to assess whether different effectiveness of congruent and incongruent colours were due to facilitation (of the congruent colour) or inhibition (of the incongruent colour). Interpreting the effect of the neutral transient cue seems rather difficult in light of the different pattern of findings in Experiments 1 and 2. Whereas in Experiment 1, neutral and congruent cues were quite similar, in Experiment 2 the effect of neutral cues were more similar to the effect of incongruent cues. In explaining the discrepancy between the two findings, we should also keep in mind that the neutral colour was also less frequently presented and may, therefore, have had a benefit due to its relative novelty. Given that in Experiment 1 the target-response mapping was consistent across the acquisition phase and the test phase, the novelty of the neutral cue might have given it a selective advantage. By contrast, switching the target-response mapping across the two phases of the experiments means that the neutral colour is no longer the only novel aspect of the test phase. Regardless, based on the effect of the neutral cues alone, we cannot attribute the different effect of congruent and incongruent cues exclusively to facilitation or inhibition of the latter.

To conclude, the present study adds to the growing body of research that attempts to place attentional operations in a broader context that includes learning and memory (Chun & Turk-Browne, 2008; Hutchinson & Turk-Browne, 2012; Stokes, Atherton, Patai, & Nobre, 2012). The present findings point out the benefit of regarding multi-featured episodes as units of attention rather than units consisting of single features. Considering the role of prior knowledge and the complexity of events in our visual world would provide a more accurate picture of how the attentional system works in everyday situations.
Chapter 3: Sensorimotor Associative Learning (I):
Response Execution Can Guide Feature-Based Attention

Our actions, including eye- and hand-movements, continually affect the sensory input we receive from the environment. In this chapter I turn to the possibility that learning the regularities that connect our movements to their corresponding sensory outcome would serve as a guide for visual attention. It is generally believed that action-outcome associative learning allows the perceptual systems to discriminate self-caused sensory events from events caused by external sources (e.g., Hommel et al., 2001; O'Regan & Noë, 2001; Shin et al., 2010). Indeed, immediately after performing an action, a stimulus that matches the learned action-outcome typically elicits a weaker response, compared to a physically identical stimulus that does not match the learned action-outcome (e.g., Blakemore et al., 1998; Bompass & O'Regan, 2006; Cardoso-Leite et al., 2010). This finding is described as sensory attenuation and has been explain in two different ways; first, it has been explained in terms of inhibition of the predicted sensory response (Blakemore et al., 1998). Second, and paradoxically, sensory attenuation has also been explained in terms of preactivation of the predicted sensory response (Waszak, Cardoso-Leite, & Hughes, 2012). Although the two accounts are inconsistent, they both explain the phenomenon of sensory attenuation in a wide range of experimental settings. Discriminating between the two accounts is central to the present investigation, as it represents an unavoidable step in connecting the function of attention with associative learning. I begin the present chapter, therefore, with a brief review the two accounts.

The first account of sensory attenuation assumes that the predicted sensory action-outcome are inhibited (Blakemore et al., 1998; Miall & Wolpert, 1996). According to this inhibition account, a comparator mechanism subtracts the actual sensory outcome from predicted action-outcome. If the actual outcome matches the internally generated prediction, its representation will be weaker. If, on the other hand, the stimulus does not match the internal prediction its representation will be left intact (Figure 8a). Empirical findings consistent with the inhibition account include reduced sensitivity for a learned sensory action-outcome, compared to a stimulus that is not associated with the performed action.

In a series of psychophysical studies, Cardoso-Leite et al. (2010), and Roussel et al. (2013) had participants associate key-press actions with distinct visual action-outcomes during
an initial acquisition phase (e.g., gabor stimuli with left/right orientation, respectively, linked to left/right keypress). Next, in a test phase, keypress actions continued to produce sensory outcomes, although the outcomes were now either congruent or incongruent with the learned contingencies in the acquisition phase. An "action-congruent" stimulus, in the present context, is defined as an outcome whose feature matches the action-outcome association learned during the acquisition phase. In the test phase, participants reported the presence (Cardoso-Leite et al., 2010) or the brightness level (Roussel et al., 2013) of stimuli as a function of action-congruency. Both studies found reduced visual sensitivity for action-congruent stimuli, relative to incongruent stimuli. Using a similar method, Pfister et al. (2012) found that even preparing an action, without performing it, can interfere with detection of learned action-outcomes (see also, Bompass & O'Regan, 2006; Müsselfer & Hommel, 1997; Stenner et al., 2014).

![Inhibition Hypothesis](image)

**Figure 8.** Panel (a): according to the inhibition hypothesis, correct prediction of the cue feature reduces signal strength caused by the cue, which should reduce the effect of the both valid and invalid cues in our paradigm. Panel (b): according to the preactivation hypothesis, internal prediction increases the baseline activity for the predicted feature, which should strengthen the effect of valid cues. Figure reprinted from Roussel et al. (2013) with the publisher's permission.
Consistent with the psychophysical findings, event-related potentials associated with early auditory response (e.g., Bäß, Jacobsen, & Schröger, 2009) and early visual response (e.g., Roussel et al., 2014; Kimura & Takeda, 2014) have been found to be weaker for a learned action-outcome, compared to a stimulus that mismatched the action-outcome (see also, Hughes & Waszak, 2011). Again, these findings fit the inhibition account, which assumes that action-outcome associative learning enables inhibition of anticipated action-outcomes.

The second account of sensory attenuation is based on internal preactivation of learned sensory outcomes. According to this preactivation account, planning an action is thought to increase activity of the cells that represent the anticipated sensory outcome (Roussel et al., 2013; Waszak et al., 2012). This sensory preactivation, in turn, limits any further increase in the cell's activity that is uniquely caused by external stimulation (Figure 8b). Unlike the inhibition account, the preactivation account is not based on subtraction of two signals, but is based on the difficulty in detecting an external signal that is received during higher baseline activity. Akin to Weber's law, a stimulus that is encountered during preactivation evokes a weaker response, relative to a stimulus that is encountered against a lower baseline, i.e., without preactivation. Given that the preactivation account also predicts weaker stimulus representation for action-outcomes, this account equally well fits the psychophysical observations of sensory attenuation (Bompass & O'Regan, 2006a; Bompass & O'Regan, 2006b; Cardoso-Leite et al., 2010; Roussel et al., 2012).

Critically, the preactivation account makes the additional prediction that the representation of sensory action outcome, although attenuated in strength, is formed faster due to the processing head-start provided by preactivation (Hughes, Desantis, & Waszak, 2013). The inhibition account, by contrast, predicts no processing speed advantage for action outcomes compared to action-incongruent colours. In the present study, I employed a visual task where the inhibition and the preactivation accounts offer conflicting predictions based on their assumptions with regard to processing speed of action-congruent and -incongruent stimuli.

The present experiment involves a visual search task in the presence of a salient cue. Cue colours were learned sensory outcomes of observers' keypress actions. I should note that previous methods of investigation have typically associated action with a feature of the target that remains central to the task in the test phase. As such, they are primarily sensitive to the strength of stimulus representation (i.e., salience), and how it might change as a function of
known action outcomes (e.g., Bompass & O'Regan, 2006; Cardoso-Leite et al., 2010; Hughes & Waszak, 2011; Stenner et al., 2014, although see, e.g., Gozli, Goodhew, Moskowitz, & Pratt, 2013; Gozli & Pratt, 2011; Kumar, Manjaly, & Sunny, 2015). By contrast, by associating a cue feature with actions, the present paradigm affords sensitivity to both the salience of representation and the speed with which that representation is formed.

Sensitivity of the present paradigm to both salience and processing rate rests on the fact that we can investigate the effect of cues when they are spatially valid (indicating target location) and when they are spatially invalid (indicating a distractor location). On the one hand, invalid cues allow for assessing the effect of feature salience, because lower salience of a cue feature allows faster attentional disengagement from the cue (Theeuwes, 2010). On the other hand, valid cues allow for assessing the effect of processing rate, because faster processing rate of a cue feature allows faster selection of the valid cue (Bundesen, 1990). The inhibition account would predict that action-congruent colours will have lower salience and lower processing speed, compared to incongruent colours. Accordingly, compared to action-incongruent cues, action-congruent cues will be easier to disengage from when invalid, and they will be selected more slowly when valid. By contrast, preactivation account holds that action-congruent colours will have lower salience but higher processing rate, compared to action-incongruent cue. Therefore, this account predicts that action-congruent cues will be easier to disengage from when invalid, and they will be selected faster when valid (Hughes et al., 2013; Waszak et al., 2012).

Similar to previous work on sensory attenuation, the present experiment consisted of an acquisition phase and a test phase. During the acquisition phase, participants' keypress determined the colour of the cue (one key producing red, the other producing green), while cue location was randomly selected from a set of four placeholders. Next, in a test phase, the same keys were followed by a cue and a search display (Figure 9). In the test phase, if a red cue appears after pressing the "red" key (i.e., the key that consistently produced red cues during acquisition), the cue colour is regarded as action-congruent, whereas it is regarded as action-incongruent if it appears after pressing the "green" key. In addition to manipulating whether cue colour was congruent with the learned action-outcome associations, I also manipulated the spatial validity of the cues. For one group, the cue was always invalid (indicating a distractor location), whereas for the other group, the cue was always valid (indicating the target location).
In summary, the inhibition account and the preactivation account both predict a smaller cost for action-congruent cues in the invalid-cue condition, but they make opposite predictions in the valid-cue condition. With valid cues, the inhibition account predicts smaller cuing with the action-congruent colour, whereas the preactivation account predicts larger cuing with action-congruent colour due to the speeded processing caused by feature preactivation.

Experiment 4

Participants. Thirty-six University of Toronto undergraduate students participated in the experiment in exchange for course credit (18 per condition). They all reported normal or corrected-to-normal vision, and were unaware of the purpose of the study. All experimental protocols were approved by the Research Ethics Board of the University of Toronto.

Stimuli. The display structure and a sample sequence of events are shown in Figure 9. Except for the colour cues, all stimuli were presented in white (CIE XYZ = 33.60, 26.59, 93.80) against a black (XYZ = 0, 0, 0) background. The cue colours were red (XYZ = 41.24, 21.26, 1.93) and green (XYZ = 6.72, 13.43, 2.24). Placeholders were four squares (2.4° x 2.4°; frame width = .16°) that appeared above, below, left, and right of the display center (distance from center = 8°). When a placeholder turned into a colour cue, its frame width increased to .24° of visual angle. The target was a tilted line ("\" vs. "/"; length = 1.4°; width= .1°) that would appear inside one placeholder. Each distractor was a letter "X" that appeared in a non-target placeholder. Participants performed two types of responses. The responses that produced visual effects were performed using the index and middle fingers of left hand and the 'Q' and 'A' buttons on the keyboard. I associated left-hand responses with action-outcomes on the basis of prior research that suggests action-outcome associative learning may be stronger for left-hand responses (Melcher et al., 2008; 2013). The search responses were performed with the right hand, using the left and right arrow keys (in response to "\" and "/" target, respectively).

Acquisition Phase. Each trial of the acquisition phase began with the presentation of the fixation cross and the four placeholders. After a random delay of 1000-1500 ms, the fixation cross flickered (i.e., it disappeared for 100 ms and then reappeared). I instructed participants to press either the 'Q' or the 'A' key upon noticing the flicker. Moreover, I instructed them to make their selection spontaneously, try to avoid patterns, and try to select the two keys equally
frequently. As soon as a keypress was recorded, the colour cue appeared at one of the placeholders. The colour of the cue was determined by the response (red or green after 'Q' or 'A', respectively). The location of the cue was randomly chosen, as any of the four placeholders were equally likely be the cue location. The cue remained on display for 300 ms, after which the next trial began. If participants pressed a key other than 'Q' or 'A', or if they pressed more than one key, they received visual feedback ("MISTAKE!"). If the response was faster than 100 ms, they also received visual feedback ("TOO FAST!"). No colour cue was presented on error trials. The search task was not included in this phase.

![Figure 9. Sequence of events in a sample trial of Experiment 4. The sample trial is from the test phase of the "invalid cue" condition.](image)

**Test Phase.** Similar to the acquisition phase, every test trial began with the presentation of the fixation cross and the four placeholders (Figure 9). Participants performed their first, left-hand response when they noticed the fixation mark flicker. There were three equiprobable types of trials, based on the colour cue. The cue could be congruent with the response-outcome associations during the acquisition phase, incongruent, or absent (i.e., no cue/action-outcome). After a 100-ms delay, the cue was followed by the appearance of the search items in the placeholders. The cue and the search display remained on display until a response was recorded.
For the search display, I instructed participants to find the tilted line among the distractors ("X"s) and identify the target tilt using the left/right arrow keys. Upon pressing an incorrect key, or pressing more than one key, participants received a visual feedback ("MISTAKE!"). Finally, if the first response was a mistake, participants received visual feedback and no search display or cue was presented.

**Design.** Participants were randomly assigned to the valid-cue or invalid-cue condition. The two conditions had the same acquisition phase. In the test phase, the cue was either valid or invalid. An invalid cue never coincided with target location and had to be ignored. A valid cue always coincided with target location and, thus, had to be selected. Each participant completed 200 trials in the acquisition phase and 128 trials in the test phase. Each phase was preceded by 15 practice trials.

**Results**

**Acquisition.** Before calculating mean response times (RT), I excluded error trials and trials in which RT fell 2.5 SD beyond the total mean. Mean RT and percentage of errors (PE) in the acquisition phase were 379 (SE = 63 ms) and 3.0% (SE = .5%). Furthermore, participants selected the two keys with equal frequency (51% and 49%, respectively, for the 'Q' and 'A' keys, t[35] =1.71, SE = .02, p = .10).

**Test.** For the first voluntary keypress made with the left hand, mean RT and PE in the test phase were 352 ms (SE = 59) and 2.6% (SE = .4%). Furthermore, participants continued to select both keys, although they slightly favoured the 'A' key, corresponding to the green cue, over the 'Q' key, corresponding to the red cue (44.8% and 55.2%, t[35] = 2.16, SE = .05, p = .038).

For the search task, mean RTs were submitted to a 3 x 2 mixed ANOVA with Cue Colour (absent, congruent, or incongruent) as the within-subject factor and Cue Validity (valid vs. invalid) as the between-subject factor (see Figure 10). This analysis revealed a marginal main effect of Cue Validity (F[1,34] = 3.90, p = .056, η² = .103), a main effect of Cue Colour (F[2,68] = 9.28, p < .001, η² = .214), and a two-way interaction (F[2,68] = 27.65, p < .001, η² = .449). The interaction indicates the rather trivial fact that the effect that the cuing effect depended on cue validity. That is, when comparing cue-present trials (including both cue colours, i.e., two-
thirds of the trials) with cue-absent trials (the remaining one-third), I found a 41-ms benefit for valid cues (Cohen's d = 1.43), and a 17-ms cost with invalid cues (Cohen's d = .72).

Since our primary interest is in the difference between action-congruent and -incongruent cue colours, I submitted the RT data to a 2 x 2 ANOVA, leaving out the absent trials. This analysis revealed a significant main effect of Cue Colour ($F[1,34] = 12.43, p = .001, \eta_p^2 = .268$). Regardless of Cue Validity, an action-congruent cue colour ($548 \pm 21$ ms) led to faster responses, compared to an action-incongruent cue colour ($M \pm SE = 562 \pm 22$ ms; Cohen's d = .59, for the main effect of Cue Colour). Most importantly, there was no two-way interaction ($F[1,34] = .008, p = .931, \eta_p^2 < .001$), showing that congruent colours conferred a benefit both when they were valid and when they were invalid. Lastly, I also found a main effect of Cue Validity ($F[1,34] = 6.68, p = .014, \eta_p^2 = .164$), based on faster responses with valid cues ($512 \pm 28$ ms) compared to invalid cues ($597 \pm 19$ ms).

**Figure 10.** Response time data from Experiment 4, graphed as a function of cue colour (absent, action-congruent, or action-incongruent) and cue validity (valid vs. invalid). Percentage of errors for each condition is presented at the base of the corresponding bar graph. Error bars represent 95% within-subjects confidence intervals.

I submitted mean PEs from the search task to the same 3 x 2 ANOVA, which did not reveal a significant effect of Cue Colour ($F[2,68] = .30$), Cue Validity ($F[1,34] = 1.59, p = .21$,
\(\eta_p^2 = .045\), or a two-way interaction (\(F[2,68] = .29\)). Error rates did not significantly differ across trials with congruent (5.1% ± 1.8%) and incongruent cues (4.5% ± 1.1%, \(p = .50\)), inconsistent with the possibility that the RT difference was a speed-accuracy trade-off.

**Discussion**

Sensory events that are self-caused typically evoke an attenuated response, compared to physically identical sensory events that are caused by external sources (e.g., Blakemore et al., 1998; Cardoso-Leite et al., 2010; Hughes et al., 2013). In the present chapter I investigated two accounts of sensory attenuation (inhibition vs. preactivation) using a visual search task in which the colour of a salient cue could be linked to participants' own action. In an initial acquisition phase, participants learned that red and green cues consistently resulted from their own key-press actions. In the subsequent test phase, cues continued to result from participants' action, but their colours could be congruent or incongruent with the action-outcome contingencies learned in the acquisition phase. Action-congruent cues were easier to ignore in the invalid-cue condition, consistent with both the inhibition account and the preactivation account. More importantly, action-congruent cues were more effective in the valid-cue condition, suggesting that self-caused features can be selected faster. This finding is consistent with the preactivation account, according to which self-caused features receive a processing head-start that speeds their processing, despite reducing their salience.

These findings are compatible with a recent report by Desantis et al. (2014), who found better discrimination performance with action-congruent motion stimuli, compared to action-incongruent stimuli. I reason, in agreement with Desantis et al., that although actions attenuate responses to sensory action-outcomes, whether the action-outcome is at a perceptual advantage or disadvantage ultimately depends on specific task demands. Desantis et al. (2014) made the crucial point that the meaning of "signal" is different in detection and discrimination tasks. In a two-alternative forced-choice discrimination task, for instance, each perceptual judgment is the result of a competition between two possible sensory states. Therefore, the advantage of each signal is defined in relation to the competing alternative signal. In this context, preactivation would provide the action-congruent sensory state a competitive advantage over the alternative state. By contrast, in a detection task, where each perceptual judgment depends on the unique
contribution of the external stimulus, preactivation would put the action-congruent stimulus at a disadvantage by reducing the unique contribution of the stimulus (see also, Roussel et al., 2013). Applying a similar logic to the present study, I argue that action-congruent cues had an initial competitive advantage in processing speed, provided by the internal preactivation. This initial advantage increased the benefit of valid cues. The same preactivation, however, also reduced the salience of action-congruent cues, allowing for rapid disengagement in the invalid-cue condition. Thus, whether feature preactivation confers an advantage or a disadvantage depends on specific characteristics of each task.

The role of action in shaping perception has gained increased recognition (e.g., Hommel, 2009; O'Regan & Noë, 2001; Thomaschke, Hopkins, & Miall, 2012). Actions change the sensory input we receive from the environment, and it is reasonable to suppose the perceptual systems take advantage of the regularities that connect our actions with their corresponding sensory outcomes, in a similar manner that perceptual associative learning leads to more efficient allocation of attention (e.g., Eslner & Hommel, 2001; Hommel et al., 2001; Hommel, 2004; O'Regan & Noë, 2001; Shin et al., 2010). Indeed, evidence suggests that preparing an action not only involves activity in motor-cortical areas, but also activity in sensory areas that underlie the expected outcome of the action (e.g., Gutteling et al., 2015). In a study by Kuhn et al. (2010), participants learned that their keypress actions produced house or face images. After learning the action-outcome associations, participants' actions alone activated the sensory regions corresponding to visual processing of faces or houses. In a similar paradigm, Hughes and Waszak (2014) found modulated activity in the posterior visual areas, prior to action execution, depending on whether participants selected a face- or house-generating action. Thus, sensory brain regions are involved in the coding of actions, through the learned sensory outcomes of the actions.

In the previous chapter, I reported an instance of feature-based attention, wherein attention to a particular non-target feature (colour) was an involuntary consequence of selecting a target feature (shape) that was associated with it. The findings in the present chapter, on the other hand, represents another instance of feature-based attention driven by selection of a response. Here, attention to a particular non-target feature is an involuntary consequence of selection a response that is associated with the feature. Both sets of findings fit within an event-based account of attention, according to which the units of selection are multi-featured events that
could include both perceptual and motor response (Hommel, 2004; Hommel et al., 2001). Within an event coding framework, actions do not possess a unique status in their ability to modulate the effect of visual stimuli. An attempt to reduce the effect of actions to sensory components of the action remains consistent with the notion that sensory anticipation is essential to action representation (Hommel et al., 2001). According to such a view, the learned sensory outcomes of an action, collectively, constitute the representation of the action (Hommel et al., 2001; Shin et al., 2010). Features of an action are not inherently distinct from features of perceptual events. Thus, to ask whether actions are unique in their ability to serve as a source of visual bias goes against this fundamental assumption by restoring the strong distinction between action- and perception-related features.

In line with the findings of Chapter 2, I argue that similar visual biases could be observed without the involvement of actions. Indeed, in agreement with Waszak et al. (2012), I propose that the action-driven effects in the present study are similar to the effect of repeated exposure to the same feature (e.g., Awh, Belopolsky, & Theeuwes, 2012). Similar to our findings, ignoring a consistently-invalid peripheral cue has been shown to be more efficient if cue colour repeats across trials (cue colour being a task-irrelevant feature), while selecting a consistently-valid cue is also more efficient if cue colour repeats across trials (e.g., Pinto et al., 2005; Vatterott & Vecera, 2012). Waszak et al. (2012) have attributed the efficiency that comes with repeated exposure to a feature to sensory preactivation, suggesting that preactivation is not unique to action performance.

In conclusion, a response can serve as a source of visual bias by virtue of generating sensory preactivation of known action-outcomes. In line with sensorimotor accounts of vision, in the present chapter I showed how the visual system is sensitive to the learned associations that connect dynamic features of the action systems to sensory outcomes. More broadly, this finding should be considered as another instance of the guidance of attention to features by associative. I should highlight that this effect was observed at the post-execution phase of the response. The perceptual features of a response, however, do not require execution of the response to be activated. Indeed, the event-based accounts of action selection presume that anticipation of perceptual action outcome as a critical component of selection that enables voluntary movement (Hommel et al., 2001). In the next chapter, therefore, I turn to the issue of action-driven perceptual bias for features prior to the execution of response.
Chapter 4: Sensorimotor Associative Learning (II): Response Selection can Guide Feature-based Attention

In Chapter 3, I described an instance in which response execution can generate a bias for the associative visual feature. In the present chapter, I examine visual bias following response selection, but prior to the execution of the response. The idea that perceptual outcomes are activated during response selection is supported by numerous studies (Hommel & Elsner, 2009; Shin et al., 2010). Kunde (2001), for instance, found that selecting one of four keys was faster when each key generated a visual stimulus at a spatially corresponding location (e.g., far-left key generating a visual onset at the far-left placeholder, etc.) compared to when the keys generate spatially non-corresponding visual outcomes (see also, Hommel, 1993; Pfister et al., 2010). Similarly, Pfister et al. (2014) found faster performance when each keypress generates a vibrotactile stimulus to the responding finger, as opposed to the alternative finger. Beyond response-outcome location compatibility, Koch and Kunde (2002) showed performing a verbal colour-name response (e.g., "blue") was faster when each response generated the same colour name or the corresponding colour patch, compared to when the response generated a different colour name or a non-corresponding colour-patch. Kunde, Koch, & Hoffman (2004) found that faster performance when the intensity of the outcome stimulus (i.e., soft vs. loud tone) is compatible with the intensity of the response (soft vs. hard keypress). These findings provide strong evidence that even with an identical target stimuli, response selection is influenced by different anticipated response-outcome compatibilities. It should be noted that the response features in these studies are intrinsic to the response, such as location of the responding finger, intensity of the response, or the semantic meaning of the vocal response.

Further research has demonstrated that outcome anticipation can go beyond intrinsic action features to apply to novel action-outcome learning. In an experiment by Hommel (1996), participants learned that two keys, assigned as responses to two visual targets, resulted in two respective auditory outcomes. After learning the action-outcome association, the target stimuli were presented together with an auditory tone that either matched or mismatched the correct response on each trial. When the stimulus tone matched the outcome of the correct response, performance was faster and more accurate compared to when the two were incompatible. In a later experiment by Elsner and Hommel (2001), after learning a similar response-outcome
associations, the tones were used as target stimuli in the test phase. Consistent with outcome anticipation, participants for whom tone-response mapping was congruent with the learning phase showed better performance compared to those assigned to the incongruent condition (see also, Ziessler et al., 2004; Wolfensteller & Ruge, 2011). Supporting the event-based theory of action, these findings show that perceiving a learned action-outcome can bias selection of the corresponding action. It is relatively less clear, however, whether and how response selection – prior to response execution – would generate a perceptual bias in favour of the anticipated perceptual outcome. As discussed by Kunde et al. (2004) anticipation of a response outcome can, in principle, not only serve as confirmation that a desired act was performed (post-execution), but can also serve to select a desired action (pre-execution). Indeed, event-based theories of action assume bidirectional associations between movement and their known perceptual outcome, which could then function both in action selection and action correction processes (Hommel et al., 2001; Shin et al., 2010).

In order to assess the influence of response selection on visual bias, I implemented a method where each response is selected based on a first stimulus (i.e., a response cue) but it is executed based on a second stimulus (i.e., a go signal). Critically, the go signal could appear inside two placeholders. In each trial of the test phase, after the presentation of the response cue and prior to the presentation of the go-signal, the colour of both placeholders changed (Figure 11). One of the placeholders changed to the colour that is compatible with the selected response, while the other placeholder changed to the incompatible colour. I reasoned that if the compatible colour is activated due to response selection, then detecting the go signal would be faster when it appears inside the placeholder with the compatible colour, compared to the incompatible colour.

One factor that needs to be considered in this design is the time-course of response selection. It is safe to assume that response selection proceeds from a phase with low certainty to a phase with high certainty regarding the selected response. It is also safe to assume that once response selection is in a low-certainty phase, it is more susceptible to external, irrelevant sources of stimulation (Ansorge & Wühr, 2004; Ziessler & Nattkemper, 2011). Therefore, I expect that shortly after the presentation of the response cue, when uncertainty regarding the selected response is higher, the visual system is most susceptible to external, irrelevant sources of information that might reduce uncertainty regarding response selection. In this early phase, I expect response selection to induce a more robust visual bias. By contrast, after relatively longer
delays following the presentation of the response cue, uncertainty regarding the response is
eliminated and further external stimulation will be irrelevant for response selection. In this later
phase, I expect response to exert little or no impact on visual bias (cf., Ziessler & Nattkemper,
2011, Experiment 2).

Experiments 5 & 6

Participants. Fifty-two students at University of Vienna participated in this study (26 in
each experiment) in exchange for course credit or monetary compensation (5 Euros / 30
minutes). They all reported normal or corrected-to-normal vision, and were unaware of the
purpose of the study.

Stimuli. The display structure and a sample sequence of events are shown in Figure 11.
Except for the colour cues and response-outcomes, all stimuli were presented in white against a
black background. Target placeholders were two squares (1.4° x 1.4°; frame width = .1°)
positioned along the vertical midline. The center of these placeholders deviated by 3.4° of visual
angle from the display center. A response cue was a circles (radius = .7°; frame width = .05°;
center of circle deviating by 3.4° from display center) that could appear at the left or right of
fixation. Each response-outcome was a colour patch (red or green) that filled the response cue.
The target was a white square (.3 x .3) that appeared at the center of a target placeholder.
Responses were performed using the index fingers of each hand by pressing the "C" and "M"
buttons on the keyboard which, respectively, corresponded to "left" and "right" responses.

Acquisition Phase. Each trial of the acquisition phase began with the presentation of the
fixation cross and the two target placeholders (Figure 11). After a 1000-ms delay, a response cue
appeared at the left or right side of fixation, indicating the correct response on that trial.
Experiments 5 and 6 differed only with respect to their acquisition phase. In Experiment 5, I used
direct spatial S-R mapping. That is, participants in Experiment 5 were instructed to prepare the
"left" and "right" keys, respectively, upon seeing the left and right response cues. By contrast, an
inverse spatial mapping was used in Experiment 6, and participants were instructed to prepare
the spatially non-corresponding key, namely the "left" keys for the right response cue, and the
"right" key for the left response cue. Following the response cue, after a delay of 0, 100, 200,
400, or 700 ms, a target appeared inside a target placeholder above or below fixation. I instructed
participants to press the prepared key as soon as they detected the target, regardless of the target location. Furthermore, on 20% of trials no target was presented after a response cue (catch trial), and participants were instructed to withhold response on these trials. As soon as a keypress was recorded, the colour outcome corresponding with the performed keypress appeared inside the response cue and remained for 200 ms. The colour of the outcome was determined by the actual response regardless of its correctness. In both experiments, green and red were outcomes of left and right keypress responses, respectively. Given the spatially inverse mapping between response cue and response key in Experiment 6, and given that response outcomes always appeared inside the response cue, response-outcomes in Experiment 6 also appeared at the opposite location to the response. An error feedback appeared at the end of the trial if participants made an anticipatory response, defined as response times shorter than 100 ms ("TOO QUICK"), or if they responded with an incorrect key ("MISTAKE!"). An error feedback would remain on display for 2000 ms.

**Test Phase.** Similar to the acquisition phase, every test trial began with the presentation of the fixation cross and the two placeholders (Figure 11). After a 1000-ms delay, a response cue appeared at the left or right side of fixation, indicating the correct response on that trial. For the test phase, I used direct spatial S-R mapping in both Experiments. That is, participants were instructed to prepare the "left" and "right" keys, respectively, when they saw the left and right response cues. Following the response cue and a after delay of 0, 100, 200, 400, or 700 ms, the two target placeholders changed colour, one of them turning red and the other turning green. The assignment of colour to placeholders was randomized, and both colours and locations were equally likely to contain the target. The target appeared 100 ms after placeholders changed colours, except for catch trials in which no target was presented. When a keypress was recorded, the colour outcome corresponding with the performed keypress appeared inside the response cue and remained for 200 ms. The colour of the outcome was determined by the actual response regardless of its correctness. Response-colour mapping did not change across the two phases of the experiment (e.g., the "left" key continued to produce green), although the locations associated with each colour outcome were reversed across the two phases of Experiment 6. Similar to the acquisition phase, participants received error feedback after anticipatory or incorrect responses.
Figure 11. Sequence of events on a sample trial of the acquisition phase (top row) and a sample trial of the test phase (bottom row) of Experiments 5-6.

**Design.** Each participant performed 200 trials in the acquisition phase and 400 trials in the test phase. In both phases, 20% of trials constitute catch trials (i.e., the detection target was presented on these trials). In the test phase, the target was equally likely to appear inside the placeholder with the action-congruent colour or the action-incongruent colour. All experimental variables, including target location (above vs. below), response cue (left vs. right), and SOA were randomized within each experiment and had equiprobable values.

**Results**

**Acquisition.** Before calculating mean response times (RT), I excluded errors (3.2% of trials) and RTs that fell 2.5 SD beyond the total mean (2.8% of trials). Mean RT data were submitted to a 5 x 2 mixed ANOVA, with the SOA between response cue and target (0, 100, 200, 400, or 700 ms) as the within-subject factor and cue-key spatial mapping (Experiment 5:
direct vs. Experiment 6: inverse) as the between subject factor (Figure 12). This analysis revealed a main effect of SOA (F[4, 192] = 181.97, p < .001, ηp² = .79), a main effect of cue-key spatial mapping (F[1, 48] = 4.05, p < .05, ηp² = .08), and a two-way interaction (F[4, 192] = 5.12, p < .01, ηp² = .10). The main effect of SOA reflects faster responses when participants had more time to prepare execution of the correct response. The main effect of spatial mapping reflects faster responses with direct mapping (M ± SE = 450 ± 12 ms) compared to inverse mapping (489 ± 17 ms). The advantage of direct spatial mapping (Experiment 5) was reliable at SOA = 0 ms (t[50] = 3.31, p < .01), but it was not statistically reliable at the later SOAs (t values < 1.85, p values > .07).

Percentage of errors (PE) were submitted to the same ANOVA, which revealed no main effect of SOA, a main effect of spatial S-R mapping (F[1,50] = 4.52, p < .001, ηp² = .23), and a two-way interaction (F[4, 192] = 3.42, p < .05, ηp² = .06). Consistent with the RT results, PEs were lower with direct spatial mapping (2.1% ± .2%) compared to inverse mapping (4.4% ± .6%). This advantage was statistically reliable with SOAs = 0 and 100 ms (respectively, t[50] = 2.96 and 5.10, p values < .01), but not at later SOAs (t values < 1.52, p values > .13). Finally, PEs on catch trials were quite low (3.4% ± .6%) and did not differ across the two experiments (t[50] = .46, p = .65).

**Test Phase.** Before calculating mean response times (RT), I excluded errors (2.1% of trials) and RTs that fell 2.5 SD beyond the total mean (3.0% of trials). Mean RT data were submitted to a 5 x 2 x 2 mixed ANOVA, with the SOA (0, 100, 200, 400, or 700 ms) and target location (congruent vs. incongruent colour) as the within-subject factors and Acquisition Type (Experiment 5: direct spatial S-R mapping vs. Experiment 6: inverse S-R mapping) as the between subject factor (Figure 13). This analysis revealed a main effect of SOA (F[4, 200] = 85.49, p < .001, ηp² = .63), a main effect of Target Location (F[1, 50] = 9.12, p < .01, ηp² = .15), and a two-way interaction between SOA and Target Location (F[4, 200] = 5.50, p < .001, ηp² = .10). Thus, the main effect of Acquisition Type (F[4, 50] < .001, p > .90, ηp² < .001), the SOA X Acquisition Type interaction (F[4, 200] = 1.49, p = .21, ηp² = .03), the Target Location X Acquisition Type (F[1, 50] = .10, p = .75, ηp² = .002) and the three-way interaction (F[4, 200] = .21, p = .93, ηp² = .004) were not statistically significant. The main effect of Target Location indicates overall faster responses on trials when the target appears inside the action-congruent colour (409 ± 8 ms), compared to the action-incongruent colour (413 ± 8 ms). The benefit of
action-congruent colour, however, was reliable only at SOAs of 0 ms (t[51] = 2.52, p = .015) and 200 ms (t[51] = 3.01, p < .01), while it was not statistically significant at later SOAs of 400 ms (t[51] = 1.37, p = .18) and 700 ms (t[51] = 1.70, p = .10). Interestingly, at SOA = 100 ms, there was a reliable disadvantage for the action-congruent colour (t[51] = 3.01, p = .015).

Finally, and less relevant for our purpose, is the main effect of SOA that indicates faster responses with more preparation time, except for the rise in RTs at 700 ms, which might indicate participants' tendency to regard late target onset as a catch trial.

**Figure 12.** Response time data from the acquisition phase of Experiments 5-6, graphed as the function of the spatial mapping between the response cue (which is also the response-outcome location) and the response key and the delay between the response cue onset and target onset (SOA). Error bars represents 95% between-subjects confidence intervals.
Figure 13. Response time data from the test phase of Experiments 5-6, graphed as the target location (congruent vs. incongruent colour in relation to the learned action-outcome colour) and the delay between the response cue onset and target onset (SOA). The top graph represents data from Experiment 5 (direct S-R spatial mapping in the acquisition phase), and the bottom graph represents data from Experiment 6 (indirect S-R spatial mapping in the acquisition phase). The test phase of the two experiments were identical. Error bars represents 95% within-subjects confidence intervals.
Percentage of errors (PE) on non-catch trials were submitted to the same ANOVA, which revealed only a main effect of SOA (F[4, 192] = 3.89, p < .01, η²_p = .072) and a main effect of Target Location (F[1, 50] = 7.95, p < .01, η²_p = .14). The main effect of SOA was driven by the reliably higher PES at SOA = 200 ms (2.2% ± .2%), compared to SOA = 100 ms (1.2% ± .2%, t[51] = 2.73, p < .01), 400 ms (1.3% ± .2%, t[51] = 2.95, p < .01), and 700 ms (1.3% ± .3%, t[51] = 2.86, p < .01). Otherwise, PEs did not differ as a function of SOA. More importantly, the main effect of Target Location was consistent with the RT results, with higher errors when the target appeared inside the incongruent colour (1.8% ± .2%) compared to the congruent colour (1.3% ± .2%). This finding could reflect the effect of visually selected colour on response selection, presumably because attending to a colour would activate the associated response (e.g., Elsner & Hommel, 2001). Finally, PEs on catch trials were quite low (2.1% ± .4%) and did not differ across the two experiments (t[50] = .86, p = .39).

Discussion

The purpose of the present study was to complement the findings discussed in Chapter 3. Namely, I asked whether and how response selection processes, prior to execution, would induce a visual bias based on the learned associations between the responses and their visual outcomes. Left and right keypress responses were associated with green and red colour outcomes during an initial acquisition phase. Across two experiments, the test phase revealed that response selection can modulated feature-based attention to colours. Specifically, immediately after the onset of the response cue (SOA=0 ms), and 200 ms after the onset response cue, participants showed a bias toward the colour that was compatible with their anticipated action outcome. That is, responses were faster when the go-signal appeared inside the placeholder of the compatible colour, compared to when it appeared inside the placeholder of the incompatible colour. At the 100-ms this pattern was reversed and, moreover, no bias was observed at later SOAs of 400 and 700 ms. These findings suggest, first, that even prior to execution of a response, the process of response selection can induce a visual bias driven by the response-outcome associative learning. Second, the modulated visual bias appeared only at earlier phases of response selection, presumably at later times response selection uncertainty is reduced to a point where observers are no longer
susceptible to external, irrelevant sources of information that could help discriminate the correct response (Ansorge & Wühr, 2004; Hommel, 2000).

These findings are consistent with the study reported by Ziessler and Nattkemper (2011), although their method was different in two important ways. First, Ziessler and Nattkemper used a higher number of stimuli, responses, and outcomes. Eight stimuli were mapped to 8 outcomes, but each response was mapped to two stimulus-outcome pairs, in order to dissociate the effect of response-outcome associative learning from the effect of stimulus-response-outcome learning. Second, unlike the present study, responses were determined solely on the basis of a single target stimulus and irrelevant distractors (compatible vs. incompatible with the known response outcome) could appear at various times before or after the target. Regardless of these differences, the authors found that shortly following the presentation of the target (~150 ms), participants' performance was sensitive to distractors, whereas they were not sensitive after longer delays of about 300 ms (Ziessler & Nattkemper, 2011; Experiment 1). Furthermore, additional experiments confirmed that relatively early stages of response selection renders participants tend to favour irrelevant information that matches the learned response-outcome, whereas later stages (closer to response execution) a bias against the outcome is observed. The authors formulated an important distinction between desired outcomes and anticipated outcome. Desired outcomes form the basis of response selection, whereas anticipated outcomes are consequences of response selection. As a result, the representation of anticipated outcomes may not remain activated if there is a delay between selection and execution. I propose, in agreement with Ziessler and Nattkemper, that the anticipation of outcomes take place once response selection begins, but once a response is selected then anticipated outcomes might constitute a smaller set of (desired) outcome (cf. Kunde, Koch, & Hoffman, 2004).

A number of recent studies have brought to light the possibility that action outcomes may not always contribute to action selection. Herwig, Prinz, and Waszak (2007) used a learning phase, in which participants either responded to target stimulus or voluntarily chose a response. In both conditions, responses were followed consistently by auditory outcomes. In the test phase, the tones were used as target stimuli and the tone-response pairing was congruent for one group of participants and incongruent for another group. Importantly, a performance benefit for congruent mapping was observed only for participants whose responses during the acquisition phase were chosen voluntarily, whereas no benefit for congruent mapping was observed for
participants whose responses during acquisition were chosen based on target stimuli. These findings suggest that when responses are selected based on target stimuli (stimulus-based), observers do not allocate much weight to the outcome of their actions, weakening or eliminating action-outcome associative learning (Herwig & Waszak, 2009; Pfister et al., 2011).

In the present study, however, I found reliable evidence in favour of response-outcome associative learning, despite the fact that each response was selected based on target stimuli. Unlike the study by Herwig et al. (2007), each response required two consecutive stimuli for its selection and execution, including a stimulus that guided the "what" decision and a second stimulus that guided the "when" decision (Haggard, 2008). Furthermore, as response certainty increased, with regard to the "what" decision (here, longer SOAs following the response cue), the nature of the task changed from a two-alternative discrimination response to a speeded detection response with only a single response. At the later phase, where participants were certain about their selected response, they were also less sensitive to the anticipated colour outcome.

Along with the findings of Herwig et al. (2007), the present findings support the notion of dimensional weighting (Memelink & Hommel, 2013), according to which stimulus dimensions that are helpful to task performance are prioritized at the expense of those that are unhelpful. Dimensional weighting could be characterized in terms of an optimal way of allocating resources across sources of uncertainty reduction. If the target stimulus can effectively reduce uncertainty with regard to response selection, then participants could reduce the weight given to irrelevant stimuli, including response outcomes. Importantly, the present study demonstrated two phases of response selection that were distinct based on their sensitivity to the anticipated response outcome. These findings, therefore, not only support the notion of action-outcome associative learning as a source of feature-based attentional guidance, they also provide an important boundary condition.
Chapter 5: Sensorimotor Associative Learning (III): Response Priming Can Guide Feature-Based Attention

In the previous two chapters, I discussed the possibility that both response execution and response selection can serve as a source of visual bias. The findings of these studies supported this possibility, as both prior and after response execution, participants favored the colour feature that was compatible with the anticipated outcome of their response. It should be noted, however, that in both paradigms used in Chapters 3 and 4, responses were selected voluntarily. In the present chapter, I investigate the possibility that response selection processes guide visual bias even when they are activated in a stimulus-driven manner. Response selection has been shown to be susceptible to stimulus-driven processes. Could such effects on response selection, in turn, have any influence on perception?

According to the ideomotor theory, to have a response bias is to have a perceptual bias toward the known effects of the response (Shin et al., 2010; Schütz-Bosbach & Prinz, 2007). This important implication of the ideomotor theory, however, remains to be tested. To illustrate the question, let us consider a task in which responses are defined based on two features: location (left/right keypress) and auditory effect (low/high-pitched tone). In this example, we could bias response selection in favor of the left keypress, either by presenting the observer with stimulus on the left-side, which activates the response via the spatial code, or by presenting a low-pitched tone, which activates the response via its known effect. We can then ask whether biasing the response, using the spatial dimension, can have an effect on perception of the tones. Finding that a bias in response selection leads to a perceptual bias to features associated with the response will have two important consequences. First, this finding will strengthen the ideomotor theory by demonstrating the potential for bidirectional activation between perception and action. Second, by considering action selection as a source of perceptual bias, this finding will inform our view of the mechanisms underlying perceptual prioritization.

A choice reaction task is used in the present study, wherein each response is distinct on two dimensions – location and colour. I took advantage of this dual representation by using one feature (location) to produce a response bias while I examine the possible changes to the perception of the second feature (colour). Similar to previous studies, I presented participants with irrelevant perceptual events that overlapped with response features so that these irrelevant
events can activate corresponding responses (e.g., activating left by presenting an irrelevant stimulus on the left periphery). Critically, the second (colour) feature was presented at a later point in time, when a response bias was already induced by the first (spatial) event. This allows us to examine the perceptual influence of the response bias; when responses and colours are linked in an action representation, does biasing a response alter the processing of colour?

Previous research on stimulus-induced response bias (e.g., Hommel, 1996; Elsner & Hommel, 2001) suggests that presenting a perceptual feature that overlaps with the representation of one response will bias response selection. For the present purpose, it was critical to first confirm whether I could induce a measurable response bias using a spatial prime, given the specific task characteristics. First, the possibility that responses might be represented based on colours might weaken the role of location in response representation and, consequently, weaken the effect of an uninformative spatial prime. Second, the delay between the onset of the spatial prime and the onset of the imperative stimulus was relatively long (400 ms), which might render any possible transient bias unobservable. Thus, in Experiment 7, I introduce the basic structure of our task and provide evidence for the effectiveness of spatial primes inducing a response bias. Then, in Experiment 8, I modify the task to address the issue of whether or not a colour-based perceptual bias induced by a bias in response selection.

**Experiment 7: Two Sources of Response Bias**

The goal of this experiment was to test whether an uninformative spatial prime can be effective in causing a measurable response bias after introducing an association between responses and their colour effects? In our pilot experiments, I have found that forming a response-colour association can greatly reduce or eliminate response-target spatial compatibility effects, suggesting the dominance of the non-spatial dimension is response coding (cf. Hommel, 1996; 2004). However, there is also evidence for the possibility of two-feature response coding, such that each response feature could interact with stimulus features. In particular, Kunde and Stöcker (2002) reported persistent spatial compatibility effects, when responses were defined on spatial (left/right) and a non-spatial (short/long) dimensions. Finding a similar persistence of the spatial dimension in response selection is important for the present study, since I aim to rely on the spatial dimension to produce response bias.
Once again, association between responses and their colour effects were established through an initial acquisition phase. During the acquisition phase, participants identified a target letter that was presented together with a distractor letter along the vertical axis (see Figure 14). The display also consisted of two circles positioned on the left and right periphery. Immediately after each response, the colour that was associated with the response filled one of the two peripheral circles. The location of this colour effect was chosen randomly (e.g., after the right-hand keypress, red could appear inside the left or right circle with equal probability) in order to avoid forming an association between colours and the circles.

Since each colour effect coincided with both a response and a target letter, formation of colour-response and colour-letter associations were both possible. Based on previous studies, there is reason to assume some degree of colour-response association (Ziessler, 1998; Hommel, 2004), although Experiment 7 cannot discriminate between the two kinds of associations. The question of whether any response-colour mapping is formed will be addressed directly in Experiment 8. Here, the question is whether the potential colour-response mapping would weaken the location-response mapping to such an extent that the location can no longer be used to prime responses.

After the acquisition phase, participants then performed the test phase, which involved a similar visual letter identification task (Figure 14). In this phase, however, before the presentation of the letter stimuli, two primes were presented in sequence, each corresponding to one of the two response dimensions. The first prime was a spatial prime and consisted of the onset of one of the peripheral circles at the left or right side of fixation, which later contained the response (colour) effect. Presentation of the spatial prime was expected to bias participants toward selecting the response of the same location (e.g., if the circle appears in the left, then participants would be biased toward selecting the left keypress) for two reasons. First, as an irrelevant visual stimulus that possesses the same binary spatial features as the response (i.e., two positions along the horizontal axis), the circle was thought to activate the compatible response code (Ansorge & Wühr, 2004; Kornblum, Hasbroucq, & Osman, 1990; Lu & Proctor, 1995). Second, the circle also signaled the location of an upcoming colour effect, and therefore it is possible that the anticipated location of the response effect would also activate the compatible response code (Ansorge, 2002; Kunde, 2001). I treated these factors (i.e., stimulus-response and response-effect compatibility) together as a single source of response bias. The spatial prime was
uninformative and could be either compatible or incompatible with the location of the performed response.

**Figure 14.** Sequence of events in one trial of the acquisition phase (above sequence) and the test phase (lower sequences) of Experiments 7-8. During the acquisition phase, the colour effect of each response was equally likely to appear inside the left or right circle. The colour prime, in test phase of Experiment 7, consisted of a change in both placeholders above and below fixation, whereas in Experiment 8, it consisted of a change in a single placeholder and, therefore, also functioned as a visual-attentional cue (this cue was equally likely to appear above or below fixation and it did not predict the target location, which was itself equally likely to be above or below).

After the presentation of the spatial prime at the left or right periphery, the colour prime was presented at both possible locations for the imperative stimulus (placeholders above and below fixation). The colour prime is also expected to bias performance by activating the associated response (i.e., the response that would produce the same colour as the prime; Hommel, 2004), and by activating the associated target letter (i.e., the target letter that coincides with the colour). Thus, each trial of the test phase is characterized by the relationship between
the spatial prime and the correct response, and the relationship between the colour prime and the correct response. I predicted compatible primes to reduce response times, compared to incompatible primes, for both spatial and colour primes. In particular, finding an effect of the spatial prime will justify the next step (Experiment 8), which involves examining the effect of response bias on colour processing.

**Experiment 8**

**Participants.** Twenty University of Toronto undergraduate students gave their informed consent and participated in this experiment in exchange for course credit. All participants reported normal or corrected-to-normal vision and were unaware of the purpose of the study. All experimental protocols were approved by the University of Toronto Research Ethics Board.

**Stimuli.** The display structure and the sequence of events in the experiment are shown in Figure 1. Stimuli appeared in white, except when noted otherwise, against a black background. Two vertically aligned squares (size = 2° x 2°) acted as placeholders for the letter stimuli ('S', 'H', 'E', 'U', approximately .5° 1° in size, presented in Arial font). The center of each placeholder deviated by 5° of visual angle from fixation. The placeholders were embedded within a larger vertical rectangle (4° x 14°). Outside this rectangle were two horizontally aligned circles (diameter = 2°), one of which were filled with the colour red or green immediately after a response. The center of each circle deviated by 9° of visual angle from fixation. During the acquisition phase, both circles were present throughout the entire block, whereas during the test phase only one circle appeared after the initiation of each trial, acting as the spatial prime.

**Procedure.** Each trial of acquisition phase began with the presentation of empty placeholders. Participants were instructed to begin each trial by fixating at the central cross (1° x 1°). After a 1000-ms delay, a letter appeared inside each placeholder. On each trial, one target letter ('S' or 'H') and one distractor letter ('U' or 'E') were presented. Participants were instructed to respond as quickly as possible with a right-hand keypress (the '/?' button on the keyboard) when identifying 'S' and with a left-hand keypress (the 'Z' button on the keyboard) when identifying 'H'. Letters remained on the screen until one of the two possible responses was recorded. Participants received visual feedback for anticipation ("TOO QUICK!") or late responses ("TOO LATE!"), defined as responses earlier than 100 ms and later than 2000 ms.
Immediately after a keypress, one of the two peripheral circles were filled with a colour (red and green, for the right- and left-hand keypress, respectively). This colour feedback remained on the screen for 500ms. Note, first, that the location of the colour feedback (left or right) had no systematic association with the response location. That is, for instance, the left-hand keypress was equally likely to fill the left or right circle with the colour green. This was done to prevent association between colours and location of the peripheral circles. Second, the colour effects were produced regardless of the correctness of the response. That is, when participants made an incorrect response (e.g., left-hand keypress to 'S') they received the colour feedback associated with their incorrectly performed response, which was then followed by another visual feedback ("MISTAKE!") remaining on the screen for 2000ms.

Each trial of the test phase also began with the presentation of vertically aligned placeholders above and below the central fixation mark. Importantly, the two circles were absent during the initial stage of each trial. After a 1000-ms delay one of the two peripheral circles appeared, acting as an exogenous response prime (i.e., the onset of this circle was aimed to bias the participant toward the response that matched the location of the circle and away from the alternative response). After a 200-ms delay, both target placeholders changed colour both to red or green. After another 200ms, both placeholders changed back to white and, at the same time, the letter stimuli appeared inside the placeholders. Similar to the acquisition phase, immediately after a response, the colour feedback appeared inside the circle. The mapping between the responses and feedback colours did not change across the two parts of the experiment.

**Design.** Participants completed one acquisition block of 200 trials and a separate test block of 320 trials. Both the acquisition and the test phase preceded by a separate block of 20 practice trials. During both phases, target letter ('S' or 'H'), target location (up or down), and the location of the response effect (left or right) were randomized and equally probable. In addition, during the test phase, the spatial prime and the colour prime were both uninformative, predicting response with chance-level validity.

**Results & Discussion**

**Acquisition phase.** After excluding incorrect responses and responses faster than 100 ms and those that fell beyond 2.5SDs of the total mean (3.7% of trials), mean response time (RT)
was 567 ms (SE = 15). Mean error rate was 2.8 % (SE = .6). Overall, these data suggest good performance during the acquisition phase.

**Test phase.** Before analyzing the RT data from test trials, incorrect responses (2.4% of trials) were excluded. Response latencies below 100 ms or latencies beyond 2.5SD around the total mean were also excluded as outliers (1.4% of trials). The remaining RT data were submitted to a 2 x 2 repeated measures analysis of variance (ANOVA), with spatial prime (compatible vs. incompatible) and colour prime (compatible vs. incompatible) as independent factors (α = .05). The data is graphed in Figure 15. Both main effects reached the significance level: a main effect of spatial prime \[F(1,19) = 5.33, p = .032, \eta_p^2 = .219\] indicated faster responses with compatible (M ± SE = 545 ± 13 ms) compared to incompatible (551 ± 14 ms) prime. A main effect of colour prime \[F(1,19) = 10.70, p = .004, \eta_p^2 = .360\] indicated faster responses with compatible (542 ± 13 ms) compared to incompatible (553 ± 15 ms) colour primes. No evidence of a two-way interaction was found (F < 1).

![Figure 15](image)

**Figure 15.** Response time data from Experiment 7, graphed as a function the first (spatial) and the second (colour) primes. Error rates are provided at the base of each bar. Error bars represent within-subjects 95% confidence intervals.

Error rate data from the test phase were submitted to a similar three-way ANOVA, which revealed no main effect or interaction (Fs < 1).
Both the spatial prime the colour prime successfully biased selection of the compatible response, consistent with the assumption that location and colour were included in response representation. Most important, for our purpose, the uninformative spatial prime produced a measurable bias in response selection, despite the 400-ms delay between the onset of the prime and the onset of the target, and despite the onset of a colour prime during this delay period.

The effect of the colour prime, however, is somewhat less clear. On the one hand, this effect can be interpreted as evidence for the inclusion of colour codes in response representation, consistent with previous work (Hommel, 2004; Koch & Kunde, 2002). On the other hand, colour primes could also bias participants toward their corresponding target, due to the formation of target-colour associations during the acquisition phase. Of course, these two possibilities are not mutually exclusive. In light of the findings of the next experiment, however, I reason that there is some degree of response-colour mapping underlying the main effect of the colour prime.

The absence of an interaction between the two sources of response bias is often taken to indicate the distributed nature of response representation over cortically separate feature maps, enabling each feature to form a separate association with the motor command (e.g., Colzato, Warrens, & Hommel, 2006; Hommel, 1998; Hommel & Colzato, 2004). This means that each feature can function independently as a source of response bias. Importantly, the location-colour mapping was successfully prevented through our random localization of the response effects.

After establishing that both features could be incorporated into the representation of the responses, I next modify the task used in Experiment 7, such that possible changes to perceptual sensitivity for colours, resulting from the induced response bias, could be measured.

**Experiment 8: Visual Bias**

In this experiment, I directly test perceptual bias for colours, by employing the visual-attentional cueing paradigm (Posner, 1980; Posner & Cohen, 1984). In this paradigm, participants detect or identify a target (e.g., a letter) that could appear in one of multiple possible locations. Prior to the appearance of the target, a cue appears at one of the possible target locations (e.g., an increase in brightness, a sudden motion, a change in colour, etc.). Even when the cue is uninformative (i.e., signaling target location with chance-level validity), it benefits performance when it coincides with the target location (valid cue) relative to when it appears at a
non-target location (invalid cue), indicative of an automatic shift of attention to the cued location (e.g., Abrams & Christ, 2003; von Mühlener, Rempel & Enns, 2005).

The task used in Experiment 7 could be successfully modified into a variant of a visual cueing paradigm, by presenting the colour prime at only one target placeholder. By spatially confining the colour transient, I could expect to observe a typical visual-attentional cueing effect, manifested in faster responses when the cue is valid compared to when it is invalid. Before presenting the hypothesis, a note on the use of the terms “prime” and “cue” seems necessary. Although these terms could be used in very similar senses, in the present chapter, I reserve the term “cue” and “cueing effect” only for the role of the colour transient in orienting visual attention to a possible target location. Thus, I reserve the term “priming” when referring to the activation of a response (e.g., left key press) or possibly a target identity (e.g., letter “H”).

In the present experiment the colour event is expected to have two roles: in addition to priming a response, the colour event also acts as a visuospatial cue. This cue should draw the participants' attention to one the target locations, resulting in faster performance when the target letter would coincide with the colour prime location (Posner, 1980; Posner & Cohen, 1984). Most importantly, the colour prime was presented when participants had already acquired a response bias, due to the previously presented spatial prime. I hypothesize that the response bias, caused by the spatial prime, would lead to a perceptual bias toward the matching colour, resulting in a stronger visual-attentional cueing effect by the matching colour prime. On the other hand, the induced response tendency would bias participants against the colour associated with the alternative response, resulting in a weaker visual-attentional cueing effect by the mismatching colour prime.

Method

Participants. Twenty-two University of Toronto undergraduate students gave their informed consent and participated in this experiment in exchange for course credit. All participants reported normal or corrected-to-normal vision and were unaware of the purpose of the study. All experimental protocols were approved by the University of Toronto Research Ethics Board.
Stimuli, and Procedure. These were identical to Experiment 7. The only modification made was to the colour prime of the test phase; instead of both target placeholders changing colour, only one of the placeholders changed colour before the target onset.

Design. Participants completed one acquisition block of 200 trials and a separate test block of 320 trials. Both the acquisition and the test phase preceded by a separate block of 20 practice trials. The spatial prime and the colour prime were uninformative. That is, they predicted response and target location with chance-level validity. Trials of the test phase were coded based on the relationships between the spatial prime and the correct response (compatible vs. incompatible), the relationship between the colour prime and the response effect (compatible vs. incompatible), and the validity of the colour prime, as an attention cue, in predicting the location of the visual target (valid vs. invalid).

Note that the two primes matched whenever they were both either compatible or incompatible with the correct response. Therefore, I expected stronger attentional cueing by colour when both response cues were compatible or incompatible. By contrast, the two primes mismatched whenever only one of them was compatible with the response. Therefore, I expected a weaker attentional cueing by colour when only one of the two primes was compatible with the correct response.

Results & Discussion

Acquisition Phase. After excluding incorrect responses and response latencies below 100 ms and those that fell beyond 2.5SD of the total mean (3.6% of trials), mean RT was 594 ms (SE = 16). Mean error was 4.3% (SE = .7). Overall, these data suggest good performance during the Acquisition phase.

Test Phase. Before analyzing the RT data from test trials, incorrect responses (3.4% of trials) were excluded. Response latencies below 100 ms or latencies beyond 2.5SD around the total mean were also excluded as outliers (3.7% of trials). The rest of the RT data were submitted to a 2 x 2 x 2 repeated measures ANOVA, with spatial response cue (compatible vs. incompatible), colour response cue (compatible vs. incompatible), colour cue location (valid vs. invalid), as independent factors (α = .05). The data is graphed in Figure 16. All main effects reached significance level: a main effect of the spatial prime [F(1,21) = 5.08, p = .035, ηp² = .19], indicated that compatible spatial primes (M ± SE = 546 ± 14 ms) led to faster responses
compared to incompatible spatial primes (552 ± 15 ms); a main effect of the colour prime [F(1,21) = 9.36, p = .006, Ũp2 = .31], indicated that compatible colour primes (544 ± 15 ms) led to faster responses compared to incompatible colour primes (554 ± 15 ms); a main effect of the colour prime location [F(1,21) = 7.56, p = .012, Ũp2 = .26], indicated that when the colour validly cued the target location (544 ± 14 ms) it led to faster responses compared to when it was an invalid cue (554 ± 15 ms). None of the two-way interactions reached significance (Fs < 1). Most importantly, a three-way interaction was found [F(1,21) = 6.75, p = .017, Ũp2 = .24], consistent with the prediction that cueing strength should depend on the relationship between the two primes.

As Figure 16 shows, when the colour of the cue was consistent with the spatial response cue (i.e., when both were either compatible or incompatible with the correct response), it caused an attentional cueing effect of 15 ms (SE = 4) that was significantly greater than zero [t(21) = 4.37, SE = 3.46, p < .001]. Importantly, the consistency between the two primes caused a significant cueing effect, both when the target letter was compatible with both [leftmost cluster in Figure 16; t(21) = 2.34, SE = 6.00, p = .029] and when the target letter was incompatible with both [rightmost cluster in Fig3; t(21) = 3.42, SE = 4.74, p = .003]. By contrast, when the colour of the cue was inconsistent with the spatial response cue, its effectiveness was not significantly different from zero (cueing effect = 3 ± 5 ms, p > .4). More specifically, no attentional cueing was observed when the spatial prime alone (p > .9), or when the colour prime alone (p > .3) were compatible.

Error rate data (Table 1) from the test phase were submitted to a similar three-way ANOVA. This analysis only revealed a significant two-way interaction between colour prime and location cue validity [F(1,21) = 5.00, p = .036, Ũp2 = .192]. Specifically, when the colour prime was compatible with the correct response, location validity did not reduce error rates [3.7 and 3.3% error, for valid and invalid conditions, t(21) < 1]. This was not surprising in light of the fact that, during these trials, the colour of the prime would already have biased participants to select the correct response. On the other hand, when the colour prime was incompatible with the response, location validity reduced error rates [2.8 and 4.6% error, for valid and invalid conditions, t(21) = 2.21, SE = .008, p = .038]. In short, the effect of cue validity on accuracy became noticeable when the correct response was not primed by the colour. No other main effect
or interaction in the error rate data suggests that the RT findings were not produced by a speed-accuracy tradeoff.

Figure 16. Response time data from Experiment 8, graphed as a function of the relationship between the first (spatial) response cue and the performed response, the second (colour) response cue and the performed response, and the validity of the colour cue in signaling the location of the imperative stimulus. Error bars represent within-subjects 95% confidence intervals.

The key finding of Experiment 8 was that the response bias induced by the spatial prime affected subsequent colour processing. Because participants encountered the colours after already acquiring a modest but measurable response bias, I reason that this response bias involved activating the anticipated perceptual effects of the response, including the learned colour effect. Therefore, when participants subsequently encountered a colour that matched this bias, their attention was reliably drawn to the location of that colour. Similar to the findings of Experiment 7, the present findings showed that both the spatial prime and the colour prime biased performance. The absence of a two-way interaction between the effects of the two primes supports the assumption that, although both features might have been included in response representation, they were not directly associated with each other.
Although I propose that the main effect of cue validity reveals a visual-attentional orienting, it should be demonstrated that this effect is not due to any potential stimulus-response compatibility effect. In particular, attentional orienting should be independent of any spatial compatibility effect between the visual events and the response. From previous work, we know that spatial compatibility effects can emerge that enhance processing efficiency for left-down and right-up combinations (Cho & Proctor, 2003; Proctor & Cho, 2006). If so, the combination of a right spatial prime, an up colour prime, and an up target may have resulted in what seems like an attentional cueing effect. Thus, to test the independence of our proposed attentional cueing from this orthogonal spatial compatibility, I re-analyzed the data, in a 2 x 2 ANOVA with target-response orthogonal compatibility (compatible vs. incompatible) and cue validity (valid vs. invalid) as factors. I found no orthogonal compatibility effect (p > .94). Instead, I found a cueing effect \[F(1,21) = 7.43, \text{p} = .013, \eta^2_p = .261\], which did not interact with the potential orthogonal spatial compatibility. Therefore, the attentional cueing effect seems genuine and cannot be accounted for by a spatial stimulus-response compatibility.

In discussing the findings of Experiment 7, I pointed out how the main effect of colour prime could not discriminate between letter-colour and response-colour associations. In view of the results of Experiment 8, however, I argue that some degree of response-colour association must have been formed. That is because the spatial prime, by activating its corresponding response, also caused a processing advantage for the colour associated with the response. Importantly, the effect of spatial primes on processing colours were mediated by response activation because there is no evidence that the two features were directly associated (i.e., both experiments failed to obtain a two-way interaction between the two primes). In fact, I minimized the possibility of forming a direct association between the location of the spatial prime and the colours by presenting each colour at both locations equally frequently.

Furthermore, the results strongly suggest that the perceptual bias for colour did not depend on target letter identity, but instead depended on the induced response bias. If target identity was critical in obtaining the perceptual bias for colour, then attentional cueing should not be observed when the spatial prime and the colour prime were consistent with each other but both incompatible with the target letter (i.e., the rightmost cluster of bars in Figure 16). Contrary to this expectation, I found a reliable attentional cueing (p = .003) when both primes were incompatible. That is, the spatial prime led to a perceptual bias for colour even when the
observer ultimately performed the incompatible response. It seems, therefore, the perceptual bias for colour seems to have been due to a bias in response selection, instead of being mediated by a bias toward forming a consistent target-response-effect episode (Logan, 1988, Hommel, 1998).

The observation that target letters did not modulate the perceptual bias for colour is also inconsistent with an account based on accumulated response activation by multiple features (e.g., Fournier, Eriksen, & Bowd, 1998). According to such an account, the observed RT benefits are due to the presence of visual features that are all associated with one response (e.g., right spatial prime + red + "S"). Since this account is based on response activation, it does not predict any change in the effect of colour prime based on where the prime appears. That is, it does not predict the observed attentional cueing effect. Therefore, the most parsimonious explanation of the present finding seems to require a perceptual bias toward colour (induced by response bias), and not merely accumulated response activation.

Why was the attentional cueing effect non-significant when the colour cue did not match the concurrent response bias? A plausible account for this observation is based on rapid disengagement of attention from the location of the colour transient (Theeuwes, Atchley, & Kramer, 2000). Disengaging from the location of the colour cue would have been a useful strategy, since the colour cues were uninformative regarding the location of the upcoming letter target. It is generally thought that when target presentation does not follow the cue onset within 150 ms, disengagement from the cued location becomes increasingly probable (Posner & Cohen, 1984; Theeuwes et al., 2000). It is, however, less efficient to disengage from the cued location when participants are biased toward the cue feature (e.g., Folk et al., 1992; Remington & Folk, 2001). In the present experiment, the different visual-attentional cueing effect, by colour transients that matched and mismatched the concurrent response bias, might have been due to a difference in efficiency of disengagement from the cued location.

Another possibility for the ineffectiveness of the colour cues that mismatched the concurrent response bias is that the induced response bias may have caused inhibition of the colour associated with the alternative response. Top-down feature inhibition has been shown to play a role in modulating stimulus-driven attentional orienting (e.g., Gozli & Pratt, 2012; Treisman & Sato, 1990). Such feature inhibition may have further facilitated the rapid disengagement from the location of the colour cues in the present experiment. In short, consistent
with our hypothesis, employing the attentional orienting paradigm successfully revealed the perceptual bias caused by a bias in response selection.

**General Discussion**

According to the event-based theories of action, the cognitive representation of an action involves the known perceptual outcomes of the action. Therefore, processes involved in action selection should impact concurrent perceptual processes. In the present chapter, I investigated a special consequence of this proposal: would a bias toward selecting a response result in a perceptual bias toward the known effect of the response? In particular, in the context of a choice response task, where each response produced a distinct colour effect, I tested whether a response bias would induce a perceptual bias for the associated colour. After establishing that a measurable response bias could be induced with an uninformative spatial prime, I examined the influence of this response bias on processing colour transient that appeared at possible target locations. The results revealed a reliable attentional cueing effect caused by the colour transient when it matched the induced response bias, but no such cueing effect when the colour mismatched the response bias. This observation is consistent with the event coding proposal that action selection involves selecting a set of known perceptual outcomes, and that processes underlying action and perception operate on common feature representations (Hommel, 1996; 1997; Hommel et al., 2001; Prinz, 1990; 1997).

Inducing a response bias prior to response selection is not commonly used as a manipulation in studying the influence of action on perception. It is, therefore, important to contrast our response bias manipulation from the frequently used response selection or execution manipulation (e.g., Gozli & Pratt, 2011; Kunde & Wühr, 2004; Müsseler & Hommel, 1997; Schubö et al., 2001, 2004; Thomaschke et al., 2012b; Wühr & Müsseler, 2001; Zwickel et al., 2010). Interestingly, selecting or executing an action often interferes with perceiving visual events that share a feature with the observer’s action. For instance, a leftward keypress response can interfere with visual perception of a briefly-presented leftward arrow (Müsseler & Hommel, 1997; Kunde & Wühr, 2004; Wühr & Müsseler, 2001). On the one hand, since this type of action-induced interference with perception is feature-specific, it supports the proposal that action and perception operate on common features (Hommel et al., 2001). On the other hand, by addressing the impact of action after action selection, these studies perhaps bypassed the
intervals in which ideomotor perception (i.e., perceptual bias toward action outcomes) could be revealed.

The interference of action with perceiving a featurally similar visual stimulus is interpreted to be due to the integration of the feature into the action plan, which renders it temporarily unavailable for other concurrent perceptual processing (Gozli & Pratt, 2011; Hommel et al., 2001; Thomaschke et al., 2012b; Zwickel et al., 2010; Zwickel & Prinz, 2012). An important additional assumption proposed by Stoet and Hommel (1999) is that the time interval in which the feature code is integrated into (or occupied by) an action plan is a subset of the interval in which the feature has above-baseline activation. That is, a feature can be activated without being bound to a response, during brief time intervals preceding and following the response. In the present experiments, for instance, the colour cues were presented prior to response selection, presumably during a time interval when the colour codes were not integrated into any response plan, but were activated due to response bias. If we consider the present response-bias manipulation to have caused feature activation, without integration, then the present findings would be consistent with previous work on action-induced interference with perception.

From these results, I propose that a response bias that occurs prior to response selection will result in the activation of response features without the integration of the features into an action. Examining perceptual processes during this time window can reveal characteristics of ideomotor perception, which may be masked or reversed by feature integration after response selection. Similarly, Zwickel and Prinz (2012) pointed out that ambiguity regarding the upcoming perceptual and response features may increase sensitivity to features that are consistent with concurrent bias (before selection). By contrast, certainty regarding response features may reverse the effect of bias on perception. It is possible that this ambiguity-certainty distinction of Zwickel and Prinz (2012) is similar to the distinction made between bound or unbound features (Hommel, 2001; Treisman & Schmidt, 1982).

Consistent with the present argument, it was previously shown that selecting a left/right keypress response can produce a greater disadvantage for items presented on the visual hemifield contralateral for the response location (Hommel & Schneider, 2002; Müsseler et al., 2005; Thomaschke et al., 2012b). Hommel and Schneider (2002) used a dual task experiment, in which a low- or high-pitched tone (S1) indicated the first keypress response (R1: left vs. right). Shortly
after the onset of S1, an array of four letters, including one target letter (S2) was briefly presented and masked. Results revealed a relative advantage for S2, when it appeared on the same side as R1. Importantly, this was observed only when the onset of S1 was followed by S2 onset by no later than 300 ms. In a similar paradigm, Müsseler et al. (2005) found the same pattern of results when S2 followed S1 by no later than 200 ms. As Hommel and Schneider noted, it is unlikely that this time window solely represents post-selection stage of action. Indeed, it is possible that the activation of location code, prior to integration into an action plan, gave the unrelated visual stimuli that shared this code a selective advantage. This advantage was not observed when the temporal distance of S1 and S2 was increased. Compared to those studies, our task presented the stimulus of interest within a time interval which was unambiguously before response selection. Furthermore, I showed that the response-induced advantage is not limited to location, but could also be observed for other response features.

The present findings support the main argument that associative learning can guide visual attention to features. Participants' tendency to select an action also modulated visual-attentional processing due to the knowledge of the outcome of the action, and in a manner so as to increase the effective salience of visual events that are consistent with the behavioural tendency. Together with the findings of Chapters 3-4, these findings suggest that response selection can guide attention toward the anticipated perceptual outcomes. Critically, these findings also suggest that even a stimulus-driven response bias (prior to voluntary selection of the response) can impact visual bias. Top-down modulations are often considered to enhance performance efficiency and guard against salient distractors. As suggested by the present findings, however, action-induced modulations in perception can also be the byproduct of how actions are represented via their perceptual associations.
Chapter 6: Visual Primes with Multiple Features

The overarching argument presented thus far has been that selection of a target response (Chapters 3-5), and the selection of a visual target shape (Chapter 2), has consequences for how observers attend to the features that are associated with the target of selection. When a colour is associated with the initial target of selection, then selection of the target will bias observers toward selection of the associated colour. In general, learned associations between features across different dimensions can guide the allocation of visual attention. As I discussed in the introductory chapter, the link between associative learning and attention could be viewed in relation to two sets of questions – the questions regarding the basis of selection and the question regarding the units of selection. With regard to the units of selection, the argument is that the target of selection, at any given point, is a multi-featured event and not isolated features (Duncan, 1984; Hommel, 1998; 2004; Kahneman, Treisman, & Gibbs, 1992). Selection of multi-featured events is the issue I pursue more directly in the present chapter.

To examine whether visual features are selected as an integrated unit (i.e., event file, Hommel, 2004), I used a priming experiment with uninformative arrows, which resembles several previous studies (Hommel, Pratt, Colzato, & Godijn, 2001; Pratt & Hommel, 2003; Pratt, Radulescu, Guo, & Hommel, 2010; Tipples, 2002). The difference was that the primes in the present experiment had two features (direction and colour) and were followed by two separate targets on each trial, including a peripheral detection target (T1) and a colour discrimination target (T2). Based on previous studies, I expected a colour priming for T2 (faster responses to repeating colours than to switching colours) and a symbolic spatial orienting effect for T1 (faster responses when the arrow correctly signals the location of the upcoming target).

Symbolic orienting of attention with uninformative arrows was first reported in the study by Hommel, Pratt, et al. (2001), in which participants saw a centrally presented, uninformative arrow that was followed by a peripheral target. Despite the non-predictive relation between the arrow and the target, participants were faster to respond to the peripheral target when it appeared at the location that was indicated by the arrow. What remains to be examined is whether other features of an invalid arrow (e.g., arrow colour) would be inhibited, while features of a valid arrow retain their activity. In the follow-up study by Pratt and Hommel (2003), the authors used multi-featured arrows (colour and direction), but used only a single target on each trial.
Consequently, the impact of selecting (valid arrow) or inhibiting (invalid arrow) cues on cue colour could not be examined.

The impact of cues across multiple perceptual dimensions was examined in a study by Kingstone and Klein (1991). The authors used an experiment with three placeholders and two (Experiment 1) or three (Experiment 2) possible target shapes. Each target shape was associated with a possible target location, i.e., each location had a unique most-probable target shape. Furthermore, the presentation of each target was preceded with a spatial cue. The findings revealed an interesting effect that spatial cuing had, in addition to spatial bias, on bias for shapes. Specifically, an invalid spatial cue reduced response times for the shape that was associated with the cued location. That is, inhibiting an invalid spatial cue caused inhibition of its associated shape. The findings of Kingstone and Klein are consistent with an event-based view of attention. They further lend support to the prediction that (a) an invalid spatial cue is inhibited in the processing of upcoming events and (b) the inhibition of the cue is not limited to its spatial dimension but includes other cue features.

In the present experiment, each trial consists of three events. The first event is the prime, which is an arrow that possesses two features (direction and colour). The second event is a visual detection target that appears at either the left or the right periphery. The relation between the prime and the first target is random, meaning that the arrow direction does predict the location of the detection target at chance level (i.e., valid on 50% of trials). The third event is a visual discrimination target that participants are required to respond to based on colour. The relationship between the prime and the colour target is predictable; two colours always repeat from prime to the second target, while two colours always switch. If colour is processed in separation from arrow direction, then valid and invalid arrow directions would result in the same facilitation of repeated colours over switched colours. If, however, colour is selected in conjunction with arrow direction, then colour priming should be stronger when arrow direction is selected (valid arrow direction), compared to when arrow direction is inhibited (invalid arrow direction).
Experiment 9

Participants. Twenty-seven University of Toronto undergraduate students gave their informed consent and participated in exchange for course credit. Participants reported normal or corrected-to-normal vision and were unaware of the purpose of the study.

Stimuli. The display structure and the sequence of events are shown in Figure 17. The cue and T2 appeared in red, green, blue, or yellow against a black background. Other stimuli appeared in white against the black background. The cue (coloured arrow) appeared centrally, within a 4° x 4° square frame. The length of the arrowhead (2°) was half of the total length of the arrow, and the width of the arrowhead was the same as the total length of the arrow (4°). The peripheral target (T1) was .4° x .4° a square that could appear 6° of visual angle to the left or right of the display center, along the horizontal midline. Finally, the second target (T2) was a coloured square that was centrally presented (2° x 2°).

Figure 17. Sequence of events on a sample trial of Experiment 9.

Procedure. Each trial began with the presentation of a central fixation cross ("+") that subtended .5° x .5° of visual angle and remained on display for 1000 ms, and was then replaced by the cue. After a random delay between 600 and 1200 ms (randomly chosen from the uniform probability distribution), following cue onset, T1 appeared on display and remained until a detection response was recorded. After the release of the detection response key and an additional 500 ms, T2 was presented, remaining on display until a discrimination response was recorded. Participants were instructed to use their left hand to press the spacebar in response to
the onset of T1. They were further instructed to use their right hand and press the left and right arrow keys in response to T2 colour. The left arrow key was used in response to green and blue; the right arrow key was used in response to red and yellow. The relationship between the cue and T2 was 100% predictive. Two colours (red and green for half 13 participants; blue and yellow for the other 14 participants) always repeated from cue to T2, while the other two colours always switched. Regardless of the full predictability, I predicted switching colours to cause slower responses compared to repeating colours due to simple feature priming (Logan, 1990).

**Design.** Each participant performed 20 practice trials and 240 experimental trials, with an opportunity to take a short break after 120 experimental trials. Arrow colour (repeating vs. switching colours), arrow validity, and T1 location were randomized and equally probable. The colour of T2 was fully determined by arrow colour. Trials were categorized based on T2 colour (repeat vs. switch) and cue direction (valid vs. invalid).

**Results**

**Detection Response.** Response time (RT) for the detection task was submitted to a 2 x 2 repeated measures ANOVA with arrow direction validity and arrow colour (repeat vs. switch category) as independent factors (Figure 18). This analysis revealed a main effect of arrow direction \( (F[1,26] = 23.57, p < .001, \eta^2_p = .48) \) and a two-way interaction \( (F[1,26] = 5.19, p = .031, \eta^2_p = .17) \). The main effect of colour category did not reach significance \( (F[1,26] = 2.20, p = .15, \eta^2_p = .08) \). As expected, spatially valid arrows resulted in faster responses \( (388 \pm 18 \text{ ms}) \) compared to spatially invalid arrows \( (409 \pm 19 \text{ ms}) \). The two-way interaction revealed a larger spatial cuing effect by arrows that possessed switching colours \( (32 \pm 7 \text{ ms}) \) compared to arrows that possessed repeating colours \( (12 \pm 6 \text{ ms}) \). It is conceivable that additional processing resources allocated to the arrows with switching colours – representing the more difficult cue-T2 rule – increased the effect of these arrows in spatial orienting. Although this speculative explanation merits further study, it will not be discussed further as it is not pertinent to the primary purpose of the study.

Error data from the detection task was submitted to the same ANOVA, which revealed no main effect of arrow colour or arrow direction validity \( (F \text{ values} < 1.4, p \text{ values} > .20) \), although it revealed a significant interaction \( (F[1,26] = 5.8, p = .023, \eta^2_p = .18) \). This interaction was not driven by anticipation errors \( (RT < 100 \text{ ms}; \text{ only } .3\% \text{ of trials}) \) or late responses \( (RT > 2000 \text{ ms}; \text{ only } .04\% \text{ of trials}) \).
only .1% of trials), and remained strong after excluding these two types of errors (F[1,26] = 10.44, p < .01, \(\eta_p^2 = .29\)). The two-way interaction in errors was driven by higher percentage of incorrect keypress responses with a combination of directional invalidity and repeating colours (error rate = 2%) or directional validity and switching colours (error rate = 2%), compared to when directional validity was combined with a switching colour (error rate = 1%) or when directional invalidity was combined with repeating colours (error rate = 1%).

Figure 18. Response time data for the detection task in Experiment 9, graphed as the function of the relationship between the cue direction and the first target (valid vs. invalid arrow direction) and the relationship between the cue colour and the second target (repeat vs. switch). Error bars represents 95% within-subject confidence intervals.

Colour Discrimination Task. Response time (RT) for the discrimination task was also submitted to a 2 x 2 repeated measures ANOVA with arrow direction validity and arrow colour (repeat vs. switch category) as independent factors (Figure 19). This analysis revealed a main effect of arrow colour (F[1,26] = 5.82, p = .023, \(\eta_p^2 = .18\)) and a two-way interaction (F[1,26] = 6.54, p = .017, \(\eta_p^2 = .20\)). The main effect of arrow direction validity did not reach significance (F[1,26] = 1.13, p = .30, \(\eta_p^2 = .04\)). As expected, repeating colours resulted in faster responses (656 ± 26 ms) compared to switching colours (680 ± 23 ms). The two-way interaction revealed a larger colour priming effect after a spatial valid cue (40 ± 12 ms) compared to spatially invalid
cues (7 ± 12 ms). Error data was submitted to the same ANOVA, which did not reveal any significant effect of the two factors or their interaction (F values < 1.6, p values > .2).

![Response time data for the colour discrimination task in Experiment 9, graphed as the function of the relationship between the cue direction and the first target (valid vs. invalid arrow direction) and the relationship between the cue colour and the second target (repeat vs. switch). Error bars represents 95% within-subject confidence intervals.]

**Figure 19.** Response time data for the colour discrimination task in Experiment 9, graphed as the function of the relationship between the cue direction and the first target (valid vs. invalid arrow direction) and the relationship between the cue colour and the second target (repeat vs. switch). Error bars represents 95% within-subject confidence intervals.

**Discussion**

The purpose of this experiment was to test whether selection and inhibition of an object feature has consequences for other object features belonging to the same object. Specifically, I tested whether priming across one dimension (colour) is affected by whether the same prime object correctly signals the location of another target event. To implement this test, I used arrow cues that were informative in their colour but uninformative in direction. During the time between the presentation of the prime and the colour target (T2), a detection target (T1) appeared at a location that was either signaled by arrow direction (valid arrow) or was at the alternative location (invalid arrow). Findings showed reliable colour priming when the direction of the arrow was valid compared to when it was invalid. These results fit the assumption of event-based attention; selection or inhibition of a target feature was not limited to a single dimension but was extended to the other feature that was bound into the same object.
The present findings fit those reported over two decades ago by Kingstone and Klein (1991), who showed that inhibition a location extends the inhibition to the shape that is frequently presented at that location. Similarly, the findings also fit the results reported in Chapter 2 where preparing to perceive a shape generated a bias for the associated colour and/or a bias against the colour associated with the alternative shape. It should be noted, however, that these findings have been reported in situations with strong feature associations. In the present experiment, cue colours were predictive of target colours. It remains to be seen whether uninformative colours are similarly bound to other arrow directions. Similarly, in the Kingstone and Klein paradigm, would disengaging from a shape at a location bias observers against both features, without a strong shape-location learned association?

The recent findings by Goller and Ansorge (2015) should also be discussed in light of the present investigation. Goller and Ansorge investigated the impact of peripheral cues as a function of three factors. These factors included (a) whether or not the cue feature matched the observer's perceptual goal, (b) whether or not the cue was encountered on the previous trial, and (c) whether the cue correctly identified the target location on the previous trial, i.e., cue validity. The authors found that a cue that does not match the observer's perceptual goal can capture attention if it was presented as a valid cue on the previous trial. That is, the validity of the cue on the previous trial impacts its effectiveness in biasing attention. The event-based interpretation of these findings would be based on the binding of the target and the cue. If the cue and the target are presented at different locations (invalid cue), then the two items do not constitute a single event and are not bound into a unified visual event. If, on the other hand, the cue and the target are presented at the same location (valid cue), then the two items can be perceived as a single event and can be bound into an event file. Since, the selection of this event file facilitates task performance its selection on the next trial would be facilitated, which means the cue feature would be selected more readily in the next trial. The same logic could be applied to the findings of Lambert et al. (1999; 2000), in which a shape that frequently appears in proximity of the target is selected more readily than a shape that appears far from the target. Perhaps collections of features that reliably predict target location or target shape could also be considered as belonging to an event representation that includes both the target information and the target context (Chun & Jiang, 1998; 1999).
Conclusion

In the preceding series of experiments, I examined the role of associative learning as a source of feature-based visual bias. Findings showed that associative learning impacts the guidance of attention by extending the selection of one feature to other associated features. Although theories of visual attention assume a role for top-down modulation of feature salience, the type of modulation that is driven by learned association is not commonly considered. The primary sources of top-down bias is considered to be the direct perceptual goal of the observer (Folk et al., 1992; also see Becker, Folk, & Remington, 2010; 2013). The present findings reveal an indirect way in which goal-driven attention can generate visual bias. That is, perceptual goals can determine attentional bias toward task-irrelevant features based on the association between the direct target of selection and the task-irrelevant feature. Thus, the outcome of attentional selection is a multi-featured event and not an isolated feature. Therefore, if the irrelevant feature resembles an element in the target event, its selection is more probable.

In addition to features that match the perceptual goals, features that are essential in successful movement control are prioritized (Bekkering & Neggers, 2002; Brown et al., 2002; Fagioli et al., 2007; Wykowska et al., 2009). An action-driven bias to features, however, could be framed in terms of learned associations between movement and features that are consistently selected during the particular movement. Associative learning between movements and their perceptual outcomes provides a rich source of guidance for attention (Bompass & O'Regan, 2006a; 2006b). It is important, however, to consider whether actions are unique in their ability to modulate the effect of visual stimuli. Is it the intention to perform an action that results in preactivation of the colour red, or is it, for instance, the tactile sensation association with the action that causes the preactivation? An attempt to reduce the effect of actions to sensory components of the action, remains consistent with the notion that sensory anticipation is essential to action representation. According to such a view, the learned sensory outcomes of an action, collectively, constitute the representation of the action (Hommel et al., 2001; Shin et al., 2010). In other words, features of an action are not inherently distinct from features of perceptual events. Thus, to ask whether actions are unique in their ability to serve as a source of visual bias goes against this fundamental assumption, by restoring the strong distinction between action- and perception-related features.
Furthermore, similar to a perceptual object, an action is represented in terms of its collective sensory outcomes. Decomposing the action into its elements will not result in finding an "essential element" of the action, but will eliminate it. This point was demonstrated in a study by Hoffmann et al. (2009), in which participants learned that keypress actions were associated with two distinct auditory outcomes. After this acquisition phase, auditory stimuli can generate a bias in selecting their corresponding actions. In order to examine action components, Hoffmann et al. tested whether the auditory signals were associated with a finger movement (regardless of key) or pressing a specific key (regardless of finger). Interestingly, neither the finger movements nor the keypresses alone evoked the auditory associations. It was only when the finger-key combinations were preserved that action selection was sensitive to concurrent auditory stimuli. In other words, dissecting the action into components eliminated the action-outcome associations.

Recently, two other factors have been put forth as sources of visual bias. First, feature salience has also been shown to increase when the feature is associated with reward (Anderson et al., 2011). Second, repeated exposure to a feature has also been shown to increase the effective salience of the feature (Awh et al., 2012; Theeuwes, 2010; see also, Lambert et al., 1999; 2000), especially when the feature is useful for task performance in previous encounters (Goller & Ansorge, 2015). Associative learning is a more general form of reward learning and, as such, is theoretically more foundational than reward learning. Moreover, the study of associative learning lends support to an event-based view of attentional selection. According to such a view, repeated exposure to features cannot on its own determine perceptual bias in future encounters (Hommel, 1998; Goller & Ansorge, 2015). Instead, it is the selection of bound features or unified events that influences bias in future perceptual encounters. In other words, associative learning provides a much more parsimonious framework of explaining the two new factors.

The idea that one cognitive process can automatically set in motion a stream of processes was introduced by Schneider & Shiffrin (1977; Shiffrin & Schneider, 1977), who compared conditions where targets and distractors were defined consistently and inconsistently. They argued, for instance, that a stimulus that has been consistently treated as a target could later recruit attentional resources in a way that cannot be controlled. In other words, the cognitive representation of the target can become stronger over repeated exposure, causing efficient activation of the corresponding response. In a similar way, it could be argued that consistent
association between perceptual features will cause formation of cognitive representations strong enough that a feature cannot be selected without an automatic bias toward an associated feature (Gibson, 1969; LaBerge, 1973). Logan (1988) emphasized the role of repeated exposure to consistent episodes, which supports development of an automatic response to a stimulus (cf. Hommel, 2000). Within this framework, extending attentional bias to other features of an episode, besides the target feature, can be regarded as an automatic attentional response.

Of course, examining the impact of repeated exposure to structured stimuli is not unique to the present thesis. Two important lines of investigation into the role of memory in attention are contextual learning (e.g., Brockmole, Castelhano, & Henderson, 2006; Brockmole & Henderson, 2006; Chun & Jiang, 1998; 1999) and memory-guided attention (e.g., Soto et al., 2005; 2008). Due to the conceptual similarity between my thesis and the lessons drawn from those paradigms, I should briefly highlight some important distinctions. In contrast to the demonstrations of memory-guided attention, the present work demonstrated the possibility of a visual bias that is driven by selection along a different visual dimension (e.g., selecting a shape causing bias for a color), or is driven by selecting a motor response. Another important aspect of the present investigations, especially in contrast to previous work on contextual cueing, is the notion that attentional bias is not necessarily driven toward a feature that is physically present on the display. Instead, vision can be biased toward features that are anticipated, but not yet encountered on the display. In this sense, the present work fits within the sensorimotor framework of O'Regan and Noë (2001).

Feature co-occurrence is ubiquitous in natural settings, and an anecdote by Ann Treisman highlights this point effectively. She described how in her lectures, she "used to flash a slide of a magazine picture of a woman in a red blouse sitting in a flowered chair on a striped rug in a room with a yellow lamp, to demonstrate to the audience that natural scenes are not immune to conjunction errors". After losing that particular slide during her move to Berkeley, she was surprised by the difficulty in replacing the picture. "I have failed to find a picture with more than one or two arbitrary, exchangeable properties." (Treisman, 1988, p.215) What allows us to rely on regularities in feature co-occurrences to guard against erroneous binding of features? While Treisman (1988) assumed that erroneous bindings are dismissed due to later processes of recognition that rely on long-term knowledge, the present results favour an earlier modulation of feature processing. Associative learning ensures that activity levels along different feature maps
(e.g., shape, colour, size) are not independent from each other. Such a dependence, consequently, would benefit certain feature bindings over others.

Taking advantage of such real-world associations, Moores et al. (2003) found observers who looked for a visual object (e.g., table) prioritized distractors related to the target object (e.g., chair) compared with unrelated distracters (e.g., chicken), an effect that could be attributed to visual, semantic relatedness, and perhaps in some cases even feature similarity. For instance, although tables and chairs do tend to co-occur in our visual world, they are also semantically related, and resemble each other in some visual features. Given that arbitrary mappings were used between features in the present experiments, the results can only be explained in terms of associative learning, and not semantic relatedness or visual feature similarity.

Another way in which associative learning guides attention is due to the nature of semantic knowledge. While the theory of event coding (Hommel et al., 2001; Hommel, 2004) gives an excellent account of a wide range of simple sensorimotor tasks, theories that describe sensorimotor features as ground for higher cognition (e.g., the perceptual symbol systems, Barsalou, 1999; 2008; Barsalou et al., 2003; see also, Gallese & Lakoff, 2005) can extend the scope of event coding to higher cognitive processes. If the building blocks of higher cognition consist essentially of sensorimotor events, then the logic of event coding should apply to higher cognitive processes. Evidence supports this view. For instance, abstract concepts referring to manual actions, e.g., “push,” “pull”, evoke event codes that include the corresponding motor actions (Glenberg & Kaschak, 2002). Abstract concepts such as “bird” and “sky” evoke event codes that include an upward visuomotor bias (Gozli, Chasteen, & Pratt, 2013). Consistent co-activation of the linguistic representation of “green” and the perceptual representation of the colour green, or the linguistic representation of “bird” and the perceptual representation of above, are likely the reason that perception can be biased through activation of symbolic meaning (e.g., Ansorge & Becker, 2012; Estes, Verges, & Barsalou, 2008; Fischer, Castel, Dodd, & Pratt, 2003; Goodhew, Kendall, Ferber, & Pratt, 2014; Gozli, Chow, Chasteen, & Pratt, 2013; Huffman & Pratt, in press; Sun, Shen, Shaw, Cant, & Ferber, 2015; Yee, Ahmed, & Thompson-Schill, 2012).

Viewing semantic representation as involving sensorimotor event codes and associated features gives rise to specific predictions regarding the way concepts guide attention. Notably, examining the time course of the interaction between spatial concepts and visuospatial attention,
Gozli, Chasteen, and Pratt (2013) found that words interferes with visual attention briefly after the word appears on display whereas after some delay processing is facilitated at the compatible location. The same time course of interaction has been found between language processing and selection of manual actions (Boulenger et al., 2006; Sato et al., 2008). This time course resembles the action-induced modulations of visual processing and could be explained with the same theoretical assumptions (Hommel et al., 2001). Thus, selecting unified, multi-featured events in perception and action in a manner that is guided by associative learning is reflected, not only in simple sensorimotor tasks, but also in higher-level cognitive processes.

There are several benefits to allocating attention to an event rather than just to a relevant target feature. Keeping track of an object is easier if we rely on multiple features of the object (Hollingworth & Franconeri, 2009). When necessary, secondary features of an event can help disambiguate which type of event the observer is currently experiencing, enabling more accurate perceptual decision and response selection (Trapold, 1970; Urcuioli & DeMarse, 1996). Moreover, secondary features of a task can possess positive or negative valence and including them into a unified cognitive representation will have adaptive value (Dayan et al., 2000; O'Brien & Raymond, 2012). Thus, the concept of attentional weight that is often used in regard to the treatment of individual features (Bundesen, 1990; Bundesen et al., 2005; 2011; Memelink & Hommel, 2013) could be extended to describe the treatment of perceptual events that contain multiple features.

In summary, the present series of experiments inform two lines of questioning with regard to the function of visual attention. With regard to the basis of selection, I argued that associative learning provides an important source of bias. The logic of associative learning is foundational and provides a framework for understanding other causes of attentional bias, including mere repetition, reward-based salience, action-based salience, and interactions between conceptual and sensorimotor tasks. With regard to the units of selection, I argued that events, and not isolated features, are the targets of selection. Bias for a feature depends, therefore, on the selection of the event to which the feature belongs. Extending feature selection to event selection equips the observer to benefit from the regularities that exist in the world of perception.
References


