Bangor University

DOCTOR OF PHILOSOPHY

The effects of probability and uncertainty on spatial attention

Close, Alexander

Award date:
2012

Awarding institution:
Bangor University

The Effects of Probability and Uncertainty on Spatial Attention

Alex Close

A thesis submitted in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy

Bangor University
Bangor, Gwynedd
2012

Acknowledgements

I would like to thank my supervisors Giovanni d'Avossa and Ayelet Sapir for their continued support, input, drive and inspiration in completing my PhD. I would like to thank my lab mates Katherine Burnette, Biddy Andrews, Joe Butler and Zia Ul Haq for their input and support and Marik Ludwikow for technical assistance. I would like to thank my parents in addition to Reuben Thornton, Chris Yeomans, Patrick Birchall and Nicola Savill for emotional support and counsel. I would also like to thank my external and internal examiners and committee chair, Ben Vincent, James Intriligator and Martin Bracewell respectively, for their constructive criticism and making four hours of Viva relatively painless. I have been lucky enough to receive ESRC funding during the course of my PhD, in addition to an EPS Grindley grant for conference attendance.

Maximum likelihood analyses in chapter 2 were carried out by Giovanni d'Avossa. All other work presented in the thesis was carried out by Alex Close or in collaboration.

Contents

Figures

Equations

Tables

Abstract

The thesis examines how spatial expectations affect endogenous attention using visual psychophysics and information theoretic models. We consider three main questions (1) how probabilistic cues affect response latencies and discrimination accuracy and whether these effects exhibit the behavioural hallmarks of decision under risk. (2) Whether there are limitations to how spatial attention is distributed across multiple locations. We consider specifically the role of working memory capacity in setting these limitations when attending multiple locations. (3) Finally we investigate whether the reliability of endogenous, attentional cues can be learned.

We carried out a number of experiments using central endogenous cues indicating one or more locations. Each location contained a random dot kinematogram (RDK), but only one RDK contained coherent motion. Experiment 2 was a simple reaction time task requiring a simple, speeded response to the appearance of supra-threshold expanding coherent motion. Experiment 6 employed a fine motion discrimination task. In all other experiments, a coarse discrimination task was used.

Experiments 1 and 2 used probabilistic cues, ranging from a non-informative 25% to a highly informative 86% reliability, in a motion discrimination accuracy task and a motion detection reaction time task, respectively. We found that cue reliability modulated the size of the validity effect for both accuracy and reaction times. However, a two-process model was consistent with a probability matching strategy in the motion discrimination task but an under-matching strategy in the speeded task.

Experiments 3-5 compared motion discrimination performance following probabilistic, one location cues and multiple location cues, which provided the same amount of spatial information. With four RDKs performance was very similar for

single and multiple location cues. We concluded that attention could be flexibly distributed over multiple locations with no costs or limitations. However, when six RDKs were used, motion discrimination accuracy was lower following cues indicating multiple locations than information matched, one location probabilistic cues, suggesting that limitations in distributing spatial attention do occur with more locations. When participants were required, in the same block of trails, to either recall the cued locations or discriminate motion at the cued locations, their recall of the cue locations was worse than blocks when they were only required to recall the cued locations. We conclude that attention to the cued locations interferes with spatial working memory of those same locations.

Experiment 6 used a probabilistic cue, whose reliability was not explicitly communicated to the participants and which changed several times across an experimental session. We find that only some of the participants showed behavioural evidence that they had learned changes in cue reliability. In those participants, who showed evidence of learning, learning took place over sequences greater than four trials.

Chapter 1

Animals extract information about their environment through the senses. Advance knowledge about the state of the environment, allows efficient gathering and use of sensory data. Helmolthz (1867) explicitly suggested that perceiving may be analogous to inferential processes used in rational scientific discovery, where the experimental results and theories are continuously compared and updated. An obvious consequence of this idea is that without internal models, sensory data would not be sufficient to sustain coherent perception. Interestingly, he equated perception of a specific environment with the generation of the set of sensory consequences forecast to result from potential actions of the observer in that environment. In other words, perception is the set of possible sensory experiences that could arise in a particular situation. This idea has come to inform later ecological (e.g. Gibson, 1986) and computational accounts of perception.

A less radical view of the relation between expectations, sensation and perception stems from the idea that the amount of information that the brain can process is limited. In this view, expectations guide selection of those aspects of the incoming sensory data likely to be most pertinent and informative about the current goals of the observer, at the expense of task irrelevant stimuli. For example, if the observer is hungry and looking for an apple, she may selectively attend to round, red or green objects, but ignore square boxes.

An important methodological advance, which has provided a convenient way to manipulate expectations in the laboratory, is based on the use of endogenous cues (Posner, 1980). Endogenous cues inform the observer about the likely location or some other distinguishing feature of a visual target (Sekuler and Ball, 1981), which facilitates its processing.

The work described in this thesis specifically addresses the relation between endogenous cues, their reliability and effects on behavioural performance. Given that spatial attention elicited by endogenous cues is considered to be a voluntary process requiring effort on the part of an observer. It seems reasonable that decision processes may play a role in utilising probabilistic information and prior knowledge of spatial uncertainties in a visual scene. As such perceptual performance may show similar sub optimal characteristics to the known departures from optimality that can be seen in decision-making literature. Such departures from optimality include overestimation of small probabilities, underestimation of large probabilities and use of probability matching strategies.

Our goal is to understand: a) the nature of the internal representation of cue reliability; b) what limits the use of cued spatial information and c) whether participants exposed repeatedly to cue target pairings can learn the internal representation of the reliability of an endogenous cue. In the following sections we summarise current models of visual attention, mostly as they pertain to our understanding of cueing effects.

Models of attention

Attention selects sensory information or facilitates sensory-motor mappings

Several theories of attention have been based on the assumption that cognitive resources, including perceptual ones, suffer from capacity limitations (Broadbent, 1958; Kahneman, 1973; Moray, 1967; Wickens, 1984). One obvious implication of this viewpoint is that that allocating resources to one task or stimulus reduces those available for another. The purpose of attention is then to prioritize stimuli for further processing according to their relevance to the current goals of the observer.

One of the main pieces of evidence supporting the idea that perceptual processes may have limited capacity comes from the set-size effect in visual search tasks. Under certain conditions, increased reaction times and reduced sensitivity is observed in detection tasks when the number of irrelevant stimuli in the scene is increased (Bergen and Julesz, 1983; Triesman and Souther, 1985). One interpretation of this result, consistent with the limited capacity viewpoint, is that when multiple stimuli compete for a limited pool of resources then each stimulus cannot be processed as quickly and accurately as when a smaller number of stimuli are presented. This effect is not observed when the distractors and the target do not share common features (Treisman and Gelade, 1980) according to Feature Integration Theory (FIT), suggesting that limited capacity processing is only engaged when there are subtle perceptual differences between the target and the distractors (Humphreys and Duncan, 1989). However, if the target is distinguished from the distractors either by the conjunction of multiple features such being a red ‘T’ among green ‘T’s and red ‘L’s, or the target and the distractors differ in some manner, which requires close inspection to be noticed (Humphreys and Duncan, 1989), then the set size effect is observed. In these cases, the increase in detection latency with number of distractors contained in the display yields itself to a simple and elegant interpretation, namely that attention is moved serially through the display, one stimulus at a time (Luck and Hillyard, 1990; Treisman and Gelade, 1980).

A different view of the role of attention comes from seminal studies where cues, which indicate the location in which a target stimulus is likely to appear, are used to study the volitional control of attention (Posner, Nissen and Ogden, 1978; Posner, 1980). A highly consistent finding is that reaction times (RT) to detect a target are briefer when the target appears at the cued location than when it appears at an uncued location or is not preceded by a spatial cue. The initial interpretation was that for a motor response to be executed, attention must be aligned with the stimulus. On valid trials, in which the alignment can be anticipated, response latencies are brief, on invalid trials, in which attention is initially oriented to a location different than the one occupied by the target, additional time is required to reorient attention and response latencies therefore are longer.

While the initial studies of endogenous cueing examined the latencies of simple detection responses to highly visible stimuli and therefore did not address the effects of spatial expectations on visual sensitivity, additional evidence has shown that endogenous attention affects visual sensitivity mainly under conditions of high external noise (Dosher and Lu, 2000a; 2000b; Lu and Dosher, 1998; 2000). Other authors have concluded instead that endogenously driven attention can increase contrast sensitivity under low noise conditions (Ling and Carrasco, 2006a).

It is possible that this apparent discrepancy may result quite simply from the fact that for low contrast visual targets empty regions of the display can contribute to external noise just as distractor stimuli do (Cameron, Tai, Eckstien and Carrasco, 2004; Morgan, Ward and Castet, 1998; Shiu and Pashler, 1994).

Attention affects visual sensitivity by decreasing uncertainty about a perceptual decision

A number of authors have endorsed a view different from the one based on either the idea that attention is used to select sensory data or that it is required for implementing basic stimulus response mappings. The underlying assumption is that attentional effects on visual sensitivity and response latencies are ultimately related to changes in the observer's uncertainty, either about the nature of the stimulus or the response required. In visual search tasks a standard model to describe the computations performed to generate a response, e.g. "target present" vs. "target absent", sensory data from distractors and targets are integrated into a decision variable. If the magnitude of the decision variable is greater than a pre-specified threshold then the observer will report "target present", otherwise "target absent". According to this simple model of the decision process, detection accuracy will be limited not only by the sensory distinctiveness of the target, but also that of the distractors. Set size effects on response accuracy become then readily explainable, since the probability that one, or more, of the distractors is mistaken for the target must increase with the number of distractor stimuli, as well as the effect of cued spatial expectancies, since they limit the impact of visual data, from locations unlikely to contain the target, on the decision variable (Lappin and Uttal, 1976; Palmer, Ames and Lindsey, 1993).

In this vein psychophysics has made use of Signal Detection Theory (SDT) (Tanner and Swets, 1954). SDT considers attention to be a parallel, unlimited capacity process followed by a decision rule and does not entail any conception of allocating limited capacity resources as Treisman and Gelade's (1980) seminal study. The basic profile of responses to a target are considered to be a distribution centred on a mean response value that varies due to internal noise within the visual system at the level of basic detectors that are attuned to given stimulus features. This assumption has strong physiological underpinnings, the spiking of neurons that detect for example a given orientation, when a given orientation is presented on a number of trials generates responses that vary about a mean spike rate with a given variability about that mean (Bradley, Skottun, Ohzawa, Sclar and Freeman, 1987; Newsome, Britten, Movshon and Shadlen, 1989). However with the presence of distractors of some other orientation, spikes will also be generated for these distractors that are not the target orientation, generating spikes of a much lower mean spike rate and a larger variance (Bradley et al, 1987). This detection system becomes uncertain about the response required only when the distributions generated by internally variable representations of the target (signal) and distributions generated by irrelevant stimuli (external noise) from distractors overlap. This overlap does not increase based the number of distractors when these distractor stimuli differ greatly from the target because this increase only changes the shape of the distribution such that it is representative of sampling numerous times from the basic variance around a mean value. The distractor distribution is still based around the same mean, distant from the target distribution, leading to little overlap and thus little difficulty in discriminating target signal from the distractor's irrelevant signals (noise). However when distractors are similar to the target the two distributions are closer together causing an overlap and creating uncertainty about which stimulus is the target. Under these circumstances increasing the number of distractors also increases the overlap with the target distribution leading to further uncertainty in discriminating the target.

One particular area in which there is major contention between the two stage limited capacity models and decision models is that of search asymmetries. For example when a tilted line is searched for among vertical line distractors compared to a vertical line target among tilted distractors search time differs between these two with shorter RT to tilted targets among vertical distractors (Treisman and Gormican, 1988). Treisman and colleagues explain this as being the result of tilted stimuli being coded as the presence of vertical and tilt, but vertical stimuli coded as the presence of vertical and absence of a tilt. Though this can be better explained as inherent uncertainty within these different types of stimulus, such that when the target distribution has a larger variance search is easier than when the distractor is drawn from a large variance distribution, with greater variance being present in the tilted line stimulus. This accounts well for search asymmetries in that accuracy is shown to decrease as a function of the ratio of standard deviations of the distractor to that of the target distribution (sigma ratio) (Palmer, Verghese and Pavel, 2000). Eckstein, Thomas, Palmer and Shimozaki (2000) show that FIT (Treisman and Gelade, 1980) can be explained without the assumption of capacity limitations in attention, performance being predicted by the sum of discriminabilities along multiple feature dimensions, explaining search with conjunctions of features without any bottleneck or serial second stage required. Even using differing models (Bayesian optimal observer, max rule or heuristic decision rule) with a one stage parallel model based on the uncertainty of sensory information, any of these differing statistical approaches can yield accurate performance predictions (Vincent, 2011).

A recent model of attention that has illuminated a great deal about the physiological underpinnings of decision processes and variability in reaction times is that of Carpenter and colleagues LATER model (Linear Approach to Threshold at an Ergodic Rate). The LATER model (Carpenter, Reddi and Anderson, 2009) considers how differing reaction time distributions occur for detecting easily visible targets and discriminating more difficult targets and how accuracy rates are affected in the case of difficult discriminations. With simple detection of a contrast or luminance stimulus the rate of approach to a decision threshold represents a diffusion process (Ratcliff, Van Zandt and McKoon, 1999) i.e. it has a stochastic variation in the generation of the signal. The variability in reaction times, rather than being considered to be a result of sensory noise, is intentionally inserted to make behaviour less predictable, which may have biological advantages under naturalistic conditions (Carpenter, 1999;

Ratcliff et al., 2001). The LATER model assumes two stages; detectors of low-level features occur according to the diffusion process described, independent of higher-level influences such as prior probability. Then information from these separate detector units are integrated in a second stage that estimates the likelihood the entirety of a particular stimulus to be discriminated is present. This second stage produces a linear rise to threshold that varies from trial to trial, triggering a response, and is permeable to decision parameters such as expectation urgency and reward that are not applied at the first stage. The influence of prior probability is described by the LATER distribution for high visibility cues because in this instance the contribution of the first stage is negligible (Carpenter, Reddi and Anderson, 2009). However by varying the difficulty of detection by varying stimulus visibility, in addition to manipulating prior probability reveals the independent contributions of the separate stages of the process, as the influences of each stage sum linearly and independently of each other (Carpenter, 2004). In addition to providing an account of RT variation the model also predicts the incidence of errors given the expected distributions of reaction time when errors are not omitted from consideration, it is possible to determine the expected error rate from the upper asymptote of the distribution on a reciprobit plot, simulations of these distributions indicated asymptotes representative of objective data (Carpenter, 2009).

The finding that the latency of an orienting response, such as a leftward or rightward saccade, depends on the probability that the target will appear either on the left or the right, suggests that likely responses are selected more readily than unexpected responses (Carpenter and Williams, 1995), converging with the SDT framework discussed and suggesting a roll for probability in spatial attention based on internal biases. A feature of models of attention based on uncertainty reduction, is that they do not imply that perceptual or motor processes are capacity limited to explain the effects of either set size, target or response probability. Whether these models can fully account for the effects of set size in complex tasks is uncertain (Palmer, 1994).

As mentioned briefly above, simple computational implementations of uncertainty based models treat cueing information as weights applied to sensory evidence prior to its integration into a decision variable (Lappin and Uttal, 1976; Palmer, Ames and Lindsey, 1993). Attentional cues thus bias the value of the decision variables to reflect information from sources of sensory data likely to contain the target at the expense of those likely to contain distractors (Lappin and Uttal 1976; Shaw 1984; Shiu and Pashler, 1994; Sperling and Dosher, 1986). The difference between performance in valid and invalid trials is due to the fact that on valid trials, contribution to the decision variable of sensory data from sources containing irrelevant information is minimized, while on invalid trials the decision is based mostly on irrelevant sensory data from sources other than the one(s) containing the target.

Uncertainty based models can account for the finding that when the observer is uncertain about where a visual target has appeared, cueing the target location can improve visual sensitivity, even when the spatial information is provided after the target. Strong evidence for uncertainty based models of attention is thus provided by the finding that post-cues, presented after the target, can improve visual sensitivity to a similar extent as pre-cues, presented prior to the target (Kinchla, Chen and Evert, 1995). A number of additional results further support the idea that the effects of expectancies can be best understood as affecting the integration of sensory information rather than its quality. For example, when multiple visual masks are presented at all locations including the one containing the target, awareness of the target location becomes uncertain and pre-cues are found to improve performance

(Shiu and Pashler, 1994; Henderson, 1991). However, when the target location is the only one masked and participants are aware of that, there are no effects of pre-cueing on visual sensitivity. This suggests that uncertainty about the target location is the crucial factor in cueing effects (Van der Heijden, Brouwer and Serlie, 1992; Shiu and Pashler, 1994), since changes in the quality of the sensory signals cannot be invoked to explain these findings. In other words, because the mask both cues the target location and interrupts sensory processes at the target location, the quality of the sensory data can no longer be modified.

Moreover, the effects of pre-cues on visual sensitivity are no different when targets and distractors are either presented sequentially, over longer time windows, or simultaneously, as long as the target duration is the same. Given that conditions, which differ in the amount of sensory data that are processed simultaneously, exhibit similar cueing effects, it has been concluded that cueing does not change the amount of sensory data processed but simply reduces the influence of irrelevant data on the decision made (Kinchla, Chen and Everet, 1995, Eriksen and Spencer, 1969; Shiffrin and Gardner 1972). More specifically these data limited accounts suggest attention appears to operate by reducing the uncertainty about the spatial location at which the stimulus occurs not the appearance of irrelevant stimuli that could be confused with the target. The studies reviewed here have concluded that the size of the cueing effects on human observers' performance is no greater than those predicted by uncertainty reduction, the implication being that uncertainty reduction suffices to account for cueing effects. Though decision models have also provided some support for increases in contrast sensitivity across all spatial frequencies, in accordance with the idea attention operates via signal enhancement rather than exclusion of noise (Carrasco, Penpeci-Talgar and Eckstein, 2000). Others challenged this conclusion and provided evidence that not all the effects of spatial cueing can be attributed to uncertainty reduction. For example, Luck, Hillyard, Mouloua, Woldorff, Clark and Hawkins (1994) found effects of pre-cueing the target location in a luminance change detection task, even when spatial uncertainty was minimized in the immediate post-target period, by the use of a visual mask at the target location. Later, Luck, Hillyard, Mouloua and Hawkins (1996) proposed that a number of methodological shortcomings in the preceding literature including the use of highly discriminable stimuli, short Stimulus Onset Asynchrony (SOA) between cues and targets, the use of peripheral rather than central iconic cues, may have led to the conclusion that cueing only affects uncertainty. Using longer SOAs led to improvements in performance even when only the target location was visually masked. This was so even when using alpha-numeric characters that could be distinguished from each other by simple features. Luck and colleagues suggested that the use of highly discriminable stimuli was one of the reasons for finding no cueing effects in previous studies with single masking (Van der Heijden, Brouwer and Serlie, 1992; Shiu and Pashler, 1994). However this did not appear to be a factor in preventing cueing effects under single mask conditions when the SOA and central cue were used, the less discriminable stimuli were also used and also lead to effects of cueing (Luck et al., 1996). Verghese (2001) concluded that it is possible that attention may work to increase discriminability by both signal enhancement and exclusion of noise and that the weighting assigned to relevant and irrelevant stimuli may explain both of these accounts.

Attention as labelling sensory lines

A different account of the functions carried out by attention was developed by Cavanagh and a number of his collaborators. In a seminal paper He, Cavanagh Intriligator (1996) examined the effects of crowding on orientation after effects and discrimination. While participants were very poor at discriminating the orientation of Gabor stimuli presented in the visual periphery, when they were flanked by similar Gabor patches, the same stimuli were still effective at inducing an orientation after effect. The authors inferred that the presence of visual flankers did not affect the ability of the target to drive early, orientation tuned, visual filters, hence causing the orientation after effect, but they did affect the ability to select task relevant information, because the of the crowding effect on orientation discrimination. Based on these findings the authors concluded that the resolution of attention is less than that afforded by visual acuity. Intriligator and Cavanagh (2001), capitalizing on the idea that the spatial resolution of attention is limited, suggested that crowding arises because the ability to select stimuli in the visual periphery does not have the necessary spatial resolution, but rather attentional mechanisms mix sensory data arising from the target with those arising from the task irrelevant flankers.

We think that the basic idea, put forward by Cavanagh, Intriligator and their collaborators, can be aptly summarized by using a computer analogy, namely accessing sensory data is like transferring information contained within a data structure on an electronic computing machine, which is achieved through the use of addresses indicating the location of the stored record in the physical memory. Similarly, the main function of attention is to transfer information from low-level sensory representations to higher level perceptual and decision processes.

Since crowding, according to the hypothesis set forth by Intriligator and Cavanagh (2001), demonstrates limitations in the spatial resolution of attentional mechanisms, it has been disappointing that a number of attempts to demonstrate cueing effects on crowding have generally either yielded negative results or shown no effect of cueing on the critical spacing, namely the maximal distance between flankers and target at which the former still affect the processing of the latter (e.g. Morgan, Ward and Castet, 1998; Scolari, Kohnen, Barton and Awh, 2007), contrary to the prediction that attention resolution affects the grain of selection.

Freeman and Pelli (2007) were able to demonstrate cueing effect using an experimental design, which combined crowded stimuli with a spatial uncertainty manipulation. Participants had to perform a change detection task while viewing displays that contained either widely or narrowly spaced alphabetic characters. When viewing narrowly spaced characters, participants were less accurate in detecting a change in the identity of a character than when viewing widely spaced characters. Moreover, a peripheral cue improved detection for both display types, and equalized performance when viewing narrowly and widely spaced characters. Freeman and Pelli (2007) also found that when the characters used were drawn from the Armenian alphabet, which was unfamiliar to the participants, there was a substantial drop in detection performance. They concluded that the main factor limiting detection in this task was the participants' ability to remember the letters identity. This was relatively easy when letters were drawn from a familiar set, i.e. the Latin alphabet, but much more difficult when they were drawn from an unfamiliar one, i.e. the Armenian alphabet, and more speculatively, when they appeared unfamiliar because of crowding.

Notwithstanding the fact that some have disputed the suggestion that crowding reflects limited attentional resolution, the idea that attention selects information has gathered notable neurophysiological evidence in its support. In extra-striate visual regions, where neurons have large receptive fields, the neural response evoked by simultaneous visual stimuli is less than the neural response evoked by the preferred stimulus presented in isolation or the sum of the responses evoked by each of the stimuli presented alone, suggestion non-linear, competitive interaction between visual stimuli (Luck, Chelazzi, Hillyard & Desimone, 1997; Reynolds, Chelazzi & Desimone, 1999). However these effects, due to competition between stimuli, are greatly diminished if the animal's attention is allocated to one of the stimuli, suggesting that attention can bias neural responses toward encoding the properties of the attended stimulus at the expense of the unattended stimuli. Interestingly, similar effects on visually evoked response in extra-striate visual areas can be replicated by electrical stimulation of cortical areas of the frontal cortex thought to be also involved in guiding eye movements and attention (Armstrong, Fitzgerald & Moore, 2006).

Exogenous and Endogenous attention

William James (1890) suggested that two types of attention exist. One is a reflexive, involuntary process, which is engaged “passively” by external, salient stimuli. The second is an “active” process, under voluntary control, which allows the observer to focus on goal relevant information. Posner and Snyder (1975) and later Jonides (1981) proposed a taxonomy of attention based on the dichotomy between “bottom – up” and “top – down” driven cueing effects. Jonides summarised these thusly: “*Peripheral cues do not draw heavily on cognitive resources in comparison with central cues. Peripheral cues induce attentional shifts that are difficult to suppress especially in comparison to central cues. Peripheral cues maintain attention capturing properties when they are unexpected which is not true of central cues. Peripheral cues also produce larger costs plus benefits in processing times and accuracy than central cues*” (Jonides, 1981, p200). These two types of cue have well defined neural substrates, the endogenous system activating intra-parietal cortex and superior frontal cortex of both hemisphere and the exogenous system activating temporo-parietal and inferior frontal cortex of the right hemisphere (Corbetta and Shulman, 2002; Shulman et al, 2010).

The exogenous system has been defined as a bottom up system that re-orients attention automatically and quickly to unexpected events. Exogenous cues are thought to activate an alerting system within 50ms from their onset, and draw attention to the cued location within 100-120ms. Exogenous cues are also difficult to ignore (Jonides 1981; Yantis and Jonides, 1996; Giordano McElree and Carrasco, 2009) and exert their effects even when they are not beneficial to task performance (Prinzmetal McCool and Park, 2005; Pestilli and Carrasco, 2005; Yeshurun and Rashal, 2010). Exogenous cueing effects are transient and decay quickly (Hein Rolke and Ulrich, 2006; Ling and Carrasco, 2006a; Liu, Stevens and Carrasco, 2007; Muller and Rabbit, 1989a), and can be followed by a period of decreased detectability for targets appearing at the cued location, a phenomenon referred to as Inhibition of Return (Posner and Cohen, 1984).

On the other hand, the endogenous system allows for sustained attention to be deployed at spatial locations well in advance of the stimulus appearance. Following an endogenous cue, at least 300ms are required for attention to be fully deployed at the cued location (Coull and Nobre, 1998). Endogenous attention has been shown to affect performance only under high noise conditions, whereas exogenous attention affects performance under both high and low noise conditions (Dosher and Lu, 2000a; 2000b; Lu and Dosher, 1998; 2000), even though more recent studies have challenged these conclusions (Ling and Carrasco, 2006a).

A further difference between exogenous and endogenous attention was suggested by Prinzmetal, McCool and Park, (2005) who examined the effects of endogenous and exogenous cues on speeded detection and accuracy tasks. They found that both exogenous and endogenous cues effected response latencies in a speeded letter discrimination tasks, but only endogenous cues affected letter discrimination accuracy. Furthermore, endogenous, non-informative cues had no behavioural effect, suggesting that observers could ignore task irrelevant cues. Others found that the benefits in response accuracy gained from endogenous cues increase with cue reliability (see below for a thorough review) whilst costs in response accuracy remain relatively constant, while for response latencies, cue reliability modulates the size of the behavioural costs, but not benefits (Giordanno, McElree and Carrasco, 2009). On the other hand, exogenous cues give comparable benefits and costs regardless of the cue reliability suggesting that computation of the target probability takes place following endogenous, but not exogenous cues (Giordano, Mcelree and Carrasco, 2009).

In summation the literature on endogenous and exogenous cues implies that endogenous cues may be more appropriate to study the effects of probability and uncertainty. The voluntary deployment of attention should intuitively be more permeable to the influence of probabilistic information about spatial contingencies. Thus our studies employ central endogenous cues to investigate manipulations of probability and uncertainty. In the empirical chapters to follow we use endogenous cues with duration of 1000ms throughout our experimental paradigms. The time course of endogenous cueing with probability manipulations has been investigated previously (Gottlob, Cheal and Lyon, 1999) and been shown to reveal consistent target discrimination accuracy across the timing of cue stimulus onset asynchronies when probability is manipulated. The authors reasoned that the effects of time after cue and probability were reliant on different processes. In addition this study used blocked conditions allowing participants to consider the same spatial expectations without adjustment, across a number of trials. However we used cues that varied in reliability across trials. Thus we reasoned that participants may require a longer period of time to interpret the cue (interpreting the probability associated with the cue's colour for probabilistic cues for example) as the spatial contingencies associated with the endogenous cue stimuli varied across each trial, instead of an entire block of trials.

Given that the access of endogenous attentional processes to effortful volition, it seems reasonable to assume that decision processes may play a role in the utilisation of probabilistic information during perceptual behaviours. There is a wealth of literature regarding how probability affects decision-making that is relevant for consideration in addition to how it may affect attentional processes, which we consider below. Though a very recent paper has suggested that exogenous cueing may be elicit modulations of perceptual performance with known spatial expectations (Vincent, 2011). This paper uses a four location spatial alternative forced choice localisation task (4-SAFC) requiring localisation of a briefly presented (133.33ms) target oriented gabor that was similar to the distractor gabors and thus insufficient to reduce all uncertainty about target location. This paper finds similar modulations of target localisation according to spatial probabilities with both endogenous and exogenous cueing under these particular circumstances. Never the less the bulk of the evidence would seem to suggest that endogenous central cues are more appropriate and certainly at least equivalent to exogenous cues in eliciting modulations of performance by probability.

How do known probabilities and uncertainty affect behaviour?

How the reliability of endogenous cues is represented in the brain and how it modulates visual processing is an interesting issue in view of a large body of literature that has dealt with the nature of the representation of cue probability and its relation to choice behaviour in domains other than attention, such as reward related choice and financial decisions (Kahneman and Tversky 1979, Platt and Glimcher, 1999; Sugrue, Corrado and Newsome, 2004; Hsu, Krajbich, Zhao and Camerer, 2009).

One prominent idea is that the representation of outcome probability, given alternative options offered to an actor, is separate from processes involved in determining the actor’s choice and subsequent actions. This separation is crucial to allow for flexibility in choice behaviour. For example, the computation of a particular option’s value may need to incorporate changing needs and thus rely on different utility functions, which translate a fixed prize into a flexible measure of its utility, according to the changing needs of the actor. Moreover, the relative expected utility of different options may determine choice in a stochastic rather than deterministic manner, either because value representations are noisy or because decision processes themselves are. The obvious implication of this state of affairs is that in order to understand the nature of the effects of probability and cue reliability on choice, one also needs to consider the nature of the decision processes that determine choice.

Distorted probability representation

Early studies in the field of behavioural economics revealed the existence of biases in representing outcome probabilities associated with alternative choices, as revealed by the pattern of decisions taken by individuals facing risky alternatives, that is choosing under conditions in which the outcome probabilities are known.

Kahneman and Tversky (1979), who formulated prospect theory, also found that people often underestimate the probability of highly likely events and conversely overestimate the probability of unlikely events, confirming previous findings (Allais, 1953). They suggested that that the expected value of different options is based on the product of two non-linear functions of the outcome probability and value, respectively.

Basso and Wurtz (1995) were the first to show that the spatial probability of a target is encoded in single unit activity during the preparatory interval before the onset of a saccadic target. Platt and Glimcher (1999) provided initial evidence that single neurons in the lateral intraparietal area of behaving non human primates exhibit neural responses, which correlate both with the expected gain from a particular trial, as well as the probability that a specific action will be instructed. Sugrue, Corrado and Newsome (2004), examined choice behaviour and neural activity in monkeys trained to saccade to one of two colour coded cue stimuli. A variable interval reward schedule was associated with each colour cue, the reward consisting of some orange juice. Under these conditions, monkeys engaged in a probability matching strategy, choosing the colour cue with a probability that closely matched the relative rate of reward obtained at each cue. Moreover, the monkeys were able to quickly adjust their choice behaviour to changes in contingencies between cue and relative rate of reward, indicating that they kept track of the reward history at each cue. Finally, parietal neurons, whose neural activity was tuned to saccade direction and cue location, also demonstrated modulations of their delay period activity by the expected value of the reward indicated by the cue.

f-MRI data have provided evidence for physiological signals, whose amplitude is scaled by the probability of behavioural relevant events, in the human brain as well. Tobler, Christopoulos, O'Doherty, Dolan and Schultz (2008) reported that a region in the left Middle Frontal Gyrus showed significant effects of the probability of obtaining a reward following the presentation of visual cues, whose reward contingencies the study participants were informed about prior to performing a "cue attractiveness" judgement task in the scanner. Task performance had no consequence on the likelihood of receiving a reward or its magnitude. The most notable aspect of the results was that the relation between reward probability and the amplitude of the Blood Oxygenation Level Dependent (BOLD) signals showed an inverted 'S' profile consistent with overestimation of low probabilities and underestimation of high probabilities. Moreover, Tobler et al., (2008) also found in Inferior Frontal Gyrus

BOLD signals, which exhibited the reversed probability distortion, while the amplitude of the BOLD signal scaled linearly with probability in the Globus Pallidum. Hsu, Krajbich, Zhao and Camerer (2009) confirmed, in a subsequent fMRI study of the BOLD correlates of outcome probability in a gambling choice task, that a number of brain regions exhibit BOLD responses modulated by probability, including Anterior Cingulate Gyrus, Insula, Middle and Inferior Frontal Gyri. Interestingly, the relation between BOLD amplitude and outcome probability was found to be non linear also in regions of the Striatum, including the left Caudate Nucleus and right Globus Pallidum. Overall, these results suggest that there are multiple representations of probability across the brain, which may affect utilisation of probabilistic information and behavioural outcomes.

Deciding under risk: probability matching and stochastic choice

When an animal needs to take repeated decisions about where to seek access to some limited food source, simple empirical laws have been found to describe the pattern of choices. Specifically, animals choosing between two options associated with a concurrent, variable interval reward schedule, allocate their choices with a relative frequency that matches the relative reinforcement frequency obtained from each option (Herrnstein, 1961; Williams, 1988), a behaviour that has come to be known as *probability matching*. These findings indicate that animals are able to learn basic aspects of the environment statistics from repeated exposures to outcome associated with specific cues or actions.

Under different conditions, such as when the animal is faced between choices associated with concurrent variable rate reward schedules, other patterns of choice are observed. Specifically animals tend to adopt a maximization rather than a probability matching strategy, consistently choosing the option associated with the highest rate of reward (Herrnstein, 1960).

Interestingly, in human participants who need to predict the outcome of a binary process, choice behaviour frequently results in probability matching, even though a maximization strategy would result in optimal performance (Bush & Mosteller, 1955). These deviations from optimal performance depend, to some extent, on whether and how the correct choice is rewarded (Vulkan, 2000). Regardless, the finding has puzzled investigators in behavioural economics, since it challenges the idea that humans behave as rational actors (Simon, 1959).

Several hypothesis have been put forward to account for the finding that humans adopt a strategy that is, in some way, non-optimal. S*tochastic learning theories* have been based on two assumptions, first that the decision process that leads to a choice is stochastic and that the probability of a making particular choice is updated on each trial depending on whether the last trial's choice was reinforced. When the updating rule is a simple linear function of reinforcements, it is possible to show that the asymptotic choice behaviour results in probability matching (Estes, 1950). It should be noted that this idea, while simple, does assume that the decision process is intrinsically stochastic, whereas probability matching observed when animals choose between concurrent variable interval schedules may result from the aggregate outcome of a set of deterministic choices, based on changing estimates of the expected utility of the alternative choices (Herrnstein & Vaughan, 1980; Herrnstein & Prelec, 1991). Others have suggested that participants simply get bored with always choosing the same option, and thus sometime choose the alternative one (Brackbill & Bravos, 1962). Finally, some have suggested that participants seek discernable sequential patterns and therefore avoid a maximization strategy as they attempt to guess the outcome of the next trial (Restle, 1961). In support of this hypothesis, it is observed that when participants become aware that outcomes are truly random, they switch to a maximization strategy, mostly choosing the option associated with the highest utility. Only when the workings of the outcome generating mechanism remain obscure, do participants keep engaging in a probability matching strategy (Morse & Rundquist, 1960).

Deviations from optimal behavior may also result from incomplete knowledge of the probability and utility of various outcomes or shifting preferences over time (Mcfadden, 1980).

Effects of cue probability on spatial attention

When the probability of a luminance target appearing at a given location does not change over trials, the accuracy in detecting it is proportional to the probability of the target appearing at that location (Shaw and Shaw, 1977). The model developed by Shaw and Shaw (1977) has been extended to reaction times data (Shaw, 1978) and assumes that spatial expectations work by allocating limited capacity resources. Though it is also possible to gain attentional facilitation in a region of the visual field by indicating the probability of a given location containing the target on each trial with the use of endogenous attentional cues. The first studies to manipulate probability of spatial, endogenous cues were conducted by Jonides (1980, 1983), who questioned and expanded on the idea of a single spotlight of attention, with a fixed focus (Posner, 1980). Jonides (1980) firstly reasoned that exhaustive serial search over all items, with attention focusing on only one item at a time, where all items are searched regardless of whether a target has been located and identified does not occur.

Under this assumption valid and invalid cue conditions should be identical with one another since an exhaustive analysis of the array would require the same number of item examinations in both conditions. Exhaustive search over non-cued locations, should mean neutral and invalid cue performance is equal because both require all eight locations to be scanned, as such valid benefits should be equal to time required to examine seven items. Terminating search over an array of items assumes the invalid trial performance reflects the average of serially scanning the remaining locations, predicting benefits when valid and cost when invalid but with magnitudes for benefits being greater than costs yet they found that the opposite was true. Finally If nearby locations receive benefits in a serial process then distance from the cued position should be factor. Again they found no difference based on total distance between cued and target location or clockwise scanning distance between these locations, also when adding the assumption that people were alternating between clockwise and anticlockwise scanning. The author proposed two alternative hypotheses could account for the patterns of modulation by probability; a two-process model in which attention was either kept in a diffuse state over all locations or focused only on the cued location, and a model in which central resources were allocated differentially to the various locations. Jonides (1983) reasoned that the former model would predict that benefits should be independent of cue validity as the focus of attention would always be restricted entirely to the cued location and reaction times should not vary by the probability that the target will occur there. Having found RT costs for invalid cues and benefits for valid cues that were modulated by the reliability of the cue, which was varied over three levels (30%, 50% and 70%) he initially concluded that the evidence supported the second hypothesis. He also suggested that if the two-process model was in fact the process at work then a probability matching strategy must be used to determine the proportion of trials on which attention was focused at the cued location. With this addendum to the model it was possible to account for modulations to reaction time benefits in addition to costs as a function of probability as performance over valid trials reflected a larger proportion of trials focusing on the correct location.

Eriksen and Yeh (1985) Used a letter discrimination task with target letters located at one of four locations, cueing one location (the primary location) indicating it would contain the target with a given probability. The diametrically opposed location was the secondary location whose probability of containing the target was either equal to the primary location (40% primary, 40%, secondary) or unequal, (70% , primary 10%, secondary). The other two possible locations had a 10% probability of containing the target letter. They tested the probability matching hypothesis in the context of the two-process model by assuming RTs on 100% valid trials represented latencies obtained with fully focal attention and RTs on neutral trials represented latencies with fully diffuse attention. They then estimated the predicted performance in partially valid cue conditions by assuming that attention was deployed focally on a number of trials proportional to cue reliability, namely the proportion of valid trials, and that attention was maintained diffusely in the remainder of the trials. The expected value of the predicted response latency as a function of cue reliability could thus be easily estimated. Their results indicated that performance following probabilistic cues closely resembled the values predicted by the probability matching model. They also hypothesised that the diffuse and focal states of attention are in fact the extremities of a continuum of possible variations in focus size, with reductions in processing power in the focus as size of the focus increases. This idea later developed into the zoom lens model (Eriksen and St. James, 1986).

Madden (1992) replicated this experiment, and found data which suggested that participants attended the cue location more often than predicted by a probability matching strategy. Also they estimated regression coefficients that reflected probability matching using 100% cued RT and neutral RT as predictors of the obtained RT for the cued location. These coefficients reflected focused attention being deployed on a larger proportion of trials than the actual cue probability would suggest if participants were probability matching.

Having considered the shape and location of the RT distribution in a multinomial maximum likelihood mixture analysis (MMLM). This approach allows a more rigorous test of the underlying assumption of the two-process model; that participants are either focusing their attention at the cued location or diffusely over all locations on any given trial. The focused attention trials would create a distribution of fast responses and the diffuse trials a distribution of slow responses, therefore neutral cues should result in a distribution representative of a diffuse state on every trial and 100% valid cues should result in a distribution representative of a focused state on every trial. A two process model with probability matching assumes that performance with a probabilistic cue should create a third distribution that is a mixture of the fast responses with focused attention and slow responses with diffuse attention, with the proportion of focused distribution trials being proportional to the cue reliability and the remaining trials being representative of the diffuse distribution. The MMLM analysis determines if the empirical data distribution is in fact representative of a mixture of these two distributions in the expected proportion of trials. Johnson and Yantis (1995) argued that there was not enough evidence to support the probability matching hypothesis, in the context of a two process model, for half of the participants and, for the other half, there was not enough statistical power to accept the hypothesis. They also provided evidence for an alternative one-process model, which assumes that the sampling rate of local sensory information reflects the probability of observing the target at that location. In this model attention is a resource allocated in parallel but non-uniformly according to the attentional priorities based on spatial expectancies. A limited number of perceptual samples are allocated according to priorities determined by top down factors of expectancies and search strategies, in theory being liable to change based on new evidence from perceptual samples (though this paper did not consider this possibility for changing sampling weights). The accumulation of evidence takes the form of a random walk or diffusion process similar to the first stage of Carpenters LATER model (Carpenter, 2009) though Johnson and Yantis (1995) assume that attention utilises probability by affecting what is essentially the first stage of the LATER model and alters the detection stage perceptual processes before their integration into a decision variable instead of being a component of decision variables derived from. Models of attention that consider sampling rate of visual information often propose that the processing power for sampling objects is derived from a limited capacity resource and the sampling density is divisible over multiple objects (d'Avossa Shulman, Snyder and Corbetta, 2006). More recent evidence suggests that endogenous attention operates by increasing the rate of information accrual but also by improving the overall discriminability of stimuli such that a greater asymptotic degree of visual sensitivity measured by d' is achieved in addition to an increased rate of processing (Giordanno McElree and Carrasco, 2009). This model produced simulated data similar to those obtained in participants, though it underestimated costs on invalid trials. Thus, they concluded that a graded allocation of attentional resources provides a better account of probabilistic cueing. They also suggested that the analyses used in previous studies, based on mean RT data, cannot be used to conclusively support the two process model (Jonides,1983; Eriksen and Yeh, 1985; Madden, 1992; Van der Heijden, 1989). The idea that attention may combine a graded strategy and a switching strategy was first proposed by Sperling and Melchner's (1978) study, where participants searched for two letters or numbers contained within the two concentric arrays containing 20 alphanumeric characters, and were given instructions to divide their attention, in varying proportions, between an outer and an inner array prior to each block of trials. The authors suggested that participants switched attention between arrays while performing the task because the successful identification of a target in one array was negatively correlated with the identification of the target contained in the other array. However, attention could also be divided between the two arrays, based on the finding that the probability of identifying the two targets was greater than what would be expected if they only attended a single array.

How can attention be distributed spatially?

Recently it has been argued that attention can be divided among multiple locations (see Jans, Peters and DeWeerd, 2010 for a review). Many authors have attempted to refine the metaphor of the attentional spotlight either by suggesting that attention can be deployed either focally or diffusely (Jonides, 1981) or that the area of attentional facilitation can be flexibly expanded and contracted, with a reduction in attentional effects as the area attended is broadened (Eriksen and St, James, 1986). Others have suggested that attention is a spatially continuous function, which is best conceived as a gradient (Downing and Pinker, 1985).

A number of studies have explicitly examined whether multiple, non-contiguous locations can be attended separately. Eriksen and Yeh (1985) found that when discriminating a letter, which could appear at the cued location with an equal or unequal probability to the location opposite the cued one, response latencies were faster when the target appeared at the cued than the opposite location. This was the case even when the probability that either location would contain the target was equal, implying that attention is either moved from location to location or that attention is divided unequally between locations equally likely to contain the target. However, since these authors used a fast onset cue, it is possible that some of the attentional effects were exogenous rather than endogenous (Eriksen and Yeh, 1985).

Castiello and Umilta (1992) cued two locations simultaneously, using placeholders of different sizes, which were equally likely to contain the target. Reaction times increased with the size of the placeholder containing the target, suggesting that attention was distributed over an area whose size varied with that of the placeholder. A replication of this experiment, which also examined cueing effects at locations intermediate those cued, revealed that targets in locations between those cued, were also attended (Mckormic, Klein and Johnston, 1998) suggesting that attention was not divided between the two locations, but rather encompassed a contiguous region, as predicted by the zoom lens theory of attention (Eriksen and St. James, 1986). To examine whether multiple, non-contiguous regions can be attended without attention spreading to intervening regions, one study used a coloured grating to mark attended areas separately from intervening areas not to be attended (Gobell, Tseng and Sperling, 2004). The task was to localise a target in a cued region, among distractors presented in both the cued and uncued regions and false targets, visually identical to the target, but presented in uncued regions. As the spatial frequency of the color grating was increased, target localization performance worsened, suggesting that participants could attend non-contiguous locations separately, but that the ability to divide attention had limited spatial resolution. Interestingly, when the visual angle of the display was changed by changing the distance of the screen from the observer. Kramer and Hahn (1995) reasoned that target onset for stimuli located at unattended location may elicit exogenous capture, confounding the interpretation of a number of previous studies suggesting that attention spreads to locations intermediate those cued. Participants had to judge whether target letters were identical or different. The target locations were cued endogenously using placeholders. Distractor stimuli were either identical or different from the target letters. When distractors that resambled the target were presented at a location intermediate the placeholders, performance in identifying the targets was reduced. However if the visual onset of the distractors was masked, to prevent exogenous capture, distractors had no effects on performance . Hahn and Kramer (1998) expanded on this finding, by manipulating the quality of letter stimuli so that the two target stimuli were both low, both high or one low and one high quality stimulus with low quality stimuli taking longer to process. This manipulation was used to determine if both stimuli were simultaneously processed in parallel or if one stimulus was processed followed by the other. Under the parallel processing assumption processing time should be equivalent to time taken to process the most difficult stimulus in the pair but not be dependent on the quality of the other stimulus. Alternatively if stimuli were processed separately one after the other, then the total processing time should reflect the sum of the time taken to process both stimuli. Thus performance with two high quality stimuli should be fastest, followed by one high and one low quality, then two low quality stimuli taking longest. When there was no sudden onset the reaction times reflected the parallel processing of both stimuli in accordance with the former prediction, only the presence of low qualitiy stimuli mattered not the combination of both stimuli. This was in addition to a lack of interference by distractor stimuli placed between these targets, thus attention appeared to be at multiple non contiguous locations and be operating in parallel at these locations.

Bichot, Cave and Pashler (1999) used an array of eight shape outlines containing different letters. Two of the shapes, indicated by their colour, were the targets for the shape matching task. Participants also reported as many letters as possible from the array. Performance on the shape matching task was less accurate when targets were not adjacent. More importantly, participants recalled none of the letters contained in distractor shapes and while they could recall the letters contained within the target shapes. Also recall of letters at between target locations was no better than at other distractor locations, supporting the idea that attention was deployed only at the target locations.

Awe and Pashler (2000) presented participants a visual array containing two target numbers, among distractor letters. Two spatial cues preceded the letter array and had an 80% likelihood of indicating the target locations. On invalid trials, one of the targets was presented between the two locations cued while the other target appeared at a distant location. Identification accuracy was greater for targets at the cued locations than for a target at the intermediate location, while the latter was identified more accurately than a target at the far location. These findings suggest that attention was less, but not entirely absent between the cued locations and are consistent with the view that separate, broad attentional foci were used to attend the cued locations.

Pylyshyn and Storm (1988) devised the attentional tracking paradigm. They found that participants can track, with perfect accuracy, up to four or five moving objects among identical distractors, and inferred that attention can label as many objects simultaneously (Cavanagh and Alverez, 2005). This conclusion was challenged by subsequent multiple object tracking studies, which suggested either that perceptual grouping of visual targets facilitates tracking (Yantis, 1992) or that attention is quickly switched among the targets (Oksama and Hyona, 2004). The model put forward by Pylyshyn and his collaborators (Pylyshyn and Storm, 1988; Pylyshyn 2003; Sears and Pylyshyn, 2000) while emphasizing the importance of position as an index for accessing other dimension of visual sensory data, bears analogies to the proposal that visual information is represented mainly in the form of object files (Kahneman, Triesman and Gibbs, 1992) since it rests on the assumption that when attention is divided the parcelling is done among objects rather than locations.

How does learning cue target contingencies affect behaviour?

How statistical regularities of the environment affect behavioural performance has been extensively investigated . Animals can learn the probability of obtaining a reward associated with diverse courses of action when foraging for food (Herrnstein, 1961; Williams, 1988). Human observers can learn spatial contingencies as shown by studies which examined the effects of repeated visual displays on target detection (Chun and Jiang, 1998). Learning the probability that a target will appear at a given position in a spatial configuration is referred to as contextual cueing. Contextual cueing can also be found when the spatial location of the target varies across configurations, but its position in relation to the visual array remains constant (Miller, 1988). This suggests that fairly complex scene statistics can be learnt. These effects cannot be attributed to priming since they are observed even when several novel stimuli have been interleaved between repetitions. If probabilistic contingencies in visual scenes can be implicitly learned and exploited by observers, the same may be true for cue-target contingencies.

Droll, Abbey and Eckstien (2009) specifically investigated learning of unknown cue probabilities in an attentional task requiring detection of a luminance contrast increment. This study used peripheral cues in the form of coloured annular place-holders, each colour being associated with a different probability of containing the target information. The cues' spatial location was changed over trials. The authors used various feedback conditions including: 1) an unsupervised condition in which no feedback was given. 2) A reinforcement condition in which only feedback about task performance was given and 3) a supervised condition in which information was given about whether the target had appeared and its location. Learning was greater for fully supervised than reinforcement learning, with the latter more effective than unsupervised learning. There was actually no evidence of learning in the unsupervised learning condition. In addition, explicit reports of the cue probability by participants, were quite accurate, even for the unsupervised condition, showing either that knowledge of cue probability by itself does not affect performance, contrary to uncertainty based models of cueing, or that explicit and procedural knowledge of the cue probability may be acquired separately.

We follow three main lines of inquiry in the empirical chapters to follow:
Firstly we investigate the effects of varying the reliability of an endogenous central cue on the discrimination of poorly visible motion and the detection of highly visible motion in separate tasks. We also consider if the two-process model with probability matching can explain data for both the response latencies to motion detection and accuracy of motion discrimination, the latter not having been considered previously in the literature. A difficult discrimination task may be a more likely candidate for

needing to have a more focused beam of attention and thus employ a two process model.

Secondly we consider how attention is distributed across multiple locations, whether there is there a cost for this division when spatial uncertainty is accounted for and if there is a cost, from where do these limitations in distributing attention arise? To address this question we use probabilistic cues and multiple location cues requiring differing distributions of attention but which require the same spatial information. We then employ working memory and discrimination tasks requiring recall of multiple cued locations and an information theoretic model to consider how spatial working memory processes may contribute to limitations in dividing attention.

Thirdly we investigate whether it is possible to learn and adapt to changes in cue reliability and consider over how many trials this learning process takes place with only recent trials taken into account or a larger history of trials contributing to the learning. We employ periodic changes in cue reliability that are unknown to the participants and consider the effects of these changes on perceptual performance. In addition we run bootstrap simulations to determine whether performance of those who do adapt to changes in reliability could create the same effects when performance is relative to only the present and three preceding trials.

Chapter 2

Effects of Cue Reliability on Accuracy and Reaction Times

Theoretical and empirical arguments suggest that utilization of attentional, probabilistic cues may be a form of choice under risk. We examined the relation between cue reliability, namely the proportion of validly cued trials, and performance in two spatially cued tasks: 1) direction discrimination of poorly visible motion and 2) speeded detection of highly visible motion.

In the discrimination task, for a given cue reliability overall accuracy was maximized by the corresponding accuracies in valid and invalid trials, rather than those observed with cues of a different reliability, except for cues of low reliability. A stochastic model, which assumes participants either attend the cued location or maintain attention diffusely, was used to estimate the probability of attending the cued location. In the discrimination task, the estimated probability matched the cue reliability, but in the detection task it showed a compressed range. These results indicate, under conditions of high target visibility, either a compressed representation of the cue reliability, if cue utilization is deterministic, or a probability matching strategy, if cue utilization is stochastic. Either account suggests a functional homology between attentional and reward related cue utilization.

The spatial expectancy, generated by a cue indicating the likely location of a highly visible target, has systematic effects on target detection latencies (Posner, 1980). The initial interpretation of this finding was that the alignment between novel sensory events and an internal attentional axis is required for a behavioural response. In valid trials, where this alignment can be anticipated since the target location is correctly indicated by the cue, response latencies are faster than on invalid trials, where attention and the sensory event are misaligned because the target has appeared at an uncued location. Later, Posner and Petersen (1990) further elaborated this model and suggested that directing attention may require a sequence of operations, which include disengaging attention from the currently attended location, moving attention to a new location and reengaging attention at that location.

Cueing has been used to study how expectations influence not only response latencies, but also the accuracy of perceptual reports (Sekuler and Ball, 1977; Sapir et al., 2005; Prinzmetal, Park and McCool, 2005). Two theoretical viewpoints have informed accounts of cueing effects on visual sensitivity. One regards sensory processes as limited capacity channels, unable to accommodate ever-increasing amounts of information, without the internal representation of the sensed data incurring some degradation. Accordingly, the main role of attention is that of gatekeeper, which filters sensory data, so that limited resources can be devoted to the analysis of task relevant information (Broadbent, 1958). An alternative view is that the early analysis of sensory data has no intrinsic limitations (Eriksen and Spencer, 1969; Shiffrin and Gardner, 1972). The role of attentional cueing is to weigh the available sensory evidence once a perceptual decision has to be made Kinchla, Chen and Evert, 1995). Either view emphasizes the role of attention in selecting or prioritizing goal relevant sensory information (Broadbent, 1958; Sperling 1960; Von

Wright, 1970; Driver, 2001; Remington and Folk, 2010).

The proportion of validly cued trials, which we shall refer to as cue reliability, has been shown to modulate the magnitude of the validity effect, namely the difference between response latencies on valid and invalid trials (Jonides, 1980; Madden, 1992; Eriksen and Yeh, 1985; Roggio and Kristner 1997; Bowman et al., 1993; Vossel, Thiel and Fink, 2006). This finding indicates that participants adjust the spatial distribution of attention according to the reliability of the cue. On the other hand, the effects of cue reliability on accuracy have only recently been examined. For example, Giordano, McElree, Carrasco (2009) found that both overall amount and rate of target information accrual are affected by cue reliability.

The effects of cue reliability on performance may help understand whether endogenous shifts of attention depend on a dedicated neural system or on processes shared with other functions. For example, functional imaging has indicated that the control of eye movements and attention depends on largely overlapping cortical networks, suggesting that attentional operations in the spatial domain share processes involved in oculomotor functions (Corbetta and Shulman, 2001) in agreement with inferences drawn from behavioural observations (Hoffman and Subramaniam, 1995). In a similar vein, it has been argued, on logical and methodological grounds, that the selection of task relevant sensory information and choice of current goals are difficult to separate in behavioural paradigms (Maunsell, 2004). This point naturally leads to the inference that neural processes involved in setting attentional priorities may generalize to other cognitive domain, such as reward.

There is growing experimental evidence to support this view. Sapir et al. (2005) reported that in regions of frontal cortex associated with reward processes, trial to trial variations in the amplitude of preparatory signals recorded during the delay period between the cue and target presentation, predict the accuracy of motion discrimination at the time of target presentation. The interpretation of this result was that trial to trial changes in cue utilization may depend on trial to trial changes in the estimated cue utility. Single unit studies (Platt & Glimcher, 1997, 1999) found that the expected value of a primary reward and attentional cueing can influence neural activity to a similar degree in attentional regions, such as the lateral intraparietal area of macaques. Furthermore, detection thresholds and RTs are modulated in a similar fashion by either reward or attention (Leon and Shadlen, 1999; Kobayashi et al., 2002; Kawagoe et al, 1998; Lauwereyns et al, 2002; Hollerman et al., 1998; Ramnani & Miall, 2003). These findings are consistent with the overarching idea that attention and reward index largely overlapping processes (Maunsell, 2004).

If the internal representations of cue reliability and reward probability are shared then orienting attention following a cue is the same as choosing among variable rate reward schedules. Given that the way humans choose among alternatives associated with different reward probabilities can deviate substantially from the behaviour of a mechanisms that simply maximises the expected reward (Preston & Baratta, 1948; Kahneman & Tversky, 1979; Tobler et al., 2008), the implication of the above hypothesis is that the effects of cue reliability on performance may also be less than optimal.

The first model to explicitly address the behavioural effects of cue reliability, suggested that attention is shifted to the cued location with a probability that matches the cue reliability, while, on the remainder of the trials, attention is maintained diffusely. The model obviously implied a less than optimal strategy of cue utilization (Eriksen & Yeh, 1985; Jonides 1983; Madden, 1992; Van der Heijden, 1989). The proposal that attention reflects the superposition of a focused and a diffused state was heralded in a seminal study of divided attention (Sperling and Melchner, 1978). The authors concluded that when simultaneously attending two visual arrays, participants' performance reflects the combination of two strategies, the first based on mixing trials in which either one of the two arrays is attended, and the second on dividing attention between the two arrays.

However, a number of other models have been proposed since to account for the behavioural effects of partially valid cues. Johnson and Yantis (1995) suggested an early selection model of cueing. According to these authors, spatial expectancy determines the sampling density at each potential target location (see chapter one for more detail on this one process model). Others suggested that spatial expectancy is used to weight sensory evidence, leading perceptual decisions to be preferentially based on sensory evidence garnered at the cued location (Eckstein Shimozaki and Abbey, 2002; T, Eckstein, Pham, and Shimozaki, 2004). These models are consistent with an optimal use of cue reliability.

The purpose of the present study was twofold. First, we assessed whether spatial attention optimizes cue utilization given its reliability, since departures from optimality would support the hypothesis that cue utilization depends on reward related decision processes. Second, we examined whether the probability matching model of cue utilization accounts for the effects of cue reliability on performance in a difficult discrimination task and a speeded detection task.

Experiment 1 - Cue reliability effects on motion discrimination

Apparatus

Participants gave written consent prior to participation. The School of Psychology Ethics committee at Bangor University approved the protocol,

Six participants underwent five sessions of 600 trials each on consecutive days. The five participants took part also in Experiment 2.

The experimental paradigm was generated on an Apple Mac Pro 1.1, using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) for Matlab© (Mathworks, 2008a, Natick, MA). Stimuli were presented on the screen of a LaCie Electron 22blue CRT monitor with a refresh rate of 60Hz, placed at a distance of 70cm from the participant. Participants' head position was restrained by a chinrest.

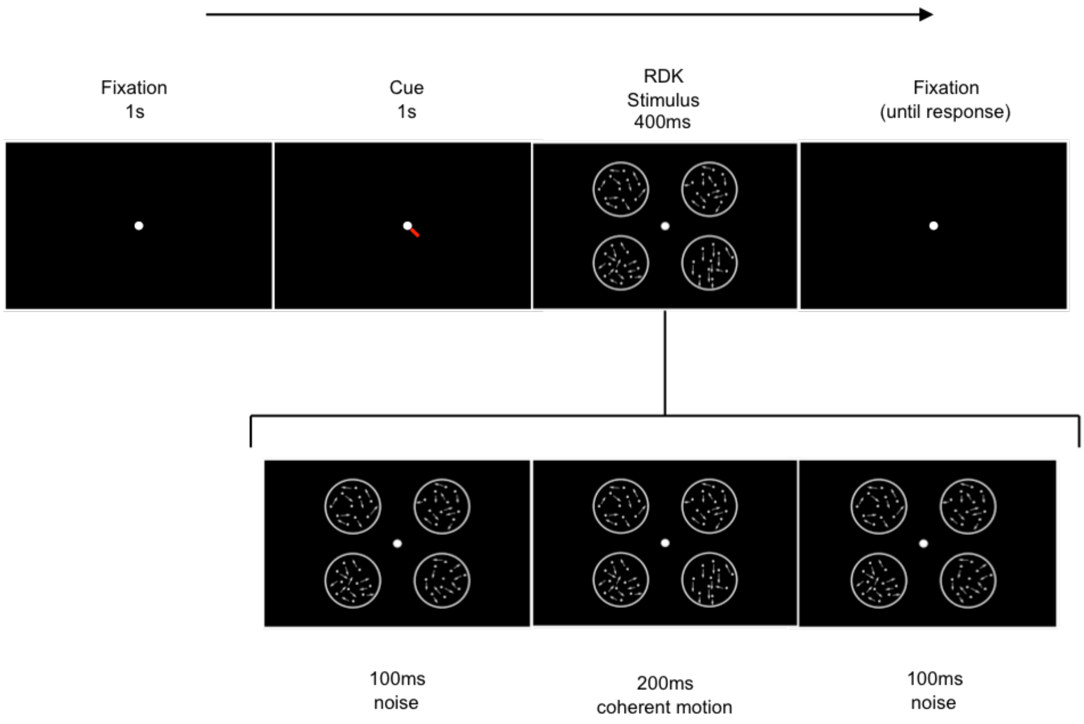
Procedure

The trial structure is illustrated in figure 1. Initially a fixation point was presented for 1.0s. This was followed by a coloured line at the centre of the screen for 1.0s, which indicated the likely target location. Four random dot kinematograms (RDKs) were displayed for 400ms immediately after the cue offset. Each RDK contained 100 dots of 0.15°, in positive contrast, within a circular aperture of 7.26° in diameter, centred at an eccentricity of 7.26°. The target RDK had a 200ms period of coherent motion in one of four directions (left, right, up or down), in between two 100ms periods of incoherent motion. The non-target locations contained random motion for 400ms. Dots' speed was 6.17°/s. Dots' lifetime was two frames. Participants reported the direction of coherent motion by pressing one of four keys.

Prior to the main experimental session, participant became acquainted with all aspects of the task during a one hour long training session. The percentage of coherently moving dots to be used in the experimental session was then determined. In this session, the RDKs were presented following a non-directional cue, consisting of a colour change of the fixation mark lasting 1.0s. Feedback about discrimination accuracy was provided visually on each trial. Threshold motion coherence was determined in blocks of 60 trials using a 2 to 1 staircase procedure with a 2% coherence step. Once stable performance was achieved, the coherence that yielded 70% discrimination accuracy was estimated by data interpolation and used in the experimental sessions.

In the main session, partially valid, spatial cues were used. On valid trials the cue indicated the target location correctly. In the remainder of the trials, i.e. invalid trials, it indicated a location different from the one containing the coherent RDK. Cue reliability, namely the probability that the cue indicates the location of the coherent RDK, was varied over four levels, 0.25, 0.6, 0.75 and 0.86, respectively. The cue colour indicated its reliability and was either red, orange, green or white. Participants were informed about the relation between cue colour and cue reliability and were given the possibility to internalize the mapping during training. A neutral cue was presented on 10% of trials, and consisted of a colour change of the fixation mark lasting 1.0s.

Figure 1. Motion discrimination task. A valid trial with coherent motion in the bottom right aperture. The colour of the cue conveys the probability that the indicated location will contain the target motion. The RDK stimulus consisted 100ms of random motion followed by 200ms of coherent motion in a given direction (left, right, up or down) after which another 100ms of random motion was presented. Arrows indicate the direction of motion of a dot.



The four reliability values spanned equal increments on an information theoretic measure of spatial uncertainty. Since coherent motion can appear in any of four locations, initially the participant is uncertain about the location of the task relevant stimulus. This uncertainty can be quantified in terms of bits, namely the number of binary choices required to uniquely specify the target location (Shannon, 1948). Following the cue, the uncertainty of the participant decreases by an amount, which depends on the reliability of the cue. This amount, that is, the information transmitted (IT) by the cue, is the difference between the participant's spatial uncertainty, before and after the cue.

The formula below is used to compute the entropy, or initial uncertainty, H(L), for the discrete random variable l, namely the location of coherent motion, from its probability distribution over the four possible locations, namely p(l).

$$\mathrm{H}(L) = -\sum_{i=1}^{4} \mathrm{p}(l_i) \cdot \log_2 \mathrm{p}(l_i) \tag{1}$$

In this task, the initial spatial uncertainty is 2.0 bits. The reduction in spatial uncertainty, that is the information transmitted (IT) by the partially valid cue is a function of the cue reliability, rel, and is given by:

$$\mathrm{IT}(rel) = \mathrm{H}(L) + \left[rel \cdot \log_2 rrl + (1 - rel) \cdot \log_2 \frac{1 - rel}{3} \right] \qquad (2)$$

For a cue reliability of 0.25, the IT is zero, since the cue does not change the spatial uncertainty. A cue reliability of 0.6 results in an IT of 0.4 bits. 0.8 bits of information are transmitted by a cue of 0.75 reliability, whereas 1.2 bits are transmitted by a cue of 0.86 reliability. The cue properties are illustrated in Table 1 below

Cue Type	**Initial Uncertainty (bits)**	**Information Transmitted (bits)**	**Uncertainty remaining (bits)**
1 location 0.25	2	0	2
1 location 0.6	2	0.4	1.6
1 location 0.75	2	0.8	1.2
1 location 0.86	2	1.2	0.8

Table 1. Information theoretic quantities of cueing conditions in four location stimulus set up. Illustrating cue type in terms of locations cued and probability of indicating target location, initial uncertainty in a stimulus with the amount of information transmitted by the cue in information theoretic terms (bits) and the remaining spatial uncertainty in the stimulus (bits)

<u>Fitting discrimination accuracy data</u>

To gain further insights into the central mechanisms underlining the effect of cue reliability on motion discrimination performance, we modelled the motion discrimination accuracy data using a two-process model (Jonides, 1983; Johnson and Yantis, 1995). This model assumes that participants either attended the cued location or attended all locations simultaneously on a given trial and that performance represents a mixture of these two trial types, with each occurring on a given proportion of trials. This analysis is employed to consider whether the observed data distribution could occur from a mixture of trials drawn from a distribution representative of focused attention, on a proportion of trials matching the cue reliability and drawn from a distribution representative of diffuse attention on the remaining proportion trials. Evidence in favour of the two-process model would suggest that this is the mechanism by which attention is deployed when discriminating poorly visible motion and show that probability matching occurs in perceptual behaviour. Such a probability matching strategy would be in accordance with the idea that decision processes are involved in utilising probabilistic information. The probability, m, of attending the cued location on a given trial depends on the cue reliability, rel. According to this model the discrimination accuracy, g(C), on valid and invalid trials is then:

$$\begin{aligned}
\mathrm{g}\left(C|rel\right)_{valid} &= \mathrm{m}\left(rel\right)\cdot p\left(C\right)^{attended} + \left[1-\mathrm{m}\left(rel\right)\right]\cdot p\left(C\right)^{diffuse} \\
\mathrm{g}\left(C|rel\right)_{invalid} &= \mathrm{m}\left(rel\right)\cdot p\left(C\right)^{unattended} + \left[1-\mathrm{m}\left(rel\right)\right]\cdot p\left(C\right)^{diffuse} \\
\mathrm{g}\left(C\right)_{neutral} &= p\left(C\right)^{diffuse}
\end{aligned} \qquad (3)$$

Point estimates of the free parameters were obtained using a maximum likelihood estimation (MLE) procedure. The posterior probability of observing the data given the model where:

$$\begin{aligned}
&\mathrm{p}\left(data_{valid}\left|p\left(C\right)_{attended}, p\left(C\right)_{diffuse}, \mathrm{m}\left(rel\right)\right.\right) = \prod_{rel}\left[\frac{n\cdot rel!}{n\cdot rel\cdot p\left(C|rel\right)_{valid}!\, n\cdot rel\cdot\left(1-p\left(C|rel\right)_{valid}\right)!}\right]\cdot \mathrm{g}\left(C|rel\right)_{valid}^{p\left(C|rel\right)_{valid}\cdot n\cdot rel}\cdot\left[1-\mathrm{g}\left(C|rel\right)_{valid}\right]^{\left[1-p\left(C|rel\right)_{valid}\right]\cdot n\cdot rel} \\
&\mathrm{p}\left(data_{invalid}\left|p\left(C\right)_{unattended}, p\left(C\right)_{diffuse}, \mathrm{m}\left(rel\right)\right.\right) = \prod_{rel}\left[\frac{n\cdot\left(1-rel\right)!}{n\cdot\left(1-rel\right)\cdot p\left(C|rel\right)_{invalid}!\, n\cdot rel\cdot\left(1-p\left(C|rel\right)_{invalid}\right)!}\right]\cdot \mathrm{g}\left(C|rel\right)_{invalid}^{p\left(C|rel\right)_{invalid}\cdot n\cdot\left(1-rel\right)}\cdot\left[1-\mathrm{g}\left(C|rel\right)_{valid}\right]^{\left[1-p\left(C|rel\right)_{invalid}\right]\cdot n\cdot\left(1-rel\right)} \\
&\mathrm{p}\left(data_{neutral}\left|p\left(C\right)_{diffuse}\right.\right) = \left[\frac{n_{neutral}!}{n_{neutral}\cdot p\left(C\right)_{neutral}!\, n_{neutral}\cdot\left(1-p\left(C|rel\right)_{neutral}\right)!}\right]\cdot \mathrm{g}\left(C\right)_{neutral}^{p\left(C\right)_{neutral}\cdot n_{neutral}}\left[1-\mathrm{g}\left(C\right)_{neutral}\right]
\end{aligned}$$

(4)

The log-likelihood function was computed by summing the logarithm of the posterior probabilities for valid and invalid trials, that is:

$$\Lambda = \log\left[\mathrm{p}\left(data_{valid}|\ldots\right)\right] + \log\left[\mathrm{p}\left(data_{invalid}|\ldots\right)\right] + \log\left[\mathrm{p}\left(data_{neutral}|\ldots\right)\right] \qquad (5)$$

The log-likelihood function was maximised using a simplex algorithm (Nelder and Mead, 1965).

Results and discussion

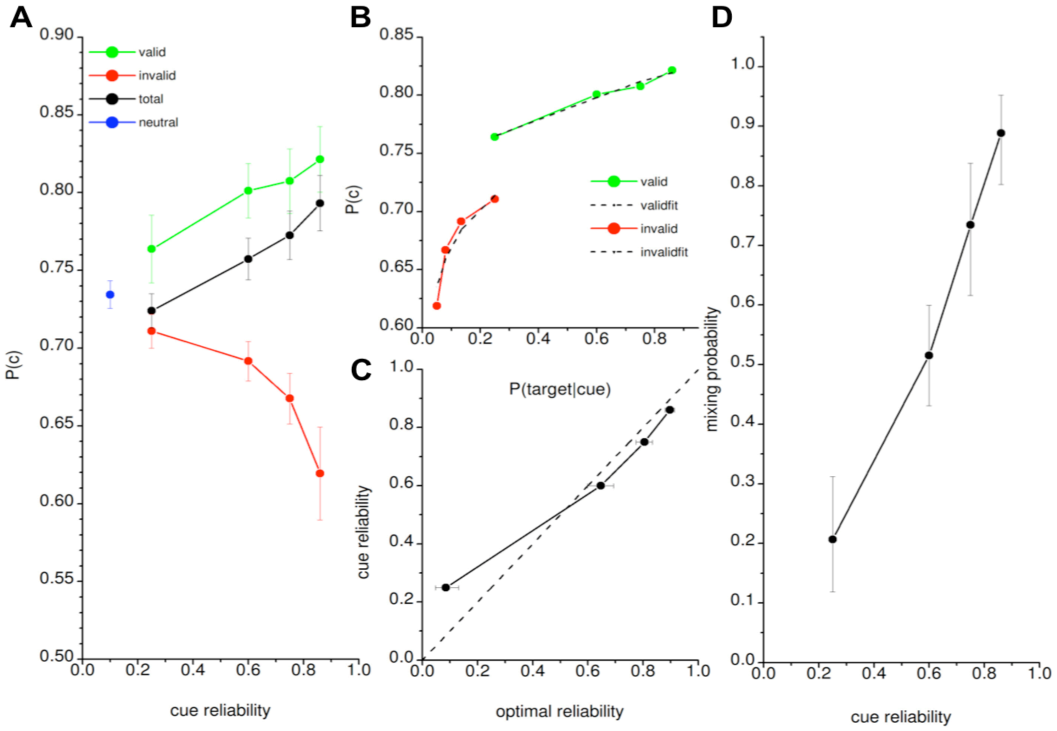
We assessed the effect of central cues reliability on motion direction discrimination. The greater the reliability of the spatial cue the greater the overall motion discrimination accuracy, as shown in Figure 2A. Furthermore, the validity effect increased with cue reliability. An unexpected finding was the higher accuracy on valid than invalid trials for the 0.25 reliability cue, even though this cue provided no useful spatial information (t [7] = 2.13, p =0.07). Moreover, the overall accuracy was lower than the accuracy for neutrally cued trials, suggesting that the use of the non-informative cue, as indexed by the validity effect, was disadvantageous though this was not statistically significant (t [7] = 1.41, p=0.2). A two-way, repeated measures ANOVA was used to test the significance of the effects of cue validity and cue reliability. The dependent variable was the inverse sine of the square root of the accuracies, a common normalizing transformation for proportional data. There was a significant main effect of validity (F [1,4] = 23.14, p <0.01), but no significant effect of reliability (F [3,12] = 1.08, p =0.39). The interaction of reliability by validity was significant (F [3,12] = 7.15, p<0.005) indicating that cue reliability significantly affected the size of the validity effect. A one-way ANOVA confirmed that reliability significantly affected the overall discrimination accuracy (F [3, 19] = 3.66, p<0.05). Figure 2B shows the relation of discrimination accuracy on valid and invalid trials to cue reliability. Each participant's accuracy data were fit using two logistic regressions, one for valid trials and one for invalid trials. The independent variable was the conditional probability of the target location given the cue, for valid trials, and the logarithm of the conditional probability, for invalid trials. Figure 2B also shows the resulting average fits and the group averaged accuracies as a function of the conditional target probability. The fits were used to estimate the effects of reliability on accuracy, for reliabilities between 0.10 and 0.95. Figure 2C shows the reliabilities estimated to correspond to the performance that would have maximized the overall accuracy for the cue reliabilities that were actually tested. For each cue reliability tested we calculated the reliability level whose associated performance on valid and invalid trials would maximize overall accuracy. If performance associated with a given reliability maximizes overall accuracy for the same cue reliability then it is inferred that participants have a veridical representation of reliability, otherwise, that the representation is biased. If the cue reliability which maximizes accuracy is greater than the actual reliability, participants underestimate reliability, vice versa, if the cue reliability which maximizes accuracy is smaller, participants overestimate it. To facilitate comparison of our data to published plots of the relation between internal estimates of event probability and event probability, used to illustrate biased representations of probability, we plotted cue reliability as a function of the reliability estimated to maximize overall accuracy. Not surprisingly, the overall accuracy, when 25% reliability cues were used, would have been maximized by performance with cues of reliability lower than 25%. On the other hand, for trials in which the cue reliability was 60% and higher, there was a trend for the overall accuracy to be maximized by performance with the objective probabilities or lower. There are two possible explanations for this pattern of results, either the representation of cue reliability is distorted, or cue utilization is based on a sub-optimal procedure that does not maximize expected accuracy as a function of cue reliability.

Next, the accuracy data were fit using a two-process model, which assumes that average performance reflects the mixture of two trial types, one in which the participants attended the cue location, the other in which participants maintained attention diffusely over all stimulus locations. The model required the estimation of seven parameters, namely accuracy on valid trials when the cued location was attended, accuracy on invalid trials when the cued location was attended, and accuracy on trials in which attention was maintained diffusely. Also, the probabilities that participants attended the cued location were estimated for each reliability level and participant. The two-process model accounted for the observed data adequately. The correlation between estimated and empirical accuracies across participants and conditions was greater than 0.97. A Montecarlo simulation indicated that the probability of observing this value was greater than 0.5 for distributions generated from the estimated accuracy values with the same set size as the empirical data. We conclude that the difference between model estimates and empirical data is statistically negligible.

Most important is the fact that group averaged mixing probabilities, plotted in figure 2D as a function of cue reliability, closely matched the cue reliability. This result may imply that departures from an optimal tuning of accuracy to cue reliability are not due to a distorted representation of cue reliability, but rather result from a less than optimal use of the cued information, that is a probability matching strategy.

Figure 2. Motion discrimination data. A) Group averaged accuracies as a function of cue reliability. Accuracy in valid trials (green circles), invalid trials (red), and overall performance (black) is displayed. Accuracy in neutral trials (blue) is also shown for reference. Error bars are between subjects standard errors of the mean. B) Logistic regression fit to accuracy data are shown as a function of the conditional probability

of the target appearing at the target location. Valid and invalid trials are fit separately. C) Relation between test cue reliability and cue reliability estimated to yield accuracies in valid and invalid trials, which would maximize overall accuracy on test cue reliability. The identity line defines the data location if performance was optimally tuned to performance. D) Group averaged estimates of mixing probabilities as a function of cue reliability.



Experiment 2 - Speeded motion detection task

The results of Experiment 1 indicated that a two-process model with probability matching accounts well for the discrimination accuracy data we obtained. This is in opposition to the literature looking at response latencies that employ maximum likelihood analyses (Yantis and Johnson, 1995), in which a two- process model with probability matching does not sufficiently account for the data when considering response distributions rather than just mean response times. Here we used the same cueing stimuli and observe their effects on the detection of a highly visible target motion stimulus. We also conduct a maximum likelihood analysis to interpret whether a two-process model with probability matching accounts well for response latencies to highly visible motion. Furthermore we plot the obtained response distributions in reciprobit plots to determine the applicability of the LATER model to these data (Carpenter, Reddi and Anderson, 2009).

Figure 3. displays the trial structure of experiment 2. Here we assessed the effect of cue reliability on speeded detection of a suprathreshold stimulus. Cue reliability levels were the same as those used in experiment 1. Following the cue presentation, lasting 1.0s, four RDK stimuli, containing incoherent motion, were displayed for up to 3200ms. During this period, expanding motion lasting 400ms could appear at one of the RDKs. The onset of the coherent motion was preceded by a non-ageing fore-period, whose latency had a truncated exponential distribution:

$$\mathrm{p}(delay = t) = \frac{1}{\tau} e^{-\frac{t-0.3s}{\tau}} \quad \text{iff } 0.3\text{s} \le t \le 2.4\text{s}$$

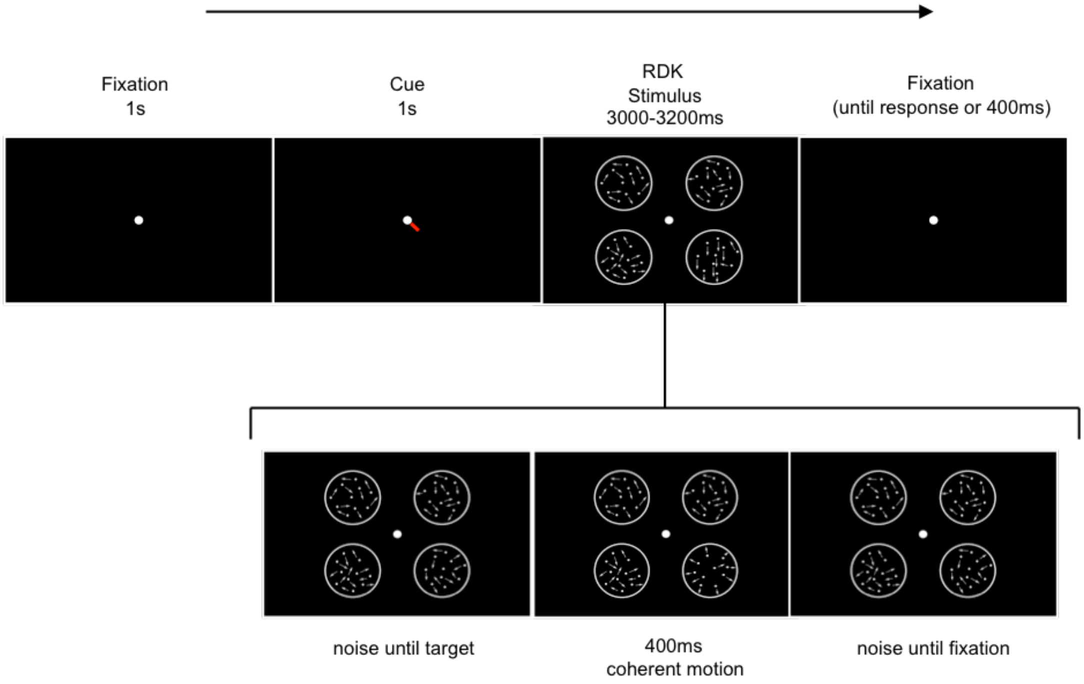
else

$$\mathrm{p}(delay = t) = 0 \tag{6}$$

Where τ is the time constant (τ=1.56) and t, the time from stimulus onset. Expanding motion consisted of linear displacements along radial directions at 100% coherence. Participants were instructed to respond as quickly and accurately as possible to the onset of coherent motion, by pressing a button on a Cedrus response box. In about 20% of the trials no target stimulus was presented to minimize anticipatory keypresses.

Figure 3. Motion detection task. The colour of the cue conveys the probability that the

indicated location will contain the target motion. The RDK stimulus consisted of a 3000-3200ms non-ageing fore period of random motion during which 400ms of expanding coherent motion would occur. Arrows indicate the direction of motion of a dot. Participants responded as quickly as possible by pressing a single button on a response box or withholding a response if no stimulus occurred during that period.



<u>Reciprobit plots</u>

To normalize the distribution of the response latencies, the RTs were transformed by computing their inverse. The group averaged cumulative distribution of inverse latencies were obtained as follows: the latencies were ranked from shortest to longest. Their rank was normalized to the interval (0-1). The probit score of the normalized rank, *r*, was computed as follows:

$$\mathrm{probit}(r) = \sqrt{2} \cdot \mathrm{erf}^{-1}(2r-1) \qquad (7)$$

where erf^{-1} is the inverse error function. Group averaged probit plots were obtained by interpolating each participant data with a cubic spline and resampling, in the interval [-2 +2] for valid trials, and [-1.5 1.5] for the invalid trials in equal steps of 0.075 and 0.05 probit scores respectively. The resulting cumulative distributions were then averaged across participants.

Fitting RTs in a simple speeded detection task

The reciprocal response latencies, a normalizing transformation previously shown to accurately describe RTs in speeded choice tasks (Carpenter and Williams, 1995; Carpenter, Reddi and Anderson, 2009) was fit using two models. The main question was whether the effects of cue reliability and validity on response latencies were more consistent with a switching strategy, in which subjects alternated between attending the cued location and maintaining attention across all stimulus locations, or rather attention is controlled deterministically since the same strategy is used in trials of a given cue reliability, for example by computing different decision thresholds for detecting targets appearing at cued vs. uncued locations. The first model included ten parameters, that is the means and standard deviations of the normal distributions of reciprocal response latencies on trials in which 1) the participant attended the cued location and the cue was valid, 2) the participant attended to the cued location and the cue was invalid and 3) the subject maintained attention diffusely over all possible stimulus locations.

$$\begin{aligned}
\mathrm{g}\left(rt^{-1}\right)_{valid} &= \mathrm{m}\left(rel\right)\cdot \mathrm{N}\left(\mu_{attended},\sigma^2_{attended}\right)+\left[1-\mathrm{m}\left(rel\right)\right]\cdot \mathrm{N}\left(\mu_{diffuse},\sigma^2_{diffuse}\right)\\
\mathrm{g}\left(rt^{-1}\right)_{invalid} &= \mathrm{m}\left(rel\right)\cdot \mathrm{N}\left(\mu_{unattended},\sigma^2_{unattended}\right)+\left[1-\mathrm{m}\left(rel\right)\right]\cdot \mathrm{N}\left(\mu_{diffuse},\sigma^2_{diffuse}\right)\\
\mathrm{g}\left(rt^{-1}\right)_{neutral} &= \mathrm{N}\left(\mu_{diffuse},\sigma^2_{diffuse}\right)
\end{aligned} \tag{8}$$

The remaining four parameters were the mixing probabilities, m, for trials of the four different reliability levels. On neutrally cued trails, participants were assumed to maintain a diffuse state of attention on each and every trial.

The probability of having observed the particular latencies distributions given the model estimates were computed as follows:

$$\begin{aligned}
\mathrm{p}(data|\ldots)_{valid} &= \prod_{rel}\prod_{i=1}^{n_{valid}}\left\{\mathrm{m}(rel)\cdot\frac{1}{\sqrt{2\pi}\sigma_{attended}}\exp\left[-\frac{\left(rt_{rel,i}^{-1}-\mu_{attended}\right)^2}{2\sigma_{attended}^2}\right]+\left[1-\mathrm{m}(rel)\right]\cdot\frac{1}{\sqrt{2\pi}\sigma_{diffuse}}\exp\left[-\frac{\left(rt_{rel,i}^{-1}-\mu_{diffuse}\right)^2}{2\sigma_{diffuse}^2}\right]\right\}\\
\mathrm{p}(data|\ldots)_{invalid} &= \prod_{rel}\prod_{i=1}^{n_{valid}}\left\{\mathrm{m}(rel)\cdot\frac{1}{\sqrt{2\pi}\sigma_{unattended}}\exp\left[-\frac{\left(rt_{rel,i}^{-1}-\mu_{unattended}\right)^2}{2\sigma_{unattended}^2}\right]+\left[1-\mathrm{m}(rel)\right]\cdot\frac{1}{\sqrt{2\pi}\sigma_{diffuse}}\exp\left[-\frac{\left(rt_{rel,i}^{-1}-\mu_{diffuse}\right)^2}{2\sigma_{diffuse}^2}\right]\right\}\\
\mathrm{p}(data|\ldots)_{neutral} &= \prod_{i=1}^{n_{neutral}}\frac{1}{\sqrt{2\pi}\sigma_{diffuse}}\exp\left[-\frac{\left(rt_{i}^{-1}-\mu_{diffuse}\right)^2}{2\sigma_{diffuse}^2}\right]
\end{aligned} \tag{9}$$

The likelihood function was calculated by summing the logarithms of the conditional probabilities, that is:

$$\Lambda = \log\left[\mathrm{p}\left(data_{valid}|\ldots\right)\right]+\log\left[\mathrm{p}\left(data_{invalid}|\ldots\right)\right]+\log\left[\mathrm{p}\left(data_{neutral}|\ldots\right)\right] \tag{10}$$

The second model included eighteen parameters, that is, the mean, μ, and standard deviations, σ, of the distribution of reciprocal latencies on valid and invalid trials of the four reliability levels and the neutrally cued trials. The conditional probability of the latencies distribution on valid trials, for all reliability level, rel, was then:

$$\mathrm{p}(data|\ldots)_{valid} = \prod_{rel}\prod_{i=1}^{n_{valid}}\frac{1}{\sqrt{2\pi}\sigma(rel)_{valid}}\exp\left[-\frac{\left(rt_{rel,i}^{-1}-\mu(rel)_{valid}\right)^2}{2\sigma(rel)_{valid}^2}\right] \tag{11}$$

and similarly for invalid and neutral trials. The log-likelihood function was computed the same way as for the first model.

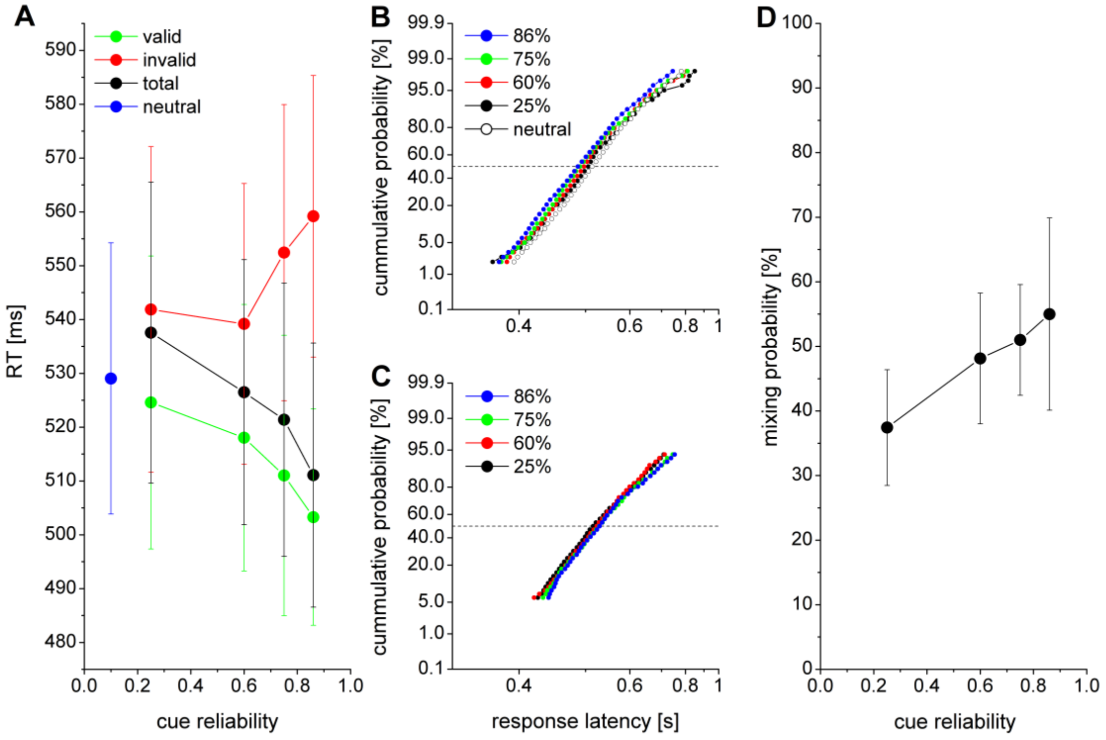
To determine which model accounted for the data better, we compared their log-likelihoods, which were expected to differ since the switching model yields bimodal RTs distributions while the deterministic model yields unimodal distributions, To adjust the comparison for differences in the degrees of freedom, we estimated the associated bias by repeating one hundred times the same procedure on pseudo-random data drawn from a normal distribution and computing the difference between the resulting log-likelihoods.

Results and discussion

The effect of cue reliability on reaction times was assessed in a task in which participants had to detect the onset of highly visible expanding motion.

Figure 4A shows the group averaged RTs for valid and invalid trials and overall RTs, for each reliability level, as a function of cue reliability. The overall RT for neutral trials, where no spatial information was provided, is also displayed. RT was faster for valid trials than invalid trials (F [1, 4] = 12.339, $p > 0.05$); this difference became larger with increasing cue reliability (F [3,12] = 0.21, $p < 0.05$). The overall RT decreased with increasing cue reliability, (F [3, 12] = 10.6, $p = 0.001$). Hence the overall effects appear similar to those obtained in the motion discrimination experiment, which used threshold rather than supra-threshold motion stimuli. Again a validity effect for the 25% cue reliability condition is found, with longer RTs than in neutrally cued trials (t [6]= 2.70, $p < 0.05$). We also observed that overall RTs for 25% reliable cues were lower than RTs to neutral cues suggesting the pattern of non-informative cues causing maladaptive behaviour extended to detection latencies, however again this was not significant (t [6]= 1.38, $p = 0.22$).

Figure 4B and 4C show group averaged reciprobit plots, in which reaction time distributions are plotted with reciprocal latency as abscissa and a probit scale of cumulative percentage probability as ordinate. These plots concern the response latencies for valid neutral trials, and invalid trials respectively. Previous work has shown that RTs distributions, in speeded choice tasks, follow a linear trajectory when thus plotted (Carpenter and Williams, 1995). These distributions seem to deviate from a straight line and do not show changes in the slope of the data with changing reliability, thus there appear to be only small effects of reliability in this study in comparison to carpenter and colleagues findings that can be observed when plotted thusly. Next we determined which of two models of cue utilization provided a better fit to the inverse RT distributions. To determine whether the reciprocal RT distributions were better fit by a two-process model, or by a normal distribution whose mean and standard deviation varied with cue reliability and validity, we computed the log likelihoods of the fits obtained using the two models. In six of the seven participants the inverse RT distributions were better fit by the two process model. Moreover, the group averaged log-likelihood of the two process model was significantly greater ($t[6]=2.47$, $p<0.05$), despite the two process model having a lower number of free parameters.. Figure 6d shows the mixing probabilities, estimated by fitting the two process model, as a function of cue reliability. Clearly the effect of cue reliability on the probability of attending the cue location is not prominent suggesting that participants undermatched cue reliability. *Figure 4. Mean reaction times as a function of cue reliability. A) Performance in valid trials (green circles), invalid trials (red), and overall performance (filled) are shown. Error bars are the standard errors of the mean. B) Distribution of response latencies for all cue reliabilities in valid and neutral trials. Data plotted in reciprobit scale i.e. on a reciprocal latency scale as abscissa and a probit scale of cumulative percentage probability as ordinate. C) Distribution of response latencies for all cue reliabilities in invalid trials on a reciprobit scale D) Group average mixing probabilities as a function of cue reliability.*



General discussion

We investigated the effect of spatial cues reliability on discrimination of threshold motion stimuli and detection speed of salient motion stimuli. For both tasks, the difference in performance between valid and invalid trials increased with cue reliability: the greater the cue reliability the greater the validity effect, replicating previously reported findings (Jonides, 1980; Madden, 1992; Eriksen and Yeh, 1985; Roggio and Kristner 1997; Bowman et al., 1993). Greater cue reliability was also associated with greater overall accuracy and faster overall reaction times, underscoring the adaptive nature of the reliability effect. However, there was also a noticeable departure from optimal use of cue reliability. In fact, a validity effect was found for trials with a 25% reliability cue, which provided no useful spatial information, in accordance with previous observations (Hommel et al., 2001; Ristic et al., 2002; Tipples, 2002; Giordano, McElree, Carrasco, 2009). Moreover, in these trials overall accuracy was also lower and reaction times longer than in neutrally cued trials (though not significantly so), giving some implication that the validity effect for a non-informative cue is maladaptive. To further examine the relation between cue reliability and performance, we estimated how closely discrimination accuracy on valid and invalid trials was matched to cue reliability. By fitting a continuous function inaccurate to the relation between accuracy and reliability, we found a good congruence between cueing effects and cue reliability. This can be inferred by the similarity between the reliability estimated to give the best performance and the observed performance at the tested cue reliabilities. However, for a 25% reliability cue overall accuracy was maximized by performance associated with cues of lower reliability. Thus validity effects are not perfectly tuned to the cue reliabilities. There are at least two possible explanations: either the representation of cue reliabilities is biased or the utilization of the cued information is based on a procedure that is less than optimal. Vincent (2011) recently found evidence for biases in perceptual performance for both endogenous and exogenous cueing. The model that best accounts for this performance considers slightly biased prior beliefs akin to those outlined in prospect theory (Khaneman and Tversky, 1979), are combined optimally with sensory evidence. Though this assumption of prior beliefs being biased implies that distortions in perceptual behaviour are universal across tasks whereas we find seemingly very different dynamics of attention strategy and deployment between tasks. It has been suggested that these biases may be adaptive in naturalistic settings according to Adaptive Probability Theory (Martins, 2005, 2006) as probabilities in the environment are often dynamic and that uncertainty about spatial expectations accounts for these distortions in choice behaviour and seemingly also in perceptual decisions. This supports the idea that the distortions we find may in fact be representative of this internal bias that also becomes apparent in financial decisions. In this vein Fennel and Baddley (2012) propose that this adaptive mechanism to assume that probabilities are non static explains the probability weighting functions outlined in prospect theory (Kahneman and Tversky, 1979). This is accounted for in terms of incorporating previous knowledge with the objective probability statement that is assumed to be inherently uncertain, though in novel contexts the influence of a uniform non biased ‘ignorance’ prior may be integrated with the usual biased ‘inference’ prior. Only with repeated exposure to an appropriate context might the observer be free from these biases’ influence. Estimates of the probabilities that participants were attending the cued location, based on the assumption that they either attended the cued location or maintained attention diffusely, closely matched cue reliabilities in the discrimination task. This result is consistent with the hypothesis that the central representation of cue reliability is unbiased, but cue utilization is limited by a suboptimal, probability matching procedure. Optimal cue utilization would be achieved, under the same set of assumptions, had participants used a maximized probability, that is on trials preceded by low reliability cues participants always maintain their attention diffusely and on trials with high reliability cues they always attend the cued location. Clearly this was not the strategy used since accuracies on valid and invalid were strictly monotonic, rather than step functions of cue reliabilities. The suggestion that inappropriate cue utilization, rather than biased reliability representation may underline departures from optimal performance may gain further support from the finding that in the speeded detection task the estimated mixing probabilities were not well matched to the cue reliabilities, contrary to what we found in the motion discrimination experiment. Since the same cueing procedure was used in the two tasks, it is plausible to infer that the representation of cue reliability in the two tasks did not differ, and hence differences in mixing probabilities were due to differences in cue utilization.

The inference that participants do not use the cue in an optimal way despite a fairly accurate representation of its reliability finds support in functional imaging data showing distinct neural representation for the information provided by a cue and its utilization (O'Doherty et al. 2004). Moreover, behavioural data also support a distinction between the representation of cue reliability and its utilization. In fact, after extensive exposure to cue target pairings, human subjects can report accurately the reliability of cues and still demonstrate little effect of cueing on perceptual performance (Droll, Abby & Eckstein, 2009). This evidence for a probability matching strategy is also in line with the idea of sub optimalities, observed in choice behaviour being present in perceptual behaviour, suggesting that decision processes are closely entwined with attentional deployment and the strategies used in perceptual behaviour. The 'sub optimality' in this setting again may be adaptive in naturalistic settings and probability matching is especially indicative of the assumption that spatial probabilities will vary in the environment in line with Martins (2005, 2006).

How is attention shared among stimulus locations?

There are two main models for how allocation of limited resources is redistributed across the visual field, following the presentation of a central probabilistic cue. The first is that attention is divided among all possible target locations (Johnson and Yantis, 1995). At each location, the attentional weight, reflecting the proportion of available central resources devoted to analyzing sensory information, is consistent across trials and reflects the participant's expectation that the target will appear at that location: the higher the expectation the higher the weighting. Others (Jonides, 1983), have suggested a two-process model, which assumes that participants alternate, in separate trials, between either paying attention to the cued location or maintaining a diffuse mode of attention. The rate at which participants engage in either attentional mode depends on the cue reliability, so that the higher the reliability the greater the proportion of trials in which participants attend the cued location. This would be akin to the behavioural effects of probabilistic reinforcement schedules on choice behaviour in animals (Hernstein, 1961). The two-process model was found to account well for the effects of reliability on motion discrimination accuracy across reliability levels. However, the comparison of RT data fits using the two models did not provide conclusive evidence that a two process model is the only strategy used by participants.

Functional imaging data have repeatedly shown that utilization of a partially valid cue preceding a threshold stimulus changes trial to trial, both in higher order regions (Sapir et al., 2005) as well as in visual regions (Sylvester, Shulman, Jack and Corbetta, 2007; Sylvester, Jack, Corbetta and Shulman, 2008) providing direct support for a two process model. On the other hand, data from choice reaction time tasks for suprathreshold stimuli, cueing the location of the target, and the associated motor response, i.e. leftward and rightward saccade, suggest that cue reliability has a graded effect on response latency distributions (Carpenter and Williams, 1995).

Conclusions

The effects of cue reliability depended on nature of the task, suggesting that cue reliability and cue utilization may not rely on the same representation.

We could not conclusively establish whether the effects of cue reliability should be attributed to a two-process mechanism or a graded division of attention for the reasons stated above. Our results seem most consistent with the possibility that attention may be both shared across locations as well switching between modes in separate trials. It should be noted that if attention is divided among potential target locations by shifts of attention during stimulus presentation (e.g. d'Avossa et al., 2005), especially if shifts can be accomplished rapidly (van Rullen, Carlson & Cavanagh, 2007) then attentional switching and attentional sharing could be accomplished using the same set of central mechanisms.

Chapter 3

Effects of probability and uncertainty on the distribution of spatial attention

What are the limitations in dividing spatial attention? We examined this question in a motion discrimination task. Several random dot kinematograms (RDKs) were preceded by central cues indicating one or multiple locations likely to contain a coherent RDK. When four RDKs were presented, discrimination accuracy was identical when cueing one or two locations. However, with six RDKs, accuracy was higher when cueing one rather than multiple locations. To test whether this finding reflects a failure in remembering the locations cued, participants were asked to recall the cued locations. Participants were at ceiling when they only had to recall the cued locations, but were inaccurate when they also had to shift attention to these same locations. This suggests that attending interferes with maintaining a representation of the cued locations in working memory. Dividing spatial attention is limited by inaccurate representations of spatial locations in working memory.

Attention comprises functions that either allocate capacity limited resources to task relevant stimuli (Pashler & Johnston, 1989; Posner & Snyder, 1975; Shaw, 1978; Shaw & Shaw, 1977) or select task relevant sensory information for further processing (Broadbent, 1958; Sperling, 1960; Von Wright, 1970). Triesman and Gelade (1980) took the former viewpoint and posited that attention allows binding of multiple features belonging to the same object. This hypothesis helped establish the view that visual perception can be divided into pre-attentive and attentive stages, the former characterized by unlimited capacity to process basic aspects of the sensory input, the latter by limited capacity to process complex aspects of the visual array. Whether attention itself is a capacity limited, as suggested by Treisman and Galade (1980), or unlimited capacity process, is an issue that has occupied numerous investigators since.

Broadbent's (1958) proposal that attention prioritizes processing of task relevant information at the expense of task irrelevant information, leads naturally to recast the issue of whether attention is capacity limited in terms of the following question: Are there limits to how sensory information can be selected?

Several studies have examined the effects of divided spatial attention to determine whether selecting sensory information from more than one location is possible. Some suggested that attending multiple objects requires selection of one contiguous area of the visual field (Posner, Snyder, & Davidson, 1980; Eriksen & Hoffman, 1973; Eriksen & Yeh, 1985; Eriksen & St. James, 1986), thus negating the possibility that spatial attention may be truly divided.

However, Shaw (1978) concluded that allocation of limited capacity processes to a particular location, depends on the probability of finding the target there, a view consistent with an unlimited ability to divide attention. To account for performance in tasks preceded by probabilistic spatial cues, Jonides (1983) proposed instead the "two-process model". The model assumes that observers either process information indiscriminately from the entire display, or prioritize information presented at the cued location, and process it with high efficiency. The likelihood of engaging in the latter strategy is determined by the probability of the target being at the cued location. In contrast, Eriksen and Yeh (1985) found that target probability was a poor predictor of attentional effects in a cued visual discrimination task. Nevertheless, the possibility that more than one location can be attended has continued to receive experimental support (Castello & Umilta, 1992; Gobel et al., 2004; Bichot, Cave, & Pashler, 1999; Kramer & Hahn, 1995; Awh & Pashler, 2000).

While the idea that attentional capacity should reflect limitation in sensory selections, for example when multiple targets or locations need to be attended, is intuitively appealing there is no consensus about what exactly defines capacity limits in attentional selection. One proposal is that limitations in attentional selection reflect the finite rate at which information can be sampled. Since the overall sampling rate cannot be increased *ad libitum*, then when more than one target is attended, each target must be sampled at a diminished rate (d'Avossa, Snyder, Shulman, & Corbetta, 2006). A similar view has been used in recent work to define the limitations of visual working memory (Bays, Catalao, & Husain, 2009). The authors, having observed that the precision in recalling the positions of visual targets decreased with the number of targets, concluded that the resolution of spatial recall is limited by a finite resource that is shared between the targets. Others have suggested that attentional capacity limits are only defined by the number of separate targets that can be attended simultaneously, over displays that minimize demands placed on attentional resolution. On these grounds it has been suggested that the number of targets that can be attended and hence the capacity of attention is potentially infinite (Franconeri, Jonathan, & Scimeca, 2010).

In the present work we explore how dividing attention among centrally cued, peripheral target locations affects performance in a demanding motion direction discrimination task. The question is whether dividing attention places any limitations on performance beyond those predicted by changes in spatial uncertainty. To answer this question, a probabilistic single location cue or a multiple location cue providing equal spatial information but in a different distribution was presented before the target display. If attention cannot be spatially divided without incurring some behavioral cost, then discrimination accuracy should decrease as the number of cued locations is increased. On the other hand, if there is no limit to the number of locations that can be separately attended, then changing the number of cued locations should not affect discrimination accuracy. In addition we employ a multiple location anti cue, to further investigate the non-informative cueing effects that we observed in Chapter 1.

Experiment 3- four locations information matched cues

Participants

Three of the authors and three naive participants (five males) completed three sessions of 600 trials each, on three consecutive days. Each session lasted approximately 45 minutes. Naïve participants gave written informed consent. The studies were approved by Bangor University's ethics board.

Stimuli and apparatus

Stimuli were generated on an Apple Mac Pro 1.1, using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) for MATLAB$^{©}$ (Mathworks, 2008a, Natick, MA).

Stimuli were presented on the screen of a LaCie Electron 22blue CRT monitor with a refresh rate of 60 Hz, set at a distance of 70cm from the participant. The participants' head position was restrained by a chinrest.

The target display contained four RDKs, lasting 400ms. Each RDK contained 100 dots in positive contrast. Each dot diameter was 0.15°. The dots' speed was 6.17°/s, with a lifetime of three frames.

Each RDK was contained within a circular aperture of 7.26° in diameter, centered at an eccentricity of 10.9° along the main diagonals. The target RDK consisted of a 200ms period of coherent motion in one of four directions (left, right, up or down). Coherent motion was preceded and followed by 100ms periods of incoherent motion.

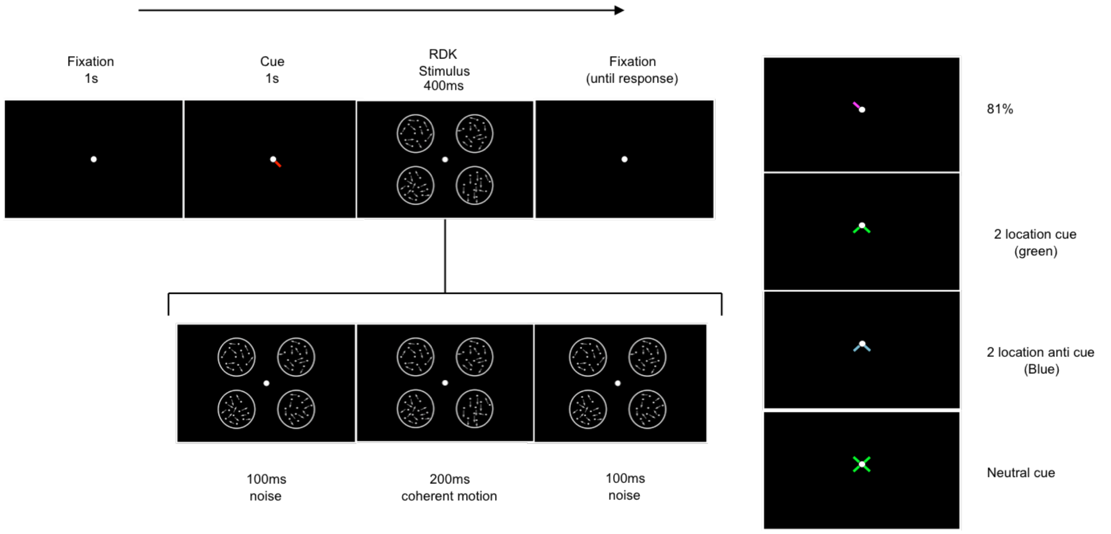
Procedure

The percentage of coherently moving dots, namely the stimulus coherence, was determined in a preliminary session. In this session, the RDKs were presented following a non spatial cue, consisting of a colour change at the point of fixation lasting 1.0s. Threshold motion coherence was determined using a 2 to 1 staircase procedure with a 2% coherence step. The level of coherence resulting in an accuracy of 70% was estimated by data interpolation, and used during the experimental sessions. During this training pre-session, participants were given visual feedback about their discrimination accuracy on each trial.

In the main experimental sessions, four cue types were used. The first cue type consisted of two oblique colored lines abutting the fixation point, which highlighted two RDK locations. One of the cued locations contained the coherent RDK. The second cue type consisted of two oblique lines but presented in a different color to the first cue, this indicated two locations that would not contain the target, we refer to this as the multiple location anti-cue. The third was a probabilistic cue, which indicated the location where the target RDK would appear on 81% of the trials. The fourth was a neutral cue, consisting of four oblique lines pointing to each RDK location and provided no spatial information. The different cue types and experimental procedure are shown in figure 5. The design of the experiment 3 is intended to compare discrimination performance in single location probabilistic conditions and multiple location cue conditions that provide the same spatial information, in information theoretic terms, but require drastically different distributions of attention. Thus, if distribution of attention for motion discrimination is limited in the manner assumed by two process models supported by chapter 2 there should be a cost for deploying attention to multiple locations. This information-matching concept is quantified below.

In addition we included a multiple location anti cue, this cue is identical to the multiple location pro cue in terms of the physical stimulus (except for color to distinguish the two cue types) and in terms of the indicated distribution of spatial information. The only difference between these two cue types was the fact that the pro cue stimulus indicated the possible target locations and the anti cue stimulus indicated the incorrect locations, which participants were explicitly informed of thus the information provided and distribution of that information was identical but the physical stimulus differed. As such if there is some aspect of endogenous cues that directs attention to cued locations regardless of the actual spatial distribution of probability, the anti cue should elicit such effects. The two location anti cue is of course different to the non-informative cues used in the first two experiments in that it explicitly indicates two incorrect locations and provides spatial information instead of indicating a location essentially at random as the 0.25 reliable cue did, which provides no spatial information.

Figure 5. The event structure of a trial in experiment 3. Participants maintained fixation for 1.0s An endogenous central cue was then presented. This was either a partially valid single location cue, a two location cue, a two location anti cue or a neutral cue. 1.0s after cue onset, four RDKs appeared. One of the RDKs contained partially coherent motion in one of the four cardinal directions. A fixation marker was presented until the motion direction was reported by a key-press.



Information theoretic measures of spatial uncertainty

To quantify the value of the cues in decreasing the participants' spatial uncertainty, we used information theory (Shannon, 1948). The entropy, or initial uncertainty, H(L), is the expected value of the logarithm of the probability of the target, p(L), appearing at each of the four locations, where the discrete random variable, L, is the location of coherent motion.

$$\mathrm{H}(L) = -\sum_{i=1}^{4} \mathrm{p}(l_i)\cdot\log_2 \mathrm{p}(l_i) \tag{1}$$

The initial spatial uncertainty is 2.0 bits. The reduction in spatial uncertainty, namely the information transmitted (IT) by a partially valid cue is a function of the cue reliability *rel*, and is given by:

$$\mathrm{IT}(rel) = \mathrm{H}(L) + \left[rel \cdot \log_2 rel + (1 - rel) \cdot \log_2 \frac{1 - rel}{3} \right] \tag{2}$$

It is easily shown that the two location cues transmit 1.0 bits of information in that with four possible target locations there is an initial uncertainty of 2 bits and indicating two locations, one of which must contain the target effectively reduces the possible number of locations to two, resulting in 1 bit of uncertainty. A probabilistic one location cue transmits 1.0 bit at a reliability of approximately 0.81, as can be derived from formula (2).

Table 2 illustrates the uncertainty associated with the number of target locations and how that uncertainty is reduced by the information provided by various cue types in chapter 3. Of particular note is the information match between single and multiple location cues. (experiment 3 highlighted in grey)

Cue Type	**Possible Target Locations**	**Initial Uncertainty (bits)**	**Information Transmitted (bits)**	**Uncertainty remaining (bits)**
1 location 0.81	4	2	1	1
2 location	4	2	1	1
1 location 0.84	6	2.5850	1.5850	1
2 location	6	2.5850	1.5850	1
1 location 0.7	6	2.5850	1	1.5850
3 location	6	2.5850	1	1.5850
1 location 1.0	6	2.5850	2.5850	0

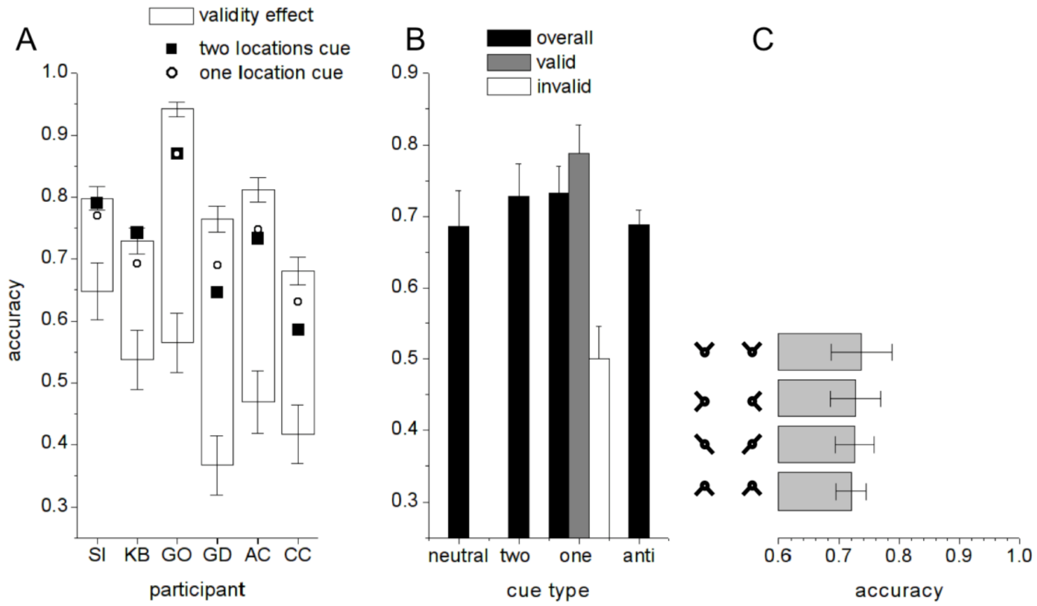
Table 2. Information theoretic quantities of cueing conditions in four location and six location stimulus set up. Illustrating cue type in terms of locations cued and probability of indicating target location, initial uncertainty in a stimulus with the amount of information transmitted by the cue in information theoretic terms (bits) and the remaining spatial uncertainty in the stimulus (bits)

Results and discussion

We compared discrimination accuracies when neutral, one or two location cues preceded the target display. The one location and two location cues were matched for the amount of spatial information they provided, but clearly differed in the target spatial probability distribution.

Figure 6A shows individual participants' accuracies and Figure 6B the group averaged accuracies. Interestingly, the one and two location cues were associated with virtually identical overall accuracies ($t(5) = -0.28$, $p = .79$), the difference being less than 5% in all participants. Nevertheless, the spatial information provided by both cue types was used, since accuracy was significantly greater cueing one or two locations than on neutrally cued trials ($t(5) = 2.61$, $p < .05$ and $t(5) = 3.50$, $p < .05$, respectively). Moreover, there was a significant validity effect for one location cues ($t(5) = 2.73$, $p < .005$). The multiple location anti-cue showed reduced accuracy compared to the multiple location pro cue though they provided the same amount of spatial information ($t(5) = 2.81$, $p < 0.05$). The detriment in the multiple location anti-cue condition reduced performance to such a degree that it was not significantly greater than neutral performance ($t(5) = 0.28$, $p =.79$). We then examined the effect of the cue spatial configuration on performance in trials preceded by the two-location cue. The cue configurations are shown in figure 6C, ordered by the associated level of performance. A Friedman's test indicated that configuration had no significant effect on accuracy (χ^2 (5) = 1.43, *p* =.92). In other words, cueing contiguous locations or locations in opposite quadrants did not seem to affect performance.

Figure 6. (A) Motion discrimination accuracy by participants. Bars indicate the size of the validity effect. The upper and lower edges are the accuracies on valid and invalid trials respectively for the 0.81 reliable cue, the circles are the overall accuracies, whereas black squares are the accuracies when the two location cue was used. Error bars represent bootstrapped estimates of standard error. (B) Motion discrimination accuracy by cue type. Grey bars indicate overall accuracies, the empty bar represents accuracy on valid trials and the filled bar on invalid trials. (C) Cue configuration and discrimination accuracies. Accuracies for different configurations of the two location cues are displayed in order of decreasing discrimination. Data are averaged over cue configurations identical up to a reflection across the vertical meridian. Error bars represent standard error of the mean (SEM) across participants



Experiment 4- six location information matched cues

Given that we observed similar performance for information matched cues in Experiment 3 we decided to increase the initial uncertainty and complexity of the stimulus by increasing the number of RDK stimuli that could be target locations to six. This served to extend on our initial findings by implementing a task which would test the limitations of dividing attention to multiple locations in different distributions. We also use two different levels of information (1 bit and 1.585 bits) each with a multiple location and information matched probabilistic single location cue. Furthermore we investigated whether the proximity of cued locations is responsible for the performance deficits that we observed when cueing multiple locations. This analysis tested the assumption that closely clustered locations would be easier to attend in accordance with a zoom lens account of attentional distribution.

Participants

Nine naïve participants underwent six sessions of 600 trials each on consecutive days.

Methods

All RDK and stimulus properties were identical to experiment 3 except that we now used six RDKs at an eccentricity of 15.9° at 60 ° increments along the radius of the circular array. One, two and three location cues preceded displays containing six RDKs. One of the RDKs contained coherently moving dots. The reliability of the one location cues varied over three levels; 0.70, 0.84 and 1.0. Both the 0.84 reliable cue and the two location cue transmitted approximately 1.59 bits of information, while the 0.70 reliable cue and the three location cue transmitted 1.0 bit of information.

The cue colour indicated the reliability of one location cues. Participants learned the relation between cue colour and cue reliability during a training pre-session. Data for the pre-session were not included in the final analysis. A neutral cue, which provided no spatial information was also used, and consisted of a colour change of the fixation dot lasting 1.0s.

Again the information of probabilistic cues and multiple location cues was matched as in Experiment 3. In this instance two levels of information matched cue were used, with 1 bit of information being transmitted by a three location cue when there are six possible target locations to begin with, leaving 1.585 bits of uncertainty remaining. The amount of in information transmitted by a three location cue is matched by a 70% reliable cue (see Equation 2 for quantification of probabilistic cue information). A two-location cue transmits 1.585 bits of information leaving 1 bit of uncertainty remaining; this amount of information is also transmitted by an 84% reliable cue. Table 2 illustrates the uncertainty associated with the number of target locations and how that uncertainty is reduced by the information provided by various cue types. Of particular note is the information match between single and multiple location cues.

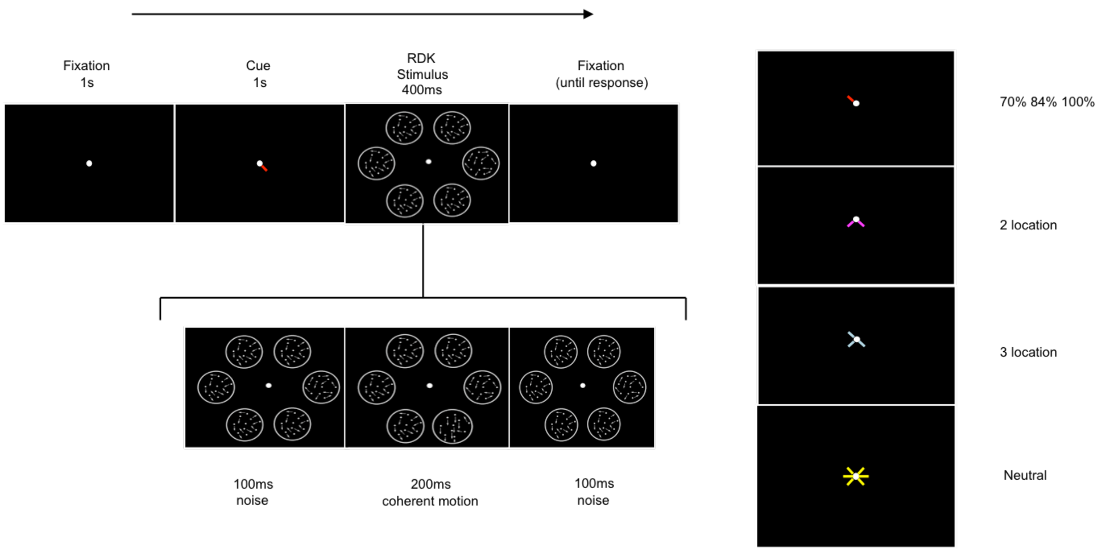
(Experiment 4 highlighted in grey)

Cue Type	**Possible Target Locations**	**Initial Uncertainty (bits)**	**Information Transmitted (bits)**	**Uncertainty remaining (bits)**
1 location 0.81	4	2	1	1
2 location	4	2	1	1
1 location 0.84	6	2.5850	1.5850	1
2 location	6	2.5850	1.5850	1
1 location 0.7	6	2.5850	1	1.5850
3 location	6	2.5850	1	1.5850
1 location 1.0	6	2.5850	2.5850	0

Table 2. Information theoretic quantities of cueing conditions in four locations and six location stimulus set up. Illustrating cue type in terms of locations cued and probability of indicating target location, initial uncertainty in a stimulus with the amount of information transmitted by the cue in information theoretic terms (bits) and the remaining spatial uncertainty in the stimulus (bits).

The order in which trials containing the various cue types was randomized within blocks. Figure 7 shows the trial set up and cue types used in experiment 4.

Figure 7. Trial structure in experiment 4. Six cue types were used. These included a 0.7, 0.84, or 1.0 reliable, one location cues, a two location cue, a three location cue and a neutral cue. The target displays contained six RDKs.

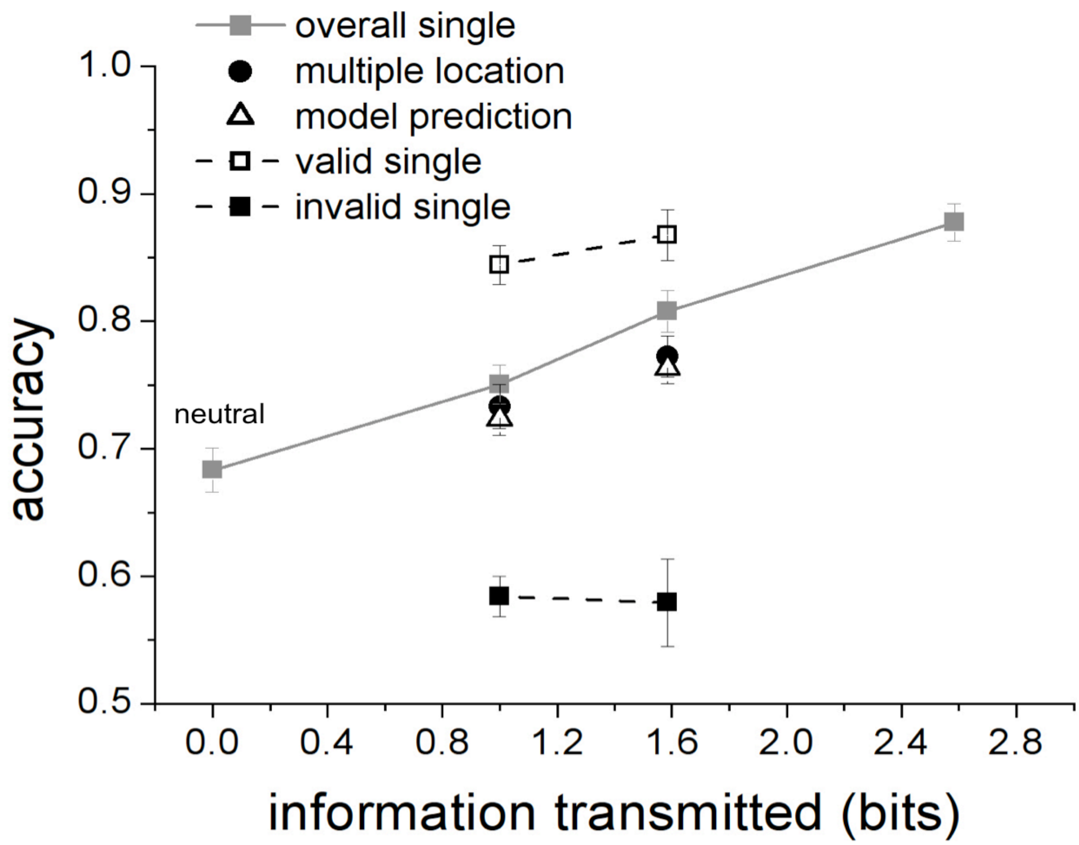


Six of the participants undertook additional testing, in which trials containing either two or three location cues were used in separate sessions. These data were used to examine the effect of cue configuration.

Results

Figure 8 shows the group averaged overall accuracy for probabilistic cues as well as the accuracies on valid and invalid trials, as a function of cue reliability. Discrimination accuracies for trials preceded by the two and three location cues are included and plotted as a function of the one location cue's reliability, which matched the information transmitted.

Figure 8. Motion discrimination accuracy as a function of information transmitted. Grey squares are overall, group accuracies following one location and neutral cues. Empty and filled squares are accuracies on valid and invalid trials, respectively. Circles are accuracies following three and two location cues. Triangles represent estimated performance when memory imprecision is accounted for as estimated from the recall data obtained in experiment 4, we estimated the effective amount of spatial information made available by two and three locations cues. Triangles are the accuracies predicted if one location cues had been used whose reliabilities matched the effective amount of information transmitted by the multiple locations cues given the memory detriments associated with them (see experiment 4). These data were obtained in six and seven participants who also completed the recall experiment (the difference in group average accuracies between nine participants and the participants who completed the recall experiment being less than 1.0% in all cases). Error bars represent standard error of the mean (SEM) across participants.



First, we examined the effects of cue validity and cue reliability using a 2 X 2 repeated measures ANOVA for cue reliabilities of 0.70 and 0.84. There was a significant main effect of cue validity ($F(1,8) = 110.13$, $p < .001$), however there was neither a significant effect of cue reliability ($F(1,8) = 0.71$, $p = .42$) nor a significant interaction of cue reliability by cue validity ($F(1,8) = 1.18$, $p = .31$), suggesting that changes in cue reliability between 0.7 and 0.84 had minor effects on cue utilization. However, there was a significant difference between discrimination accuracies in valid trials following the fully reliable and the 0.84 reliable cue ($t(8) = -2.78$, $p < .05$), suggesting that cue reliability does affect cue utilization.

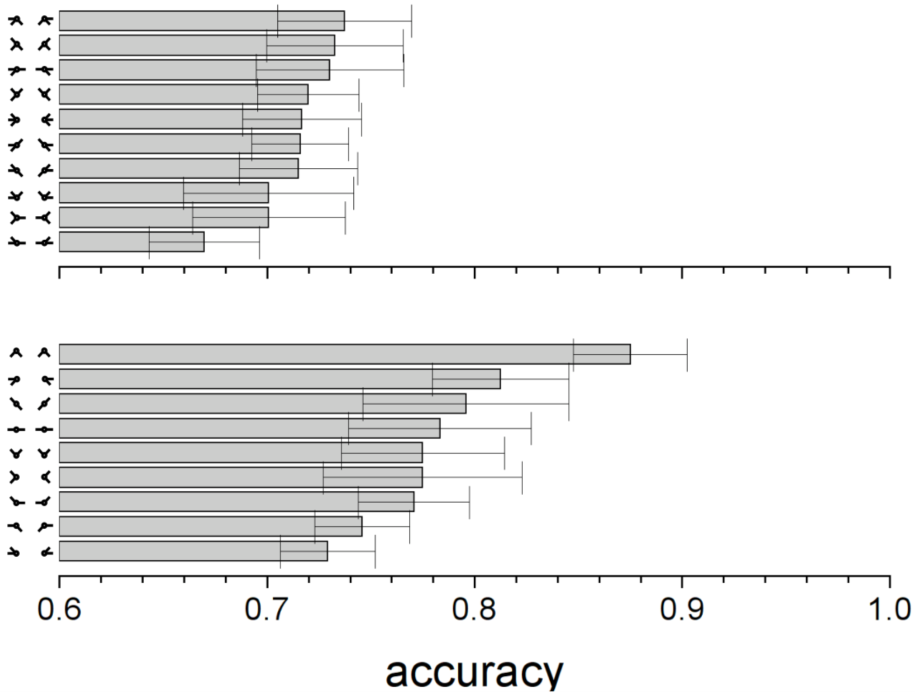
Overall discrimination accuracies on trials preceded by the 0.7 reliable cues and the three locations cues were both significantly greater than accuracies on neutrally cued trials ($t(8) = -5.65, p < .001$) and ($t(8) = -6.19, p < .001$) indicating that participants used the spatial information provided by both cue types.

Accuracies on trials in which three locations were cued were significantly lower than on trials where only one location was cued ($t(8) = 2.50, p < .05$), despite the fact that both cue types carried the same amount of spatial information. Similarly, following two location cues, overall accuracy was significantly lower than when a 0.84 reliable cue was used ($t(8) = 2.70, p < .05$). These data suggest that dividing attention between multiple locations is less effective than attending one location in displays containing six RDKs.

A Friedman's test showed no significant effect of two location cue configurations (χ^2 (14) = 11.57, $p = 0.64$), and a near significant effect of three location cue configurations (χ^2 (19) = 28.97, $p = 0.07$). When accuracies were averaged over configurations that were mirror symmetric across the vertical meridian, a significant effect of cue configuration was found for the three location cues (χ^2 (9) = 18.86, $p =$ 0.03). When accuracies were averaged over cue configurations that were mirror symmetric across the horizontal meridian, configuration of three location cues had no effect (χ^2 (11) = 6.33, $p = 0.85$). Figure 9 shows the accuracies associated with each cue configuration, where accuracies are averaged over cue configurations mirror symmetric across the vertical meridian. The accuracy rankings of cue configuration suggest that the main factor influencing performance is the number of locations cued in the upper vs. lower visual field, performance being higher when locations were cued in the lower visual field.

Figure 9. Cue configurations and performance. (A) Accuracies are displayed for three locations cue configurations and (B) two locations cue configurations.

Configurations grouped by horizontally equivalent configuration. Error bars represent standard error of the mean (SEM) across participants.



Next, we examined the effect of the target location on performance. A two factor ANOVA showed no significant effect of target position in the left and right hemifield ($F(1, 6) = 0.26$, $p = .63$) or its vertical location ($F(2, 12) = 0.59$, $p = .57$). A near significant interaction was found between hemifield and vertical position ($F(2,12) = 2.90$, $p = .094$).

Experiment 5. Attention and recall of multiple location cues

There are two possible explanations for the performance decrements observed when cueing two or three locations compared to one. The first is that dividing attention carries a cost over maintaining an undivided attentional focus. However, this account is difficult to reconcile with the finding that there was no difference in performance when one or two location cues were used and the target display contained only four RDKs. Alternatively, the worse discrimination performance could be accounted for by the increased memory load associated with attending two or three locations. That is, cueing multiple locations may have breached the capacity to recall the locations cued leading participants to either confound cued with uncued locations or ignore the cues. An objective measure of memory load, M, can be easily computed using information theory. The memory load, in bits, is the number of binary values required to uniquely specify the cued location(s):

$$\mathrm{M}(c) = \log_2 \frac{n!}{(n-c)!\,c!} \tag{12}$$

where n is the number of RDKs and c is the number of cued locations.

The table shows the IT and memory load, expressed in bits, for each cue type and demonstrates that memory load increases with the number of locations cued, resulting in different memory load for cues that provide identical spatial IT.

Cue Type	**Information Transmitted**	**Load**
1 location 0.70	1	2.5850
1 location 0.84	1.5850	2.5850
1 location 1.0	2.5850	2.5850
2 locations cue	1.5850	3.9069
3 locations cue	1	4.3219

Table 3. Information theoretic values for the information transmitted in determining probable target location and corresponding information required to maintain that representation as a memory load for various cue types. The greater number of locations cued corresponds to greater memory load but also less information transmitted.

To test the hypothesis that inaccurate recall of cued locations limits attentional benefits of multiple location cues, we asked participants to recall the cued locations in two separate experimental conditions. In the first, participants reported in separate blocks the cued location or the target motion direction. In the second, participants were required to report either the cued locations or the target motion direction in the same block.

Participants

Seven naïve participants were tested in this task. All the participants had also taken part in Experiment 4.

Methods

Two trial types were used: recall and discrimination trials. In recall trials, a spatial cue lasting 1.0s was presented. After the cue offset, six circular frames appeared, outlining the boundaries of the RDKs. The participants placed the cursor over the apertures corresponding to the cued locations. Discrimination trials were identical to those in the previous experiment.

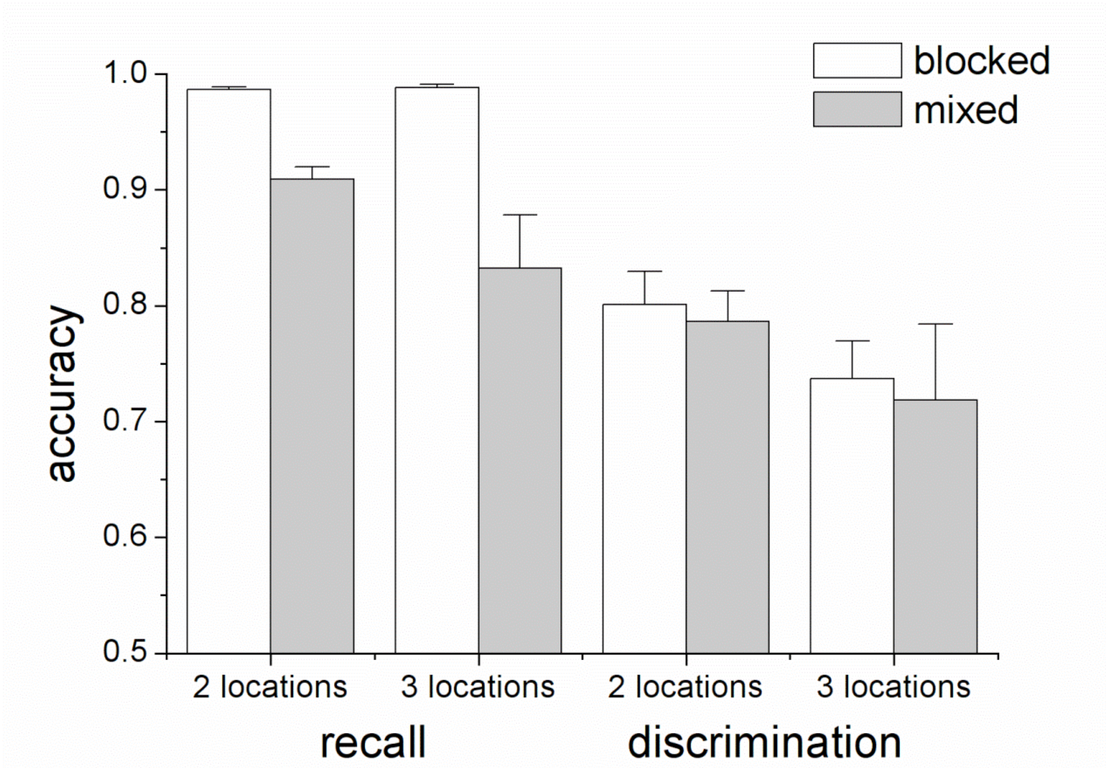
Participants performed the two tasks either in pure or mixed blocks. In the mixed blocks, participants were not aware of the trial type until either the motion or memory response display appeared. Trials preceded by two and three location cues were run in separate sessions. Recall with one location cues was not tested since preliminary testing sessions carried out in three of the participants showed performance to be at ceiling both in the pure (100% accuracy) and mixed conditions (>99% accuracy).

Results and discussion

Recall and discrimination accuracies are shown in figure 10. In blocks in which participants only had to recall the cued locations, performance was at ceiling, the average accuracies being 98.7%, recalling two and 98.8%, recalling three cued locations. However, in mixed blocks, recall was poorer, the accuracies being 90.9%, recalling two and 83.2%, recalling three locations. The effects of block type (pure vs. mixed) was significant ($t(5) = 45.50$, $p <.001$ and $t(6) = 23.56$, $p < .001$ for recalling two and three locations respectively).

In contrast, motion discrimination accuracy was not significantly different between pure blocks and mixed blocks whether two, ($t(6) = -0.33$, $p = .75$), or three locations were cued, ($t(6) = 0.23$, $p = .83$).

Figure 10. Cued locations recall and motion discrimination accuracies following two and three locations cues. The empty bars are the group averaged, recall and discrimination accuracies in blocked recall and discrimination trials. Grey bars are the accuracies when discrimination and recall trials were interleaved in random order. Error bars represent standard error of the mean (SEM) across participants.



A question of some interest is whether faulty recall of the cued locations could account for the discrimination accuracy decrements observed when two or three locations were cued. To get some insight into the issue, recall accuracies in mixed block trials were used to calculate an adjusted measure of the information transmitted by the cues, which accounted for the information loss due to memory imprecision. The following formula was used to compute the spatial information (IT2 for 2 location cues and IT3 for 3 location cues), in bits, provided about the actual target location when cueing two or three out of six possible target locations. Where P_x is the probability that a participant correctly recalled x cued locations.

$$\mathrm{IT2} = \log_2 6 + \left(p_2 \cdot \log_2 \frac{p_2}{2} + \frac{1}{2} p_1 \cdot \log_2 \frac{p_1}{4} + \frac{1}{2} p_1 \cdot \log_2 \frac{p_1}{8} + p_0 \cdot \log_2 \frac{p_0}{4} \right)$$

$$\mathrm{IT3} = \log_2 6 + \left(p_3 \cdot \log_2 \frac{p_3}{3} + \frac{2}{3} p_2 \cdot \log_2 \frac{2p_2}{9} + \frac{1}{3} p_2 \cdot \log_2 \frac{p_2}{9} + \frac{1}{3} p_1 \cdot \log_2 \frac{p_1}{9} + \frac{2}{3} p_1 \cdot \log_2 \frac{2p_1}{9} + p_i \cdot \log_2 \frac{p_0}{3} \right) \tag{13}$$

The recall adjusted IT associated with two and three locations cues was estimated for each participant and used to compute the reliability required for a one location probabilistic cue to transmit the same amount of information. Finally, the discrimination accuracy, associated with a one location cue whose reliability level matched the recall adjusted IT, was estimated by linear interpolation. The group averaged estimates are shown in figure 8. The recall adjusted discrimination accuracies were not different from those observed i*n* experiment *5* when two ($t(5) = -0.78$, $p = .47$) or three ($t(6) = 1.91$, $p = .10$) location cues were used. We conclude that imprecise recall of the cued locations accounts for accuracy differences when cueing one and multiple locations.

Discussion

Some authors argued that only one contiguous region of space can be attended at any given time (e.g., Posner, Snyder, & Davidson, 1980; Jonides, 1983; Eriksen & St. James ,1986), while others have suggested that human observers can split attention (Shaw & Shaw, 1977; Kramer & Hahn, 1995; Bichot, Cave, & Pashler, 1999; Awh & Pashler, 2000; Gobel et al., 2004).

We examined performance in a cued motion discrimination task in which the target display was preceded by cues for either one or multiple locations. The goal was to determine whether the number of locations cued modulated the effectiveness of spatial cueing.

Cueing the location of the coherently moving RDK improved discrimination accuracy. Moreover, when the display contained four RDKs, accuracy was identical when targets were preceded by one and two location cues. When stimulus contained six RDKs instead, cues indicating two and three locations were not as effective in improving accuracy as probabilistic cues indicating a single location, even though the cues were matched for the amount of spatial information provided. The finding that cueing one or two locations was equally effective when four RDKs were used, is inconsistent both with the view that a single attentional focus is simply adjusted to include all cued locations and that attention cannot be divided.

How should one account then for the lower effectiveness of multiple cues when six RDKs were displayed? One possibility is that the costs of dividing attention become apparent only when greater demands on attentional resolution are placed either by more tightly spaced stimuli (He et al., 1997; Cavanagh, Hunt, Afraz, & Rolfs, 2010; Franconeri, Jonathan, & Scimeca, 2010), or when attentional resolution is traded off against the number of attended locations (Alvarez, & Franconeri, 2007). The other is that memory demands are increased when displays contained six compared to four RDKs, leading participants to either forget the cued locations or confuse cued and uncued locations. This hypothesis is not, *prima facie*, in keeping with estimates of visual working memory capacity. These have indicated that humans can recall accurately three or four objects (Luck, & Vogel, 1997; Zhang, & Luck, 2008), suggesting that recalling two or three locations is within the capacity limits of working memory. However, using cues may require more than just remembering the cued locations. It has been argued that cueing engages processes that shift attention to the cued location and keep it there (Posner et al., 1980; Posner & Petersen, 1990). We examined the possibility that such processes, by themselves, tap capacity limited resources shared with working memory, by asking participants to recall the cued locations, under two different conditions. In the blocked condition, participants had to perform the recall task on each trial; in the mixed condition, cued locations recall trials and motion discrimination trials were randomly interleaved. Recall accuracy was at ceiling in the blocked trials, but it dropped significantly when participants could not anticipate whether they would have to recall the cued location or perform the discrimination task, and thus had to shift attention on every trial. The results suggest that shifting to and maintaining attention at the cued locations requires resources shared with working memory. Moreover, motion discrimination was not different between blocks in which participants only discriminated motion direction and blocks in which memory and discrimination trials were mixed, indicating that the demands placed on attention by adding the possibility of the recall task were no greater than those placed by the cued motion discrimination task. Finally, since in recall trials the motion display did not appear, the performance decrements observed in mixed blocks obviously cannot be attributed to increased demands posed by performing a dual task. Rather, the difference should be accounted for in terms of the preparatory processes engaged in mixed blocks, but not in recall only blocks, namely those used to attend the cued locations.

We calculated the amount of spatial information lost because of memory imprecision, to predict the discrimination accuracy that should have been observed with one location probabilistic cues, matched for the amount of spatial uncertainty. Once the effect of recall imprecision was taken into account, motion discrimination accuracy when cueing multiple locations was no worse than that predicted if an equally informative, one location cue had been used. This result shows that recall imprecision could account for the accuracy differences when cueing one vs. multiple locations.

Our findings are consistent with the widely held belief that visual working memory and attention are intimately related processes (Downing, 2000; Theeuwes, Kramer, & Irwin, 2011; Chun 2011). Some have suggested that working memory is dependent on attention to operate. For example, if attention is engaged at locations other than those stored in memory then recall is impaired (Awh, Jonides, & Reuter-Lorenz, 1998; Smyth & Scholey, 1994; Lawrence, Myerson, & Abrams, 2004). Crucially, when no spatial shift of attention is required, perceptually demanding tasks do not interfere with spatial memory (Awh, Jonides, & Reuter-Lorenz, 1998), implying that interference between attention and working memory arises when they operate on a shared spatial representation. Some authors have concluded that maintaining visual information in working memory requires continuously attending internal representations (Smyth & Scholey, 1994, Chun et al., 2011) or that maintaining a location in working memory is equivalent to attending that location (Awh, & Jonides, 2001). However, neither this nor the opposite view, as formulated by Fougnie and Marois (2006), who concluded that capacity limits for attention and working memory are largely independent (see however Zhang et al., 2010), can account for the finding that the recall of cued locations was worse when participants had also to attend those locations. Our findings suggest instead that endogenous attention requires both access to limited capacity resources shared with working memory and maintenance of a representation of the attended locations in working memory. On the other hand, working memory does not require shifts of attention to the maintained locations.

In addition we find that when a two-location cue indicates explicitly incorrect locations performance is depreciated. This leads to performance that is not different from neutral cues providing no spatial information. If people attended these incorrect locations then they should find lower performance than neutral cues. However the depreciation in performance suggests the directions indicated by cue stimuli influenced attention independent of spatial probabilities as the anti-cue was visually identical to the two-location cue except for color, the only difference between these conditions was the locations that were physically indicated by the directional stimulus. The most likely explanation is that there is some automated attentional influence of directional stimulus and that this requires inhibitory processes be used to suppress the influence to attend these locations resulting in reduced performance. This does not likely relate to a distortion in probability or a probability matching strategy as the non-informative cues in experiments 1 and 2 may have. The spatial distribution of probability was identical to the multiple location pro cue and as the performance seems quite effective with the pro cue, the anti cue deficit does not likely relate to incorrect use of probabilistic information.

Chapter 4

Learning of spatial cue reliability

Are spatial expectations generated by attentional cues shaped by experience? We examined the issue in a motion discrimination task where cues indicating the likely location of coherent motion in displays, which contained four random dots kinematograms (RDKs). The probability that the cue would indicate the location of the coherent RDK was varied within blocks over four levels. Effects of cue reliability were observed only in two of the five participants in both valid and invalid trials. In those two participants, changes in cue utilization, that followed changes in cue reliability, took more than four trials to become established, suggesting that the effects of learning, when present, either cumulate over trials or depend on evidence available from sequences whose length is greater than four trials.

The central processing of sensory stimuli, which may require either a prompt response or a difficult visual judgement, is affected by the availability of prior probabilistic information. Posner (1980) examined the effect of centrally presented cues, which indicated where a highly visible target was likely to appear, in a simple reaction time task. When the target appeared at the cued location, participants were faster reporting its presence than when it appeared at the opposite, uncued location, leading to the conclusion that preparatory processes, engaged by spatial expectancies, modulate the timing of motor responses. A number of observations in both healthy and brain injured patients lead to the suggestion that the effect of spatial cues on the timing of a motor response reflect the sequence of distinct mental operations that need to be carried out in order for a behavioural response to be executed (Posner and Petersen, 1991). An important antecedent to this work was the seminal observation that laboratory animals show an overwhelming, automatic orienting response to unexpected, novel events (Pavlov, 1925). One proposal regarding the development of endogenous attention, based on a behaviourist viewpoint, is that orienting can be initiated by unconditioned cues, which foreshadow the appearance of the conditioned stimulus (Sokolov, 1966). If endogenous attention is a conditioned response then it must be characterized by the ability to internalize, though learning, the relation between unconditioned and conditioned stimuli.

Currently, the evidence that human subjects can learn where to direct their attention given environmental cues is quite limited. Carpenter and Williams (1995) examined the effect of spatial expectancies on saccadic latencies, in a task in which participants had to quickly fixate a visual target presented either on the left or right of the fixation point. The probability that the visual target would appear on either side was systematically varied over extensive testing sessions. The learnt spatial distribution of the saccadic target was found to systematically affect saccadic latencies. When these spatial distributions change under the same conditions, it was found that people adapt to these changes in spatial expectancies (Anderson and Carpenter, 2006). In addition contextual cueing has show that underlying spatial expectancies can bias perceptual behaviour and facilitate search based on the arrangement of distractor stimuli as a visual cue, when the target location is static (Chun and Jiang, 1998). The benefits these authors observed occur without the ability to recall explicitly which visual displays are indicative of target location from displays that were not of any utility, suggesting the process is implicitly learned. Miller (1988) also found that targets could be better identified based on position relative to other stimuli when the actual retinal location changed. Therefore it is possible that endogenous cues, which indicate a spatial probability that is non static, may also be effectively learned without prior knowledge. Though when participants had to detect and localize a visual target presented within one of several locations highlighted by annular placeholders, participants' performance was not found to be influenced greatly by the probability of the target appearing at a given location, as signalled by a placeholder's colour when feedback was not provided (Droll, Abbey and Eckstein, 2009). This was true even though the participants were able to verbally report how likely placeholders of various colours were to contain the target, at the end of a behavioural session where they were exposed to repeated pairings of placeholders' colour and targets' location. Thus, one can conclude that it is possible to estimate accurately the spatial probability distribution of a target on the basis of learnt association with preceding visual cues. However, either the effects of cues on behaviour requires more than the knowledge of the target probability, or the ability to report verbally the reliability of the cued information, on one hand, and use that information procedurally, on the other, require access to different explicit and implicit representations of the cue reliability. However, this latter explanation is at odds with results obtained in detection and discrimination tasks, where the level of reliability of a spatial cue communicated verbally to the participants, prior to testing, was found to affect participants performance.

Another issue is that of whether learning effects are governed by transient mechanisms relating to only the most recent trial activity or whether a representation of cue reliability is attained. For example it has been suggested that repetition of targets occurring in high probability locations accounts for apparent spatial biases (Walthew and Gichrist, 2006). Counter to this evidence static spatial biases have been shown to be present when a broad distribution of target locations and almost no location repetition occurs (Druker and Anderson 2010).

Given that there is still sparse evidence regarding whether and how the reliability of attentional cues is learnt, and its consequences on performance, we examined how changing the reliability of a visual cue, which indicated the likely location of a target RDK containing coherent motion, affected report accuracy, in a motion direction identification task. The reliability of the attentional cues was varied within trial blocks and the effect of cue reliability on performance, in valid and invalid trials, was used as a behavioural proxy for changes in the internal representation of cue reliability.

Experiment 6 – Learning cue reliability

Methods

Participants

Six naïve, healthy participants took part in the experiment. All gave written consent. Five participants completed four training sessions of 600 trials each while a participant completed only one training session. All participant complete four daily experimental sessions of 480 trials each.

Apparatus

Stimuli were generated on an Apple Mac Pro 1.1, using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) for Matlab© (Mathworks, 2008a, Natick, MA). Stimuli were presented on the screen of a LaCie Electron 22blue CRT monitor with a refresh rate of 60Hz, placed at a distance of 70cm from the participant. Head position was restrained by a chinrest.

Procedure

The trial structure is illustrated in figure 1. At the trial onset, a fixation point was presented for 1.0s. A single coloured line, abutting the centre of the screen was then presented for 1.0s, which indicated the likely target location. The cue was followed by four RDKs, one in each visual quadrant, lasting 600ms. The RDKs contained 100 dots with a 0.15° diameter, in positive contrast. The dots' speed was 6.17°/s. Dots' lifetime was two frames. The RDKs covered circular apertures of 7.26° in diameter, centred at an eccentricity of 7.26°. The target RDK contained a 400ms period of coherent motion in one of thirty-six possible directions over a 360° range. This coherent motion was preceded and followed by two 100ms periods of incoherent motion.

Immediately after the motion stimulus a response display containing thirty-six oriented red lines was shown. To indicate the direction of motion participants moved a cursor over the line with a matching orientation and clicked on the mouse. This prompted the start of the next trial.

In the main experimental session, partially valid directional cues were used. On valid trials the cue indicated the target location correctly. In the remainder of the trials, i.e. invalid trials, it indicated a location different from the one containing the coherent RDK. Cue reliability, namely the probability that the cue indicates the location of the coherent RDK, was varied over four levels 0.25, 0.58, 0.72 (baseline probability), 0.86 during the experimental sessions. The training sessions contained trials preceded by 0.72 probability cues, neutral cues, which consisted of the fixation point turning green for 1 second, and a fully valid 1.0 probability cue. Participants were informed of the reliability of the fully valid cue, but not the other cues.

Participants familiarized themselves with the task during four training sessions of 600 trials each. In the first three sessions, two cue reliabilities were used, 0.72 and 1.0, and the cue reliability were indicated by the colour of the cue (red or green). Participants were informed initially of which colour indicated the fully reliable cue and that the second colour was associated with an unstated, but constant reliability, which they were expected to learn. In the last training session, the cue reliability was 0.72. In 10% of trials of each session, the RDKs were preceded by neutral rather than spatially informative cues.

During the experimental sessions, participants were informed that the cue reliability could vary, but they were not told either when or how. In each session, the cue reliability was changed in a stepwise fashion five times. The initial and two subsequent periods contained either seventy-six, eighty or eighty-four trials, in which the cue reliability was 0.72. These were interleaved with eighty trials long periods in which the cue reliability was either 0.58 or 0.86. The cue reliability was devalued to 0.25 in the final eighty trials of each session.

The coherence of the target motion stimulus was determined in a 60 trials long pre-session, ran before every training and experimental session. In these trials, the RDKs were preceded by a neutral cue and feedback was provided visually, after the participants' response, by presenting the word 'right' or 'wrong'. Threshold motion coherence was determined using a 2 to 1 staircase procedure with a 2% coherence step. The coherence that yielded 70% discrimination accuracy was estimated by data interpolation and used in the experimental sessions.

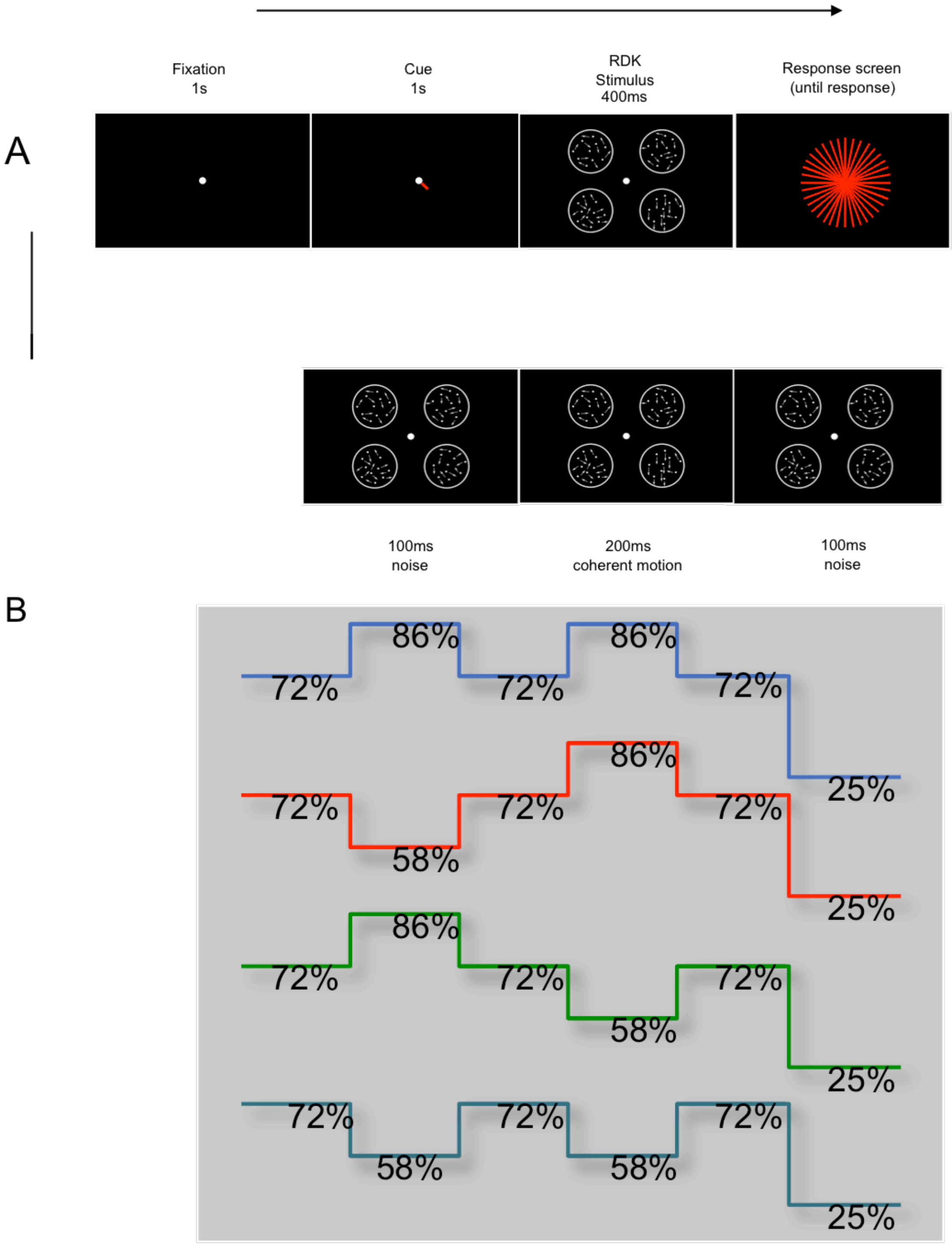


Figure 11. A. Trial structure for experiment 6. Four cue reliabilities were used, that were not known to participants. The cue varied between 0.25, 0.58, 0.72 and 0.86 reliability. There were four RDK stimuli one of which had coherent motion in one of 36 directions (10° increments around 360°). A response screen indicated the 36 possible directions of motion, participants indicated the direction of motion by

placing the cursor on one of these indicators and depressing the mouse button. B. displays the four structures of cue reliability trials that were used. The trials began at the baseline level and could vary to a higher or lower reliability at two intervals followed by a devaluation to 25% non informative reliability. The order in which these session structures were run was counterbalanced across participants.

Analysis of motion direction error size

The root mean squared error (RMSE) in reporting the motion direction was estimated using the following procedure: 1) the unsigned angular difference between the actual and reported motion direction was computed trial by trail. 2) One thousand bootstrap samples were generated by sampling with replacement the empirical error size distribution. 3) For each bootstrapped sample the 68th percentile value of the distribution was identified. 4) The RMSE, corresponding to the standard deviation of the error distribution, was obtained by averaging the bootstrap estimates. This procedure yielded a continuously valued RMSE, even though the trial-wise error size varied in increments of 10°. The relation between RMSE and cue reliability was estimated separately on valid and invalid trials and participant by participant. To this end a linear regression analysis was used, where the dependent variable was the RMSE and the independent variable the target log-likelihood:

$$RMSE_{valid} = c_0 + c_1 \cdot \log_2(r)$$
$$RMSE_{invalid} = c_0 + c_1 \cdot \log_2\left(\frac{1-r}{3}\right) \qquad (14)$$

where r is the cue reliability. The c_1 coefficients provided a numerical estimate of the relation between the size of the directional error and the target log-likelihood, whose units are degrees of the standard error per doubling of the target probability.

The group-wise statistical analysis was carried out using customary parametric tests.

Estimating the length of the learning sequence

Unsignalled changes in cue reliability were found to have systematic effects on the accuracy of motion direction estimates in two of the participants, suggesting that they adjusted their cue utilization to the learnt cue reliability. An important issue is how many trials the effects of cue validity are cumulated over, which eventually result in significant changes in cue utilization. In order to address this question, we estimated a lower bound on the length of the trial sequence sufficient to account for the behavioural effects of cue reliability. We generated synthetic sequences of directional reports by substituting, for each trial in the original data set, the directional error obtained by sampling, with replacement, distributions comprising directional reports from all the trials, which were of the same validity as the target trial and were preceded by one, two or three trials, whose validity matched the validity of the one, two or three trials preceding the target trial. To summarise, the synthetic data sets represented performance based on the most recent trial history, but was randomised with regard to the actual reliability block that was originally being responded to. Thus if perceptual changes are dependent on the most recent trial activity then the synthetic data should still reflect this factor and be similar to the empirical data. However if the synthetic data differ from the empirical data it suggests the modulation of performance that we do observe is not dependent on just the recent trial sequence but must take into account a larger history of trials. The RMSE was then computed for each of the four cue reliability levels and the effect of target log-likelihood on the RMSE estimated separately for valid and invalid trials. Ten thousand such estimates were then averaged to calculate the average effect of the target log-likelihood on the size of the RMSE, for valid and invalid trials. If the effects were smaller than those found in the empirical data set then we could conclude that learning takes place over trial sequences longer than the length of the sequence of two, three or four trials used to generate the synthetic data. Otherwise, we would conclude that learning does take place over short sequences of two, three or four trials.

Additional analyses not included

The intention of using a fine motion discrimination task was to allow for greater degrees of freedom in the metric used for performance. We filtered the data in a number of ways, first looking at trial by trial fluctuations in overall error with indications of validity for each trial. Next we used a moving average with windows of various sizes (5, 10, 15, 20, 25 and 30 trials) that took the mean of errors for trials in this window separately for valid and invalid considering all trials within the window advancing a trial or number of trials (1, 5 and 10) at a time. We also took means of each reliability block for valid invalid and overall error. These filtering and averaging techniques revealed no discernible modulations of performance and so were not given further consideration or illustrated in this chapter. The volatility of the spatial contingencies may have been too high to allow adaptations to be seen, especially given estimates of reliability must be derived from trial by trial updating based on a binary outcome (i.e. a valid or invalid cue on every trial).

Parametric statistical analyses were also conducted that provided no significant results, these included 2 way repeated measures ANOVAS on sections of each block of trials comparing the RMSE for the first, second and third portions of trials with factors of section, reliability and validity no significant effects were found even in individual ANOVAS, with validity and section as factors separately for each reliability condition.

Results

No group level evidence for learning of cue reliability

Figure 2 shows the relation between the size of the group average RMSE and the cue reliability, separately for valid and invalid trials. Valid trials showed no systematic effect of cue reliability on the size of the RMSE, however, on invalid trials, the RMSE appeared to increase with cue reliability. The significance of the effect of cue validity and reliability on the RMSE was assessed at the group level with a two-way repeated measures ANOVA. The within factors were cue validity, on two levels, valid and invalid, and cue reliability, which included four levels (0.25, 0.58, 0.72 and 0.86). The ANOVA showed a main effect of cue validity ($F [1, 4] = 10.654$, $p < 0.05$), but no significant effect of cue reliability ($F [3, 12] = 1.412$, $p =0.29$). The interaction of cue validity by cue reliability was also not significant ($F [3, 12]= 1.730$, $p = 0.21$). Thus, the data suggest that at the group level the effects of learning on cue utilization were either too small or too variable, among participants, to be significant at the group level.

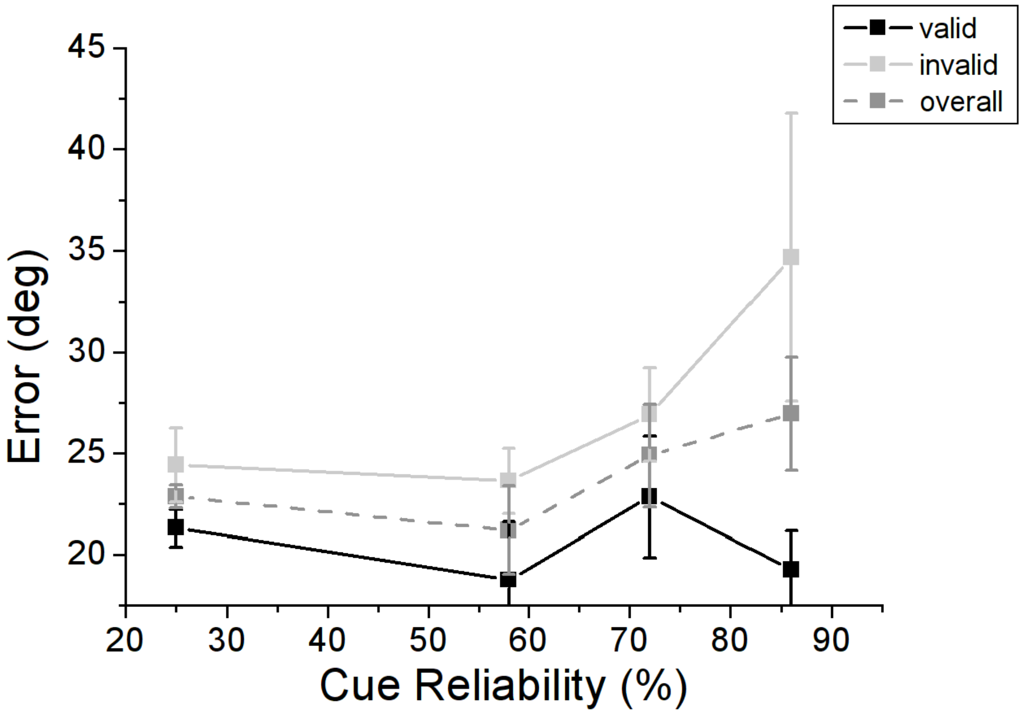


Figure 12. Group average RMSE as a function of cue reliability, the solid black line indicates performance on valid trials the solid light grey line indicates performance on invalid trials and the dashed grey line indicates overall performance collapsed across valid and invalid trials. Error bars indicate standard error of the mean (SEM)

Some participants learn cue reliability

To examine whether the lack of a significant effect of cue reliability on performance reflected effects that were consistently negligible across all subjects or rather concealed large inter-individual differences, we examined the relation between the magnitude of the effects of target log-likelihood on the RMSE on valid and invalid trials. If participants differ in their ability to learn then we expected these differences to encompass performance on both valid and invalid trials. Figure 3 shows the relation, participant by participant, between the size of the effects of target log-likelihood on the RMSE on valid versus invalid trials. A significant correlation existed between the values on valid and invalid trials ($r = 0.9$, $p<0.05$) suggesting that the lack of significant effects of cue reliability at the group level reflected effects of cue reliability that were consistent within participants, on valid and invalid trials, but which differed significantly between individuals. Only two of the participants did show in fact *bona fide* effects of target log-likelihood and, by extension, cue reliability on the size of the RMSE, while the remaining three participants did not

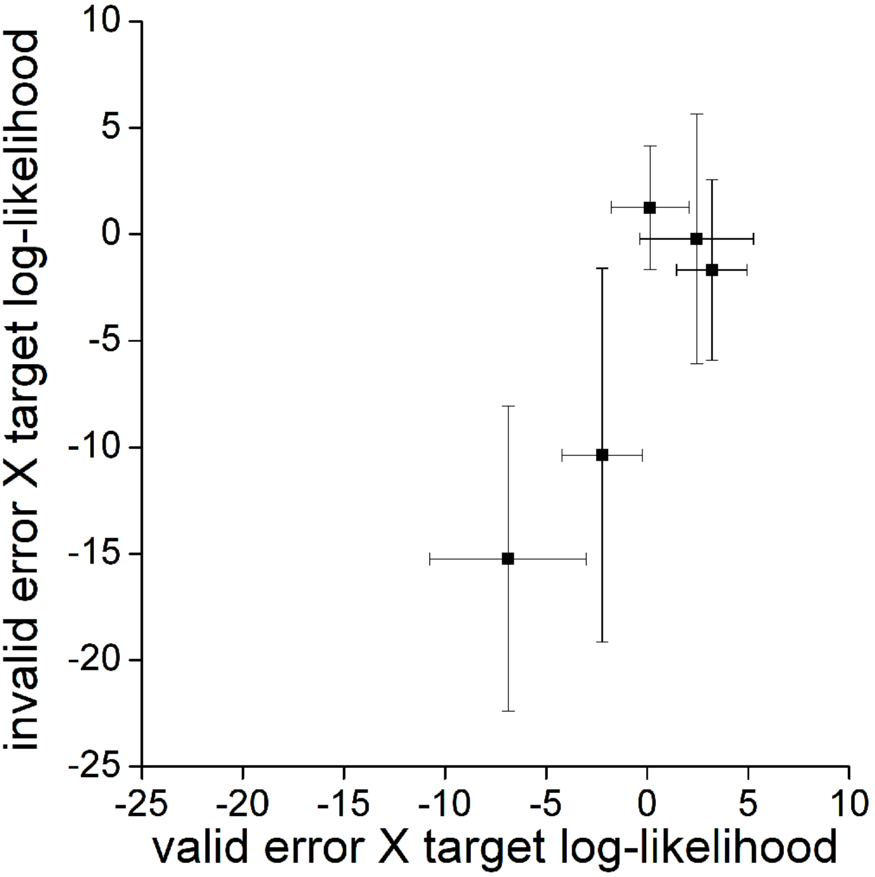


Figure 13. Effect sizes of the correlation between RMSE and target log-likelihood for valid and invalid trials by participant. Error bars indicate SEM for valid and invalid trials. Negative values indicate that a reduction in error and thus improved performance is correlated with the target log likelihood. Whereas a value of 0 indicates no correlation between performance and target log likelihood. Positive values indicate poorer performance is associated with target log likelihood. Two of the participants show performance modulations by target log likelihood, whereas the other three show no effect of target log likelihood on performance. Emphasising the inter-individual differences between participants.

show evidence of either having learnt the cue reliability or changed their performance accordingly.

Learning of cue reliability takes place over sequences longer than four trials

Finally, we examined in the two participants, whose performance was affected by changes in cue reliability, the extend to which learning effects could be accounted by changes in cue utilization that only spanned two, three or four trials, in order to establish whether the effects of learning cumulate over time or rather exhibit a "leaky integration". Figure 4A and 4B show, for the two participants, the target log-likelihood effects on the RMSE separately for valid and invalid trials. The effects were estimated from the empirical data and from synthetic data obtained under conditions in which the effects of learning were limited, by construction (see Methods), to trial sequences of length two, three and four. The results indicate that as the length of the sequence, over which learning took place, increased from two to four trials, so did the effect of the target log-likelihood on the RMSE. Nevertheless, in neither subject was the size of the target log-likelihood effect, when the sequence length was four trials, as large as that estimated from the empirical data set. If the modulations by reliability condition observed were dependent purely on the recent trial sequence, then the modulations of performance by the synthetic data (composed according to recent trial sequence) should not be different from the empirical data. At least for invalid trials, suggesting that the participants who learnt, cumulated evidence over sequences whose length was greater than four trials.

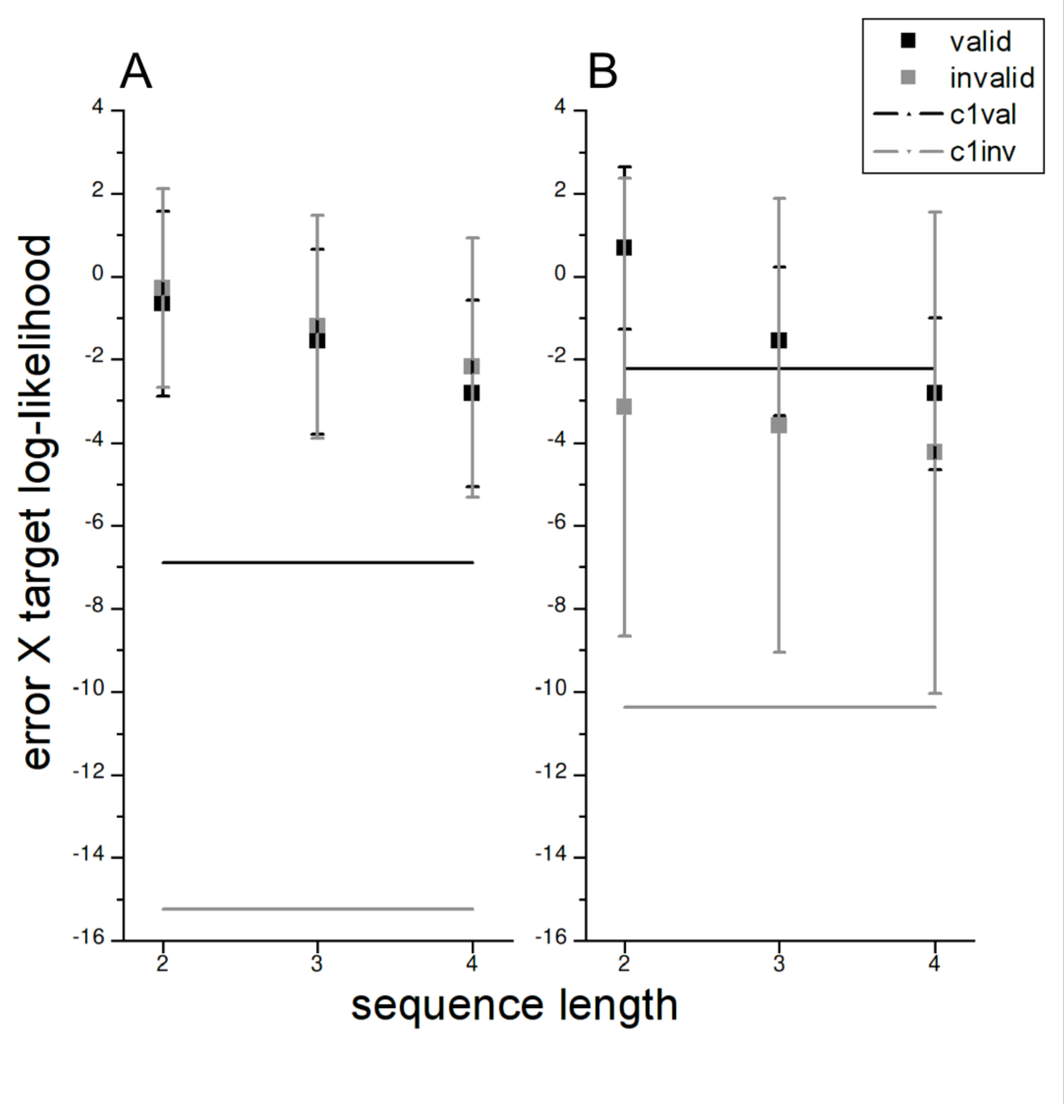


Figure 14. displays two participants' (A and B) effect sizes for of the correlation between RMSE and target log-likelihood for valid (black) and invalid (grey) trials for original data (lines) and for simulated data based on sequences considering validity of only two three or four trials (points). Error bars indicate SEM for valid and invalid trials. The difference between the synthetic data and actual data in terms of target log likelihood affecting performance suggest that more trials than the previous three need to be considered to elucidate the effects in the empirical data. Though as the number of trials considered increases so to do performance modulations by target log likelihood.

Discussion

We examined cueing effects on the accuracy of motion direction identification from displays, which contained one coherent and several non coherent RDKs and were preceded by probabilistic spatial cues. Within each testing session, the cues' reliability was changed at times unknown to the participants. We found that overall participants did not adjust their level of cue utilization on the basis of changing cue reliabilities. If participants had done so, we expected that the accuracy in identifying the motion direction would have improved with increasing cue reliability, on valid trials, and worsened, on invalid trials. However, we found no group level effect of cue reliability on motion direction identification accuracy, on both valid and invalid trials (see figure 12), nor a significant interaction of cue reliability with cue validity. In contrast to learning of spatial probabilities in contextual cueing paradigms showing underlying spatial expectancies can bias perceptual behaviour and facilitate search when the target location is static (Chun and Jiang, 1998) or when the location changes but the relative position can infer spatial expectancy (Miller, 1988). Though the lack of learning effects, at the group level, replicates previous findings. In a recent study (Droll Abbey and Eckstien, 2009), participants were asked to detect and localize a visual target that would appear, in 50% of the trials, in one of several locations, highlighted by placeholders of various colours. Each colour was associated with a particular probability that the placeholder would contain the target and this probability was unknown to the participants. It was found that when the participants were given no feedback, either about their performance, or about the presence and location of the target, they did not appear to make use of the colour information when detecting and localizing the target, despite being able to accurately report the relation between colour and target probability, following completion of the testing session. The authors concluded that cue utilization could not be learnt under conditions in which feedback about either performance or target was withheld. Since we did not collect subjective ratings of the cue reliability from our participants, we do not know whether they were simply unaware of the cue reliability or rather they did not use this information in performing the task. Regardless, our results confirm that learning the value of the cue does not happen under conditions in which no feedback about either performance or the visual target is given, even when participants are offered ample time to familiarize themselves with the task and the nature of the cue. Though it is important to note that our study reflects changes in both the location a reliability is associated with and the fidelity of that reliability that has become associated with the cued location. Changing the probabilities of unknown static spatial biases can be learned over a number of trials (Anderson and Carpenter, 2006). Though perhaps when the reliability of a cue is variable and not consistently associated with a static spatial location people are unable to adapt to this changing cue reliability.

Between participants differences in learning

While the data did not suggest that participants, as a group, adjusted their cue utilization following changes in cue reliability, we also found that some of the participants exhibited, on both valid and invalid trials, an effect of target log-likelihood on motion identification RMSE (see figure 13). The significance of the correlation, across subjects, between the size of the target log-likelihood effects on valid and invalid trials, confirmed that indeed inter-individual differences are likely to exist in the ability to learn the cue reliability or its utilization. Whether these inter-individual differences reflect differences in circuits associated with reinforcement learning (Schönberg et al. 2007), which may be vulnerable to genetic and environmental factors (Joel et al. 2005) or in the nature of the strategy underlying the use of probabilistic information (Stanovich 2003) are interesting, but at the moment, speculative inferences.

Learning takes place over more than four, subsequent trials

In two participants, who were able to adjust their performance to changes in cue reliability, we assessed whether learning could be accounted for by processes that included only information about the most recent sequence of two, three or four trials. We found that the size of the learning effects, in those participants who showed them, could not be accounted for if participants had only kept track of the cue validity in the previous one, two or three trials, suggesting that information about previous trials included sequences whose length was greater than 4 trials. This finding may suggest that learning of endogenous cue reliability depends on mechanisms whose temporal integration window is longer than that of mechanism used to learn the unconditioned spatial distribution of visual targets, the latter taking place over 5 trials or less (Walthew and Gilchrist, 2006; Rabbit, Cumming, Vyas, 1977; Rabbit, Cumming Vyas, 1979).

There are at least two potential explanations for the accumulation of information about cue reliability over trials. First, it is possible that the internal estimates of the cue reliability are incrementally adjusted on every trial given the previous trial cue validity. Then, the finding that sequences of greater than four trials are required in order to account for the effects of learning would not imply that participants hold in memory a distinct representation of each of the previous trials validity, but rather that the estimates of cue reliability are changed relatively slowly by ongoing updates. Alternatively, it is possible that a certain level of evidence regarding the possibility that cue reliability has changed, needs to be reached in order for the internal estimate of cue reliability to be updated. A number of algorithms have been formulated to adapt quickly to changes in the statistical dependencies between conditioned and unconditioned stimuli or outcomes (Yu and Dayan, 2005). Recent evidence shows that when informed the probability will not change for static probabilities, peoples saccadic behaviour reflects intentions to seek out dynamic and transitory environmental changes (Vincent, 2012), even over long periods of testing allowing for adaptation to the learned probability. Thus it may be that, humans are hard wired to assume that spatial expectations are dynamic as this is adaptive in naturalistic settings where static spatial probabilities are rare.

There are a number of changes that could be made to better answer the questions that we had in mind. Firstly a feedback condition may have allowed people to be better able to internalise and update a representation of cue validity or to utilise this representation given the assumption that people can learn these probabilistic associations but then fail to apply them to the task. A less volatile reliability block structure and/or longer periods of cue reliability may have been in order to allow adaptations to new reliabilities to take place. In addition given the large inter-individual differences that were observed, future studies of similar concern may benefit from seeking out learners who display modulations of perceptual behaviour and separating them from non-learners into separate groups. Likely allowing for more robust analysis of group level effects, especially given the small number of participants used in this study. A different approach in terms of cueing may have proved beneficial, as inferring modulations in performance from comparisons of valid and invalid trials (reliability effects) did not appear conducive to studying brief periods of adaptation. Perhaps using similar place holders to those of Droll Abbey and Eckstien (2009) could have allowed the relation between spatial location and probability to be inferred. In addition changing the instructions given to participants such that they are asked only to discern the relation between the cue and the target, giving less inclination about the volatility of cue reliability may have lead participants to perhaps trust and apply representations to alter their perceptual behaviour.

Chapter 5

GENERAL DISCUSSION

Cue reliability affects accuracy and reaction times in a similar manner

We investigated the effect of cue reliability on supra-threshold motion detection latencies and accuracy of threshold motion discrimination. There was an overall benefit in both response accuracies and latencies as the reliability of the cue increased. Moreover, in both tasks the difference in performance between valid and invalid trials was modulated by the reliability of the cue, in agreement with previous literature (Jonides, 1980; Giordano, McElree and Carrasco, 2009; Geng and Behrmann, 2005).

The finding that the validity effect increased with the cue reliability confirms that humans can adjust their cue utilization to reflect the value of the probabilistic information thus provided. The effects of cue reliability on performance appeared to be larger on invalid than valid trials, even though the difference was not significant.

Greater costs, in invalid trials, compared to benefits, in valid trials, is not a novel finding (Jonides 1980), and several hypotheses have been laid out that could explain it. Jonides (1980), having altered the probability that an endogenous cue would indicate where a target letter may appear among eight locations, all containing letters, found that valid cues improved performance and invalid cues decreased performance and that the size of this difference between cost and benefit increased with increasing cue reliability. The author also found that the size of costs (difference between invalid and neutral trials) increased more than benefits (difference between valid and neutral trials) with increasing cue reliability. The finding of greater RT costs following invalid cues than benefits following valid cues rules out the hypothesis that locations are searched serially until the target is found. since this model would predict greater effects of cue reliability on benefits than costs. Secondly, the finding that cue reliability modulates validity effects also rules out a two-process model based on a deterministic strategy, such as a maximization rule, since this would predict that the validity effects should vary with cue reliability in a step-wise, rather than gradual fashion. Therefore, either a deterministic strategy is not used or cues are not used to initialize a serial search strategy.

The modulation of the validity effect for discrimination accuracy by the reliability of probabilistic cues is similar to that found for response latencies. Cueing the location most likely to contain the target stimulus has been shown to improve RT performance (Posner, Nissen and Ogden 1978; Posner 1980), and the cue reliability can change the magnitude of the validity effect for RTs. Few studies have considered the effects of manipulating cue probability on the accuracy of discriminating threshold stimuli, which we have considered in experiment 1. Nevertheless, there is some evidence that probability effects both reaction time and accuracy in a similar manner. Giordano, McElree and Carrasco (2009) concluded that the amount and rate of target information accrual are both affected by cue reliability and increase as a function of cue reliability, in agreement with the suggestion, by Johnson and Yantis (1995), that the sampling of visual information at a location depends on the expectation of observing a target there. These results may not be obvious since several studies have found that the effects of various experimental variables on accuracy and reaction time effects can differ greatly (Santee and Egeth, 1982; Mordkoff and Egeth, 1993). Santee and Egeth (1982) found, using a flanker effect paradigm, that repeating a target letter in a display made RT faster, but diminished identification accuracy. Accordingly, it was proposed that accuracy is subject to early perceptual interference, whereas RT effects reflect biases in response selection arising at the level of decision processes (Mordkoff and Egeth, 1993). Several authors have proposed the two measures may not reflect the same underlying process (Mordkoff and Egeth, 1993; Prinzmetal, Mccool and Park, 2005). Though we used differed stimuli (threshold and supra threshold motion) and response sets (4 alternative forced choice and simple reaction time) between the accuracy and RT tasks and find similar patterns of performance modulation. Thus when detection of salient motion and the discrimination of threshold motion are the dependent variables, accuracy and reaction time measures can show similar patterns of performance.

Models, which assume the distribution of attention simply reflects the probability distribution of target locations, independently of the stimuli and task demands involved, are more compatible with the finding of similar patterns of validity effects for both discrimination accuracy and the response latencies of detection. Uncertainty models propose that differences in behavioural performance arise from uncertainty about the stimuli or the response required and do not assume that the effects of spatial expectancies depend on preallocation of resources or increases in perceptual sensitivity at target locations (Lappin and Uttal, 1976; Palmer, Ames and Lindsey, 1993; Palmer, 1994). Some of these models assume a decision stage where sensory evidence is weighted by the observer's own expectations and determines the subject's response (Eckstien, Shimozaki and Abbey, 2002; Eckstien, Pham and Shimozaki, 2004). The proposal that probabilistic cues may instead operate by increasing the rate of sampling visual information at a location, the magnitude of increase in proportion to the probability of the target being at each of the possible target locations, is also compatible with our findings (Johnson and Yantis, 1995). This model also fits well with more recent findings from Giordano, McElree and Carrasco (2009). This information accrual model could also account for the similarity of RT and accuracy data, as each could be affected by the rate of sampling. By contrast, if attention only operated by increasing perceptual sensitivity, this should exert little influence on detection of supra-threshold stimuli, leading to differences between the effects of cue probability on response latencies and discrimination accuracies.

Probability matching and the two-process model

Jonides' (1980, 1983) two-process model assumes that attention is either maintained diffusely over all locations or focused at the cued location on a given trial. Furthermore, a probability matching strategy is used which determines the proportion of trials on which attention is in a focused or diffuse state (see chapter 1 for an extensive summary of those studies). Previous studies have examined whether a probability matching strategy is how cue reliability affects cue utilization (Jonides, 1983; Eriksen and Yeh, 1985; Madden 1992). Johnson and Yantis (1995) used a Multinomial Maximum Likelihood Mixture (MMLM) procedure, that considers the entire distribution of RT responses, to test the predictions of the two process model. These authors reasoned that the two process model mandates that the empirical RT distribution is a mixture of two separate distributions, corresponding to response latencies when participants are either attending the cued location or monitoring the entire display. In order to deconvolve the two distributions from the overall distribution of response latencies, one needs to consider more than just the first moments, namely the mean RTs, but also higher order moments, such as the standard deviations of the distributions (Yantis, Meyer and Smith, 1991). In our experiments, the probabilities that participants attended the cued location were estimated for each cues' reliability level and participant. A probability matching strategy accounted well for the effects of cue reliability on the response accuracies in the motion discrimination experiment, but not for response latencies in the speeded detection task. This result suggests that, at least in the discrimination task, deviations from optimal cue use may be due to a suboptimal strategy. The two-process model with probability matching did not account well for the data in the motion detection data. The lack of evidence for probability matching with the RT data is in accordance with Johnson and Yantis' (1995) conclusions, but not those of others (Madden, 1992, Eriksen and Yeh, 1985), who only examined mean RT data.

The implication of the finding that cue reliability may have different effects on discrimination accuracy and response latencies, despite the similarities discussed in the previous section, may be, after all, that participants did utilise cued information differently in the discrimination and speeded detection tasks. Potential explanations may include that the representation of either cue reliability or, more likely, the value of attending the cued location, may differ between the two tasks. For example, multiple representations of cue probability have been found across the brain, with veridical representations of probability in the striatum and biased ones in higher order prefrontal regions (O'Doherty et al., 2004). Alternatively, if the deployment of attention depends on a decision process based on the valuation and comparison of action outcomes, e.g. the effects of attending the cued location vs. attending all location equally, then one may expect differences between the detection vs. the discrimination task, since the former outcome is associated with a latency difference and the latter with an accuracy difference.

Distortions in cue reliability

Analyses conducted in experiment 1 suggest that there may be some distortion in the representation of or utilisation of cue reliability compared to the objective probabilities of observing a target at the cued location. Though the distortions at higher reliabilities are quite small and not significantly different from the objective reliabilities used, the 25% non-informative cue showed a large distortion (though this distortion may be due to mechanisms described in the Non-informative cueing effects below). Considering that these effects do appear similar to distortions that are seen in decision under risk (Khaneman and Tversky, 1979) and that we find evidence of probability matching under a two process model for the same experiment, there is a strong suggestion that probabilistic information may be utilised in a similar manner to decision processes and display similar sub optimalities in laboratory conditions. This finding is in line with similar recent evidence in perceptual decision making (Vincent 2011), showing that a baysian optimal observer actually performs more closely to humans when the prior expectations of probability are biased. These priors are then combined with the incoming sensory evidence. This account implies that the biases seen arise from internal representations rather than from distortions in the utilisation of cue information, especially as Vincent (2011) used both endogenous and exogenous cues suggesting that the biases may generalise to various cueing types and be less a factor of how a cue is utilise and more an innate bias. This evidence is supportive of the idea that our findings may be representative of distortions observed in decision under risk; further data collection could make this finding more robust. The purposes of this bias may be adaptive outside of the laboratory environment as spatial probabilities are rarely static in the natural environment (Martins, 2005, 2006) and that participant's performance shows that humans will look for statistics within a scene that are based on the assumption that probabilities in the environment are dynamic. This is true even when there *are* static probabilities and participants are exposed to these spatial expectations for over an hour (Vincent, 2012).

Non- informative cueing effects

Validity effects were present following non-informative cues, which provided no spatial information. One would expect that since attending the location, indicated by a non-informative cue, should not provide any performance benefit, participants would restrain from doing so, contrary to what we found. Moreover, the overall performance following a non informative cue was worse, even though non significantly so, than that observed following neutral cues, in both the discrimination and speeded detection tasks, confirming that participants did not benefit from attending the location highlighted by a non informative cue. Why then are endogenous shifts of attention observed following non-informative cues? Participants, prior to commencing the experiment were explicitly informed about the probability that the cued location would contain the target motion as well as the fact that non-informative cues provided no useful information, suggesting that these findings did not simply reflect poor understanding of the task demands. On the other hand, Prinzmetal McCool and Park (2005) found no effects of non-informative cues, when trials preceded by non-informative cues were presented in separate blocks from those preceded by informative cues. This suggests that the utilization of non-informative cues may reflect some form of perseverative behaviour in conditions where trials of different cue reliabilities are mixed, as in our study. Previous studies have found validity effects with non-informative endogenous cues (Tipples, 2002; Ristic, Friesen and Kingstone, 2002) and for spatial verbal cues (Hommel, Pratt, Colzato and Godjin, 2001). It was suggested that these represent shifts of attention evoked by automated effects of decoding the cue meaning, rather than a distortion in the internal representation of the cue's value. The idea that endogenous cueing may have automatic components is given further credence by the finding that validity effects are found with cue target SOAs of 100ms (Tipples, 2002), which may be too brief for a voluntary deployment of attention. even though this remains a contentious issue (Jonides, 1981; Muller and Rabbitt, 1989; Krose and Julesz 1989; Nakayama and Mackeben, 1989; Cheal and Lyon, 1991). Jonides (1981) found that non informative cues with a short cue target SOA of 100ms did not elicit validity effects in contrast to Tipples (2002). Gibson and Bryant (2005) provided evidence that both cue duration and task demends alter non-informative cueing effects. They replicated Jonides' (1981) stimulus set up that entailed a letter discrimination task with eight locations containing letters, one being the target letter. They presented a non-informative, central arrow cue either for 25ms or 200ms in separate blocks. Both of these cue conditions were followed by cue-target SOA of 50ms, 300ms or 550ms**.** The 25ms long, non informative cue did not elicit significant validity effects at the 50ms SOA, as found in Jonides (1981), but the 200ms and 550ms SOA conditions did, overall the validity effect increasing in magnitude with increasing SOA. The 200ms duration non-informative cues elicited significant validity effects, even at 50ms SOA and increased in magnitude with 200ms and 550ms SOA. Our own results, obtained with long cue durations and SOAs, are consistent with these findings. In a second experiment, Gibson and Bryant (2005) interleaved, within blocks, the two cue durations. In mixed blocks, the 25ms cue and the 200ms cue both lead to validity effects in the short SOA trials, suggesting that the differences observed in the blocked condition reflected more than simply differences in cue duration. In a third experiment, the authors changed the visual appearance of the cue, by removing the triangular head from the arrow cue on some proportion of the trials. When this cue appeared participants had to withhold a response, thus ensuring that each cue had to be visually analyzed. A large validity effect was then found for trials with 25ms cues and 50ms SOA. Thus Gibson and Bryant's (2005) experiments suggest that the depth of cue processing, as determined by the cue duration and the demands of the task, determines the extent of a non-informative cue's behavioural effects.

Previous literature (Hommel et al., 2001; Geng and Behrmann, 2005) has emphasised the dissociation between static spatial probabilities and an explicit endogenous cue. Non-informative central cues and words indicating a location both create a validity effect in detection and discrimination paradigms that is additive and dissociable from the effects of knowing a probability that a stimulus will appear in a given location (Hommel et al 2001). Hommel et al (2001) carried out a study in which one location had an 80% probability of containing a target but a non-informative central endogenous cue indicates a random location each turn. They found that both the probability location and the central cue produced significant validity effects in RT of discriminating colours but that there was no interaction between these effects, as such both exerted an influence on attention but neither interfered with the effect of the other. The validity effects of the non-informative cues were of similar magnitude to another experiment they conducted in which there were only non-informative cues. As such location probability does not appear to alter the attentional effects of a non-informative central cue stimulus when simultaneously deployed, though the location probability effect was larger than the central cue effects.

Experiment 3 also shows further effects of cues that indicate non-informative or non target locations, in this case with multiple locations, attained in the multiple location anti-cue condition. The multiple location anti cue reduces discrimination performance to that of the neutral cue despite containing the same spatial information and even the same distribution of spatial information as the multiple location pro cue. This could be accounted for by several explanations. Firstly that there could be an automatic orienting that occurs to the indicated locations and this leads to attentional processing at incorrect locations. Secondly the indicated directions are suggestive enough that inhibitory cognitive processes need to be deployed to interpret the cues correctly in comparison to the pro cues and this may lead to a detriment to performance. This second account is perhaps analogous in a purely conceptual sense to the Stroop Task (1935), but regarding spatial information. The directional symbolic nature of the cue conflicts with the intended utility of the stimulus to performing the task. To elaborate; in the Stroop task the symbolic representation of word stimuli conflicts with the requirement to report the basic features of these stimuli in the task (the colour a word is naming or the colour of the ink the word is printed in). Though the symbolic indication and intended utility of the cue stimuli are both spatial in nature in our paradigm. This also reinforces the idea that in studies employing longer cue durations, the symbolic nature of the cue being fully processed is what drives attentional effects. Our current results using a long cue presentation of 1s in all experiments also corroborate this interpretation and may help explain why these non-informative cue effects are so robust across our experiments. Thirdly that the strategy of attending to the pro cued locations became conflated with the use of the anti cue to the detriment of performance in the anti cue condition.

The attending of incorrect cued locations and suggestive directional influence of the cue in either case are both indicative of a tendency toward using the cued information provided even when it is suboptimal. In the case of the anti cue the cued locations were explicitly and with certainty not going to contain the target information, in contrast to the 25% valid non-informative cue in experiments 1 and 2. This implies the symbolic influence of the cue that could be in operation is strong enough to elicit cue use even in the face of explicitly incorrect orienting for obeying the cue information. The anti cue stimulus was identical in terms of spatial information and distribution of that information but clearly dissociable from the multiple location cue. This implies the reduced performance observed in the anti cue condition is due to the conflict between the locations indicated by the physical stimulus and the top down knowledge that these are with certainty the incorrect locations. Therefore the anti cue findings imply endogenous cues can elicit attentional orienting that is clearly dissociable from the probabilistic contingency of observing target information at a given location. Though the 25% reliable non-informative cues of experiments 1 and 2 could be explained by a probability matching strategy in which attention is focused on the cued location for 25% of the trials, which is obviously not an adaptive strategy.

Dividing spatial attention with four locations

Recent literature has provided evidence that it may be possible to divide attention among multiple locations (Gobell, Tseng and Sperling, 2004; Kramer and Hahn, 1995; Hahn and Kramer 1998; Bichot, Cave and Pashler, 1999; Awe and Pashler, 2000), in contrast to earlier models of spatial attention (Posner, Nissen and Ogden, 1978; Posner, 1980; Posner, Snyder and Davidson, 1980; Eriksen and Yeh, 1985; Eriksen and St. James, 1986). Though the definition of divided attention proposed by some (Jans et al., 2010) considers the literature in support of divided attention to be inconclusive in providing definitive evidence of multiple attentional foci. One of their main criticisms of the literature concerning divided attention is that intervening space between foci of attention must be monitored and show no attentional enhancement. In experiment 3 our goal was to separate the effects of spatial uncertainty and the spatial distribution of the target on performance by comparing the behavioural consequences of using cues which highlighted multiple locations where the target could appear and cues which highlighted one location where the target was likely, but not certain to appear. We reasoned that the distribution of attention, following multiple locations cues, will differ from that following one location probabilistic cues.

In experiment 3 we find similar performance for cues that are matched for spatial information provided, but which require different distributions of attention across spatial locations. This implies that attention can be distributed differentially across multiple locations according to probability of finding a target as suggested by Shaw (1977). This contradicts zoom lens models of attentional distribution (Eriksen and St. James, 1986) in that neighbouring regions should show similar attentional facilitation, thus cue configurations indicating closely clustered locations should elicit the highest performance. We consider the configuration of multiple location cues in our analysis in order assess whether performance was influenced by the distance between cued locations. If closely clustered cue configurations resulted in greater performance this would imply that a single focus could encompass all nearby locations and that multiple distant locations could not be attended simultaneously. The configuration analyses suggested no effect of location proximity (see figures x y z) as the most clustered configurations were at various rankings of behavioural performance. This would suggest that the distribution of target locations did not impair performance, as configurations containing more disparate locations should decrease performance if splitting of attention into multiple foci is not possible. As such we provide some evidence that attention may be split between multiple locations at once. The three cue configuration analysis showed a significant effect when collapsed across configurations that were mirror symmetric across the vertical meridian. Thus the only effect of cue configuration we found was when collapsing across conditions such that the main factor was the number of cued locations in the upper or lower visual field. This implies that the performance benefits may be due to a benefit for processing stimuli in the lower visual field and not dependent on the spatial proximity of the cued locations, thus this makes a single focus of attention less likely to account for the findings. Observing figure x it appears that the number of cued locations in the lower visual field is a strong factor in determining configuration effects as has been previously reported, whether this is due to attentional (He, Cavanagh and Intriligator, 1997) or sensory (Carrasco, Talgar and Cameron, 2001) factors. Taking the benefit for cueing lower visual field locations into consideration in tandem with the near significant effects of lower visual field location, independent of cueing. Suggests that the lower visual field and also hemi-field of cued locations may both have contributed to the three location configuration effects. Therefore the configuration of cues does not appear to have strong effects on performance, in addition the proximity of locations in the configurations does not appear to be a factor in determining these minor effects.

Previous studies using multiple cued locations have suggested a multiple location indexing strategy whereby indexes are assigned to separate locations that are sufficiently separated and attention deployed directly to the location with the strongest signal (Wright, 1994). When two adjacent stimuli are cued it is proposed that this leads to a single focus of attention between these locations leading to reduced performance. The authors also find that separated non adjacent cued locations show similar RT performance to single cued locations, but that if a target appears in a location between two separated cued locations it does not show a benefit, thus a single large focus of attention encompassing both cued locations and the intervening space is not a viable explanation. Our results do not agree with this entirely as the clustered configurations were not the lowest ranking. These results thus suggest that multiple cueing does not fit within a zoom lens hypothesis as suggested by Kiefer and Siple (1987).

The results of experiment 3 using four possible target locations suggest that attention can be distributed flexibly across the visual field and that there are no costs for dividing attention in different and uneven distributions when spatial uncertainty is constant as neither the distribution of spatial information nor the distance between multiple spatial locations appeared to result in costs to overall performance. From experiment 3 the results imply that there was no cost for dividing attention in single cue as compared to multiple cue conditions when spatial uncertainty was accounted for. Experiment 3 may be in agreement with findings that endogenous spatial attention can be distributed unequally over multiple locations and that the differential distribution of spatial attention over multiple locations does not impair attentional performance (Gobell, Tseng and Sperling, 2004; Kramer and Hahn, 1995; Hahn and Kramer 1998; Bichot, Cave and Pashler, 1999; Awe and Pashler, 2000). The fact that single location cue performance was similar to multiple location cue performance when spatial uncertainty was matched implies that attentional allocation was scaled effectively in the one location condition to utilise the abstract representation of value afforded by an explicitly stated probability. The distribution of spatial information in the two location condition required an equal split of attention over two cued locations without any interpretation of a numerical value. Assuming the equal performance is due to equal spatial information, the probabilistic single location cue would require interpretation of the proportion of trials that were valid to attain an equal performance. If this information were not utilized the cued location would either always be attended which should lead to much greater performance than that attained. These findings are consistent with models of attention that suggest that sensory information from different locations are weighted according to the expectation that they will contain the target (shaw, 1977; Jonides 1980; Cameron, Tai, Eckstien and Carrasco, 2004; Morgan, Ward and Castet, 1998; Shiu and Pashler, 1994). Alternatively sensory information at locations could be weighted at the level of a decision process (Eckstien, Shimozaki and Abbey, 2002; Eckstien, Pham and Shimozaki, 2004). This is also consistent with the idea that a pre allocation of processing resources is made to multiple spatial locations, with the proportion of resources representing prior knowledge of the probability of finding target information (Jonides, 1980; Yantis and Johnson, 1995).

The apparent scaling of attention in probabilistic cueing conditions, strengthens the hypothesis that this scaling is how the magnitude of the validity effect was modulated by probabilistic cues in experiments 1 and 2. If probability matching is the method by which this is implemented, then it is carried out in such a way in this study as to match information theoretic measures of spatial uncertainty in utilising the probabilistic information of the single location cue. Assuming a two process model (Jonides, 1980; Eriksen and Yeh, 1985) with probability matching cannot account for these findings, as this assumes attention is entirely focused on one location or dispersed over the entire stimulus array. As such attention should be focused on one of the cued locations in the multiple cue condition, so when two disparate locations are cued there should be lower performance as attention must be at only one location distant from the other possible target location. Though it is not unreasonable to postulate that probability matching may be the strategy by which the scaling of attentional performance is achieved in only the probabilistic cue condition or that some form of probability matching that does not assume a two-process model could account for the data. Taken as a whole the findings suggest that a spatial uncertainty model that takes into account probabilistic values allows this variable distribution of attention across locations.

Dividing attention with six locations

When we increased the complexity of the task by using displays, which contained six stimuli, we found that cues matched for the amount of spatial information provided no longer elicited identical performance. Discrimination accuracy was lower following multiple locations cues compared to one locations probabilistic cue. This finding could have reflected limited ability to divide spatial attention among multiple locations, as some have suggested (Posner, 1980; Triesman and Gelade, 1980; Eriksen and Yeh, 1985). This interpretation would lead to the prediction that when the cued locations are in close proximity the performance decrements should be less severe than when the cued locations are further apart. However, we found only weak effects of the cue configuration, which did not appear to reflect an effect of cued location proximity, but more likely a visual field difference in attentional processing between lower and upper visual field (He, Cavanagh and Intriligator, 1997). If you instead interpret this finding in terms of a limitation in resources that arises from distributing these resources over an increased number of locations, then the probabilistic rather than the multiple location cues should show the deficit. The multiple location cue conditions have the same number of locations cued in a four or six location stimulus set in comparison to the probabilistic cues, which now have five other possible locations with the six location experiment instead of the three other possible locations when there are a total of four locations. Therefore the multiple cue condition would be expected to show greater performance than matched probabilistic cues, as the probabilistic cues have more locations to allocate resources to. Thus even if it is possible to flexibly allocate smaller degrees of attentional processing resources over the remaining five locations, this should still lead to a poorer performance overall, due to the poorer performance in invalid trials that would result compared to that with three remaining locations in probabilistic cue conditions for experiment 3. Therefore if any detriment should occur between the four and six location experiments due to an inability to divide resources, logically this should be in the probabilistic cue conditions as multiple location cues should be identical between the paradigms. The fact that multiple location cues rather than probabilistic cues show a detriment therefore does not support the assertion that limitations in attentional capacity are the cause of these detriments. Though there may be limitations in general processing resources rather than attention specifically (Broadbent, 1958; Kahneman, 1973; Moray, 1967; Wickens 1984).

Spatial working memory and limitations in attending multiple locations

Experiment 5 provides evidence for an alternative explanation of the performance deficits observed when cuing multiple locations compared to single locations in Experiment 4. In Experiment 5 participants were asked to recall cued locations immediately after cue presentation. Performance was at ceiling for cues indicating one, two or three locations. However when participants had to prepare to use the spatial information provided by these cues, in blocks where each trial required either recall of cued locations or discrimination of cued locations, recall of cued locations was impaired. In blocks in which participants either recalled cued locations or reported the motion direction, motion discrimination accuracy was no different than in blocks where participants only reported the direction of motion. Thus the requirement to engage attention appears to conflict with the requirement to remember the cued locations. This is in opposition to the view espoused by (Awh et al 1998; 2001), namely that attending a location allows rehearsal of that location in spatial working memory. The authors presented black dots to mark locations that had to be recalled later, a colour discrimination task was presented during the retention interval. The stimuli for the colour discrimination task were either a small circle presented at a location different from the locations of the memorised targets or a large circle encompassing the entire array of possible target positions (thus requiring no shift of attention to perform the colour discrimination task). When attention was shifted to the small circle at a different location, recall of the dot locations was worse, however there was no decrement in recall accuracy when no shift of attention was required. Also control conditions were run in which only colour discrimination or only location memory were required, memory responses showed a smaller detriment between static attentional task and no attentional task than between shifting attention and having no attentional task implying that detriments were not due to difficulty of dual task but the requirement to shift attention.

Awh jonides and Reuter-Lorenz (1998) established a functional role for attention in working memory by demonstrating that shifting attention to non memorised locations impedes recall of the memorised locations. These results are similar to those of our combined recall and discrimination cueing experiments and strengthen our hypothesis that shifting spatial attention causes the observed working memory decrements rather than cognitive demands of preparing for two possible tasks. Though Awh (Awh, Jonides and Reuter-Lorenz, 1998; Awh and Jonides, 2001) show the dependent link of spatial attention on working memory, they suggest that shifting spatial attention to an important location plays a role in spatial rehearsal in working memory and that attending a *different* location is what disrupts this rehearsal. Our present studies contradict the notion that attending a spatial location allows rehearsal of that location in working memory by showing that when multiple target locations have attention shifted to them, recall of these same locations suffers. This seems counter intuitive and certainly at odds with the proposal of Awh and colleagues that attention to a location is identical to maintaining the location in working memory. However we do find that with a single cued location, in discrimination and recall conditions, there is no impairment to recalling the cued location compared with recall only conditions.

Thus it appears that spatial attention and working memory processes conflict only when there are multiple cued locations. So our single location discrimination and recall conditions are in agreement with Awh and colleagues, but the multiple cue conditions show a conflict between the processes of spatial attention and spatial working memory.

Thus in our paradigm at least there would appear to be a particular delineation between the attentional processing and memory representation of the same locations. This is reflected in our calculation of the information theoretic cost of retaining a cue configuration in working memory (See table), which demonstrates that there is an increasing cost in terms of information for storing increasing numbers of locations. This increase in information storage cost of a cue is also independent from the amount of spatial information that the same cue conveys. It seems reasonable to postulate that since we found no detriment to recall for one location cues when also deploying attention, that the recall detriments for multiple cued locations relate to the large memory load associated with retaining multiple location cues in working memory. This has strong implications for the nature of attention in general and more specifically for the divided attention literature and in particular endogenous cueing paradigms that consider divided attention. Namely that deficits in retention of spatial information may underlie reduced performance in divided attention tasks especially if they involve endogenous spatial cueing. Our information theoretic model of predicted performance accounting for influences of memory deficits supports this notion. Also the findings imply that the internal representation of an endogenous central cue may be in conflict with the intended attentional shift that may result from that cue.

This implies that WM and attention are overlapping processes in that they may share common resources. There is much contention in the literature concerning whether capacity limits for attention and working memory are independent (Fougnie and Marois, 2006) or shared (Zhang et al, 2010). Some authors propose that the two processes draw from a shared, capacity limited, amodal resource (Fougnie and Marois, 2006). Some literature has suggested the capacity limit of working memory is reducible to that of attention, though these assertions include more than just the spatial working memory component that we consider here (Cowan 2001; Rensink, 2000a). Alternatively, for completeness of argument, one could consider the possibility that the capacity limit of attention could be reducible to that of working memory resources. Our research would seem to lend some credence to this explanation and contradict the notion that working memory capacity is reducible to that of attention. In particular the similar attentional performance between blocks containing discrimination trials only and blocks where discrimination and recall trials were interleaved suggests that attention was not depreciated by working memory processes. As such attention either by default draws on working memory resources or spatial attention took precedence in this paradigm as participants were instructed to consider the attentional task as the primary task. If anything attention would appear to be limited by working memory resources, in that working memory resources are depleted when attentional shifts are required. Alternatively it could be considered that perhaps an executive process mediates what proportion of resources is allocated to each process and in this case attention was given priority. This is partly in accordance with Fougnie and Marois (2006) who propose an amodal cognitive resource is used by both attention and working memory processes, however their claim that these processes do not interfere or share resources is at odds with our findings. In contrast to Fougnie and Marois (2006) in our paradigm it seems that there is a clear dependence of attention on working memory and an overlap in their capacity limitations, at least when the working memory component of a task requires spatial locations to be represented. Previous research has shown that when the spatial location of an array of objects is not a requisite memory component in the task there does not appear to be any interference from simultaneously executed attentional processes (Zhang et al, 2010).

People do not learn cue reliability or do not effectively use cue reliability estimates

Cue reliability has been shown to have strong effects on cueing effects (Jonides, 1980; Giordano, McElree and Carrasco, 2009; Geng and Behrmann, 2005). However when cue probability is not known before hand and needs to be learned as in experiment 6, we observed no significant group level effects of reliability on motion discrimination accuracy. This implies that either cue reliability was not learned or that the estimate of cue reliability was not utilised to fine tune preparatory processes. This lack of learning effects is in accordance with Droll, Abbey and Eckstien (2009) who reported that learning of cue reliabilities does not occur without feedback about performance or the target. We did observe that two out of five subjects showed evidence that learnt cue reliability influenced attentional processes in that invalid trials showed greater discrimination errors with greater cue reliability. This modulation did not extend to concomitant increases for valid trials and overall did not improve discrimination accuracy, thus what influence we do observe of uncertain probability on attentional processes does not appear to be adaptive. Though this pattern of results in these subjects implies some recognition and application of probability changes to attentional processes. A significant within group difference

was observed for the effect that cue reliability had on discrimination error suggesting large inter-individual differences in cue learning and or utilisation.
Bestmann et al (2008) demonstrate that visual cues that were either valid or invalid with respect to indicating a correct motor response, could be learned even in a volatile system with changing cue reliability similar to that used for our trial set up with changing cue reliability for a single visual cue, reaction times and preparatory motor output both reflected changes in probability. Thus perhaps detection of a supra-threshold stimulus may have shown greater modulation by estimated cue reliabilities. Though the finding of motor response latencies being modulated by learned probabilities in Bestman et al (2008) could be dependent on their paradigm using feedback, furthermore the association between cue use and subsequent correct responses was clearer in a motor reaction time task (we expand on the issue of associating action and outcome in learning unknown probabilities below).

For those participants who did appear to learn and utilise cue reliability we created a bootstrap simulated data set that re-arranged trials independent of reliability conditions but dependent on the validity of only the previous, one, two or three trials. This created a data set in which each trial was randomly exchanged with another trial that had the same preceding sequence of cue validities in preceding trials. Thus discrimination errors were now representative of recent trial sequence only and not the reliability block that was being responded to in the original data and trial set up. Therefore if the average discrimination error within a reliability block did not significantly alter, this error value could be attributed to the influence of short trial sequences rather than the total sequence of trials in a reliability block, however if the average error did alter with simulated data then the original value is attributable to interpreting of a longer sequence than the preceding one, two or three trials. We observed that the more trials that were considered in the sequence to compose the data set the closer the simulated data became to the actual data. Though considering the three preceding trials did not produce data that were similar to the observed data. This implies that a greater number of trials was taken into account in producing a reliability estimate or at least in utilising cue information in accordance with underlying cue reliability. This analysis also provides some evidence in opposition to the claim that the effect of spatial probability on RTs and first saccades to discriminate simple stimuli is entirely attributable to sequential dependencies (Walthew and Gilchrist, 2006; Rabbit, Cumming, Vyas, 1977; Rabbit, Cumming Vyas, 1979). Sequential dependencies refers to a previous trial's target location affecting performance on the current trial, when the trial structure was altered such that the target location in the previous four trials were different from the current trial but spatial probability was kept the same, there was no effect of spatial probability on saccades (Walthew and Gilchrist, 2006). However this is somewhat different to the non-static spatial probabilistic contingencies afforded by a spatial cue.

What is being learned: cue reliability or the cue utility value?

Droll, Abbey and Eckstien (2009) show that participants are able to give accurate estimated reports of cue reliability without showing any behavioural benefits based on these seemingly well represented cue reliabilities. This dissociation between an internal representation of cue reliability and utilisation of said cue reliability raises questions about what exactly is being learned in these paradigms. Given that this internal estimate seems to be accurate regardless of whether or not feedback about performance or information about scene statistics is provided, it could be argued that this probability estimate is not sufficient to influence performance by its self. It may therefore be necessary to learn the value of attending the location assumed to be the target, which may require more complex computations than just cue probability, with multiple possible actions and outcomes such as for example attending the target location as opposed to attending a different location or all locations diffusely. Feedback may improve this latter type of learning by providing acknowledgement that utilising the cue leads to the intended outcome; a correct response. It may be that an estimated utility of using the cue is updated more effectively with reinforcement due to association between action and outcome, and is independent to an estimate of the cue probability, which may receive no benefit from reinforcement feedback as it does not require the successful perception of the target. It may instead be that the estimate of cue probability is drawn on in determining the degree to which this representation achieves useful outcomes, as a separate value, this value is then used to determine the degree to which the representation is applied to perceptual processes. Providing feedback under this latter assumption would strengthen the association between the value of attending and the probability of the cue. In the latter interpretation there are some parallels with recent models of cue learning (Yu and Dayan, 2005). Yu and Dayan (2005) outline two types of uncertainty, ‘expected uncertainty’ being an estimate of a probabilistic contingency and ‘unexpected uncertainty’ being an estimate of whether the assumed probability is still relevant to a current context. In this case the unexpected uncertainty about this probability prevents it from being utilised at the point of implementing this information to enhance discrimination performance, which assumes something like a decision process is at work.

Is awareness of this cue reliability necessary for perceptual enhancement? Choice paradigms always show the outcome of a decision and associated value awarded and hence trial to trial updating of probability representation based on outcome and value of the outcome should have no impediment and be available to conscious awareness. However the act of making a decision and knowing the outcome was correct is not sufficient to bias choice even in full awareness if it has no consequence. This suggests that equally, in perceptual choices, if you do not receive confirmation that this outcome matters in terms of perceptual performance may not update your consideration of cue value. This may be the case even with conscious awareness of the cue's value in localising a target stimulus. The frequency of a favourable outcome in repeated choices is strikingly accurate when explicitly reported (Ungemach et al., 2009) and remains robust in memory after interference tasks, more so than when these frequencies were not experienced but explicit probabilistic information was provided (Lejarraga, 2009). Therefore it is possible to bring these representations to conscious awareness, even if they do not affect behaviour in some cases (Droll, Abbey and Eckstien, 2009).

Summary

Considering our findings as a whole, probability would appear to have a strong influence on the distribution of endogenous spatial attention, seemingly by reducing spatial uncertainty about the likely location of a target. Probability matching may be the mechanism by which accuracy of threshold motion is modulated by cue reliability, though this does not seem to account for reaction times to detect supra-threshold motion.

Endogenous cues also exert a strong influence on attention that appears to arise from the symbolic representation of spatial information, independent of the influence of spatial probability. The tendency to use the symbolic information occurs even when cues indicate explicitly incorrect locations and reduces performance, likely by requiring that the tendency to use this symbolic influence be inhibited.

Attention appears to be flexibly distributable based on the information provided by spatial cues with no costs to performance when spatial uncertainty is accounted for. Limitations in this ability to distribute attention arise in complex scenes, but do not appear to be due to attentional capacity limitations or an inability to divide attention across multiple locations. The limitations appear to arise from working memory limitations in representing multiple cued locations when simultaneously attending these same cued locations. The processes of attending spatial locations, therefore appears to interfere with working memory processes. This suggests attention draws from working memory resources or that working memory and attention draw from the same shared, amodal, limited capacity cognitive resource.

Probability and spatial uncertainty therefore have a strong effect on the distribution of spatial attention with endogenous central cues, which is explained well by information theoretic models. Though there are additional factors that influence attention such as symbolic influence of central cues that can elicit endogenous orienting. Furthermore attending multiple locations is limited by the working memory requirements of storing this information in more complex scenes.

Conclusions

In conclusion knowledge of probabilistic contingencies can be utilized to flexibly distribute spatial attention across the visual field without cost when spatial uncertainty is accounted for. Though limitations arise in the distribution of attention to multiple locations when scenes become increasingly complex, seemingly from inability to store spatial representations rather than discriminate stimuli at several locations. It is clear that factors other than probability also influence spatial attention in endogenous cueing paradigms. Symbolic representations of spatial contingencies derived from central cues appear to be processed by default and exert a separate influence on spatial attention in addition to spatial probabilities via top down mechanisms.
Thus spatial uncertainty models account well for the influence of probability on spatial attention but additional influences can be found from several sources that are attributable to processes other than attention, such as spatial working memory and higher order symbolic representations. The top-down representations that guide attention exert an additional influence and the underlying working memory representations that attention is dependent on prevent it from being optimally distributed.

Further research might benefit from utilizing models that do not assume probability matching for comparison with MMLM procedures. In addition this type of model could be used to consider how multiple location cues operate in comparison to probabilistic cues as the information matched cue paradigms provided evidence against probability matching with a two-process model. Further manipulations of spatial uncertainty utilizing our recall and discrimination task set up with four RDKs could strengthen our model of spatial working memory limitations as this number of locations showed no detriment to attentional performance with multiple location cues. An iterative Bayesian model of cue probability learning could help determine how representations of probability are updated on a trial-by-trial basis and whether cue probability representations are distorted. In addition a further consideration of the fact that only invalid trials show cue reliability modulations has yet to be explained, this may be a good starting place to consider differences in orienting and reorienting that may occur in this paradigm and if this relates to uncertain representations.

References

Allais, M. (1953) Le Comportement de l'homme Rationnel Devant le Risque Econometrica, **21**, 503–546.

Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track? Evidence for a resource-limited tracking mechanism. Journal of Vision, **7**, 1-10.

Anderson A. J., Carpenter R. H. S. (2006). Changes in expectation consequent on experience, modelled by a simple, forgetful neural circuit. *Journal of Vision*, **6**, 5, 822–835.

Awh, E., Jonides, J., Reuter-Lorenz, P.A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology-Human Perception and Performance*, **24**, 780–790.

Awh, E., Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, **5**, 119–126.

Awh, E., & Pashler, H. (2000). Evidence for split attentional foci. *Journal of Experimental Psychology: Human Perception and Performance*, **26**, 834–846.

Baddeley, A.D., Hitch, G. (1974). Working memory. In The Psychology of Learning and Motivation (Bower, G.A., ed.), 48–79, Academic Press.

Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. Visual Cognition, **10**, 949–963.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, **16**, 215–233.

Bays, P. M., Catalao, R.F., Husain, M. (2009). Precision of visual working memory is set by allocation of a shared resource, Journal of Vision, **10**, 1-11.

Bergen, J. R., Julesz, B. (1983) Parallel versus serial processing in rapid pattern discrimination. *Nature*, **303**, 696-698.

Bestmann, S.H.L.M., Blankenburg, F., Mars, R., Haggard, P., Friston, K.J., (2008) Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. *Current Biology*, **18**, 775–80.

Bichot, N. P., Cave, K. R., & Pashler, H. (1999). Visual selection mediated by location: Feature-based selection of noncontiguous locations. *Perception & Psychophysics*, **61**, 403–423.

Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. *Science*, **308**, 529–534.

Bowman E. M., Brown V., Kertzman C., Schwarz U., Robinson D. L. (1993). Covert orienting of attention in Macaques: I. effects of behavioral context, *Journal of Neurophysiology*, **70**, 431-443.

Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology,* **47**,191-196.

Broadbent D. E. (1958). Perception and Communication*. Oxford: Oxford University Press.*

Brainard D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 437-442.

Carpenter, R. H. S., Williams, M. L. L. (1995). Neural computation of log likelihood in the control of saccadic eye movements. *Nature*, **377**, 59-62.

Brackbill, Y. Bravos , A. (1962). Supplementary report: The utility of correctly predicting infrequent events. Journal of experimental Psychology ,**64**, 648-649.

Bradley, A. Skottun, B.C. Ohzawa, I. Sclar, G. Freeman R.D(1987) Visual orientation and spatial frequency discrimination a comparison of single neurons and behaviour. *Joural of Neurophysiology*, **57**, 755–772.

Bush, R. R., Mosteller, F. (1955) Stochastic models for learning. New York: Wiley.

Carrasco, M., Penpeci-Talgar, C., Eckstein, M. (2000) Spatial covert attention increases contrast sensitivity across the CSF support for signal enhancement, *Vision Research*, **40**, 1203–1215.

Carpenter R. H. S., Reddi B. A. J., Anderson A. J. (2009). A simple two-stage model predicts response time distributions. *Journal of Physiology*, **16**, 4051-4062.

Carpenter, R.H.S., Williams, M.L.L. (1995). Neural computation of log likelihood in the control of saccadic eye movements. *Nature,* **377**, 59-62.

Carpenter R.H.S. (1999). A neural mechanism that randomizes behavior. *Journal of Consciousness Studies,* **6**, 13–22.

Carpenter, R.H.S. (2004) Contrast, probability and saccadic latency: evidence for independence of detection and decision. *Current Biology,* **14**, 1576–1580.

Carrasco, M., Talgar, C., Cameron, L. (2001). Characterizing visual performance fields: Effects of transient covert attention, spatial frequency, eccentricity, task and set size. *Spatial Vision,* **15**, p 61-75

Corbetta M., Shulman G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews: Neuroscience*, **3**, 201-215.

Castiello, U., & Umiltà, C. (1992). Splitting focal attention. *Journal of Experimental*

Psychology: Human Perception and Performance, **18**, 837–848.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, **9**, 349–354.

Cavanagh, P., Hunt, A. R., Afraz, A., Rolfs, M. (2010). Visual stability based on remapping of attention pointers. *Trends in Cognitive Sciences,* **14**, 147-153.

Cherry, E.C. (1953). "Some Experiments on the Recognition of Speech, with One and with Two Ears". *Journal of Acoustic Society of America* **25** (5): 975–979.

Chun, M.M., Golomb, J.A., Turk-Browne, N.B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, **62**, 73–101.

Chun M.M., Jiang Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, **36**, 28–71.

Chun M. M., Jiang Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science,* **10**, 360–365.

Chun M. M., Nakayama K. *(*2000*).* On the functional role of implicit visual memory for the adaptive deployment of attention across scenes. *Visual Cognition,* **7**, 65–81.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, **24**, 87–185.

Downing, E. (2000). Interactions Between Visual Working Memory and Selective Attention. *Psychological Science,* **11**, 467—473.

Downing , C. J., Pinker , S. (1985). The spatial structure of visual attention. In Posner, M. I., Marin, O. S. M. (Ed.*), Mechanisms of attention: Attention and performance XI (pp. 171-187). Hillsdale, NJ: Erlbaum.*

d'Avossa G., Shulman G. L., Snyder A. Z., Corbetta M. (2006). Attentional selection of moving objects by a serial process. *Vision Research*, **46**, 3403-3412.

Castiello, U., & Umiltà, C. (1992). Splitting focal attention. *Journal of Experimental Psychology: Human Perception and Performance*, **18**, 837–848.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, **9**, 349–354.

Cavanagh, P., Hunt, A. R., Afraz, A., Rolfs, M. (2010). Visual stability based on remapping of attention pointers. *Trends in Cognitive Sciences,* **14**, 147-153.

Cherry, E.C. (1953). "Some Experiments on the Recognition of Speech, with One and with Two Ears". *Journal of Acoustic Society of America* **25** (5): 975–979.

Chun, M.M., Golomb, J.A., Turk-Browne, N.B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, **62**, 73–101.

Coull, J. T., Frith, C. D., Buchel, C., & Nobre, A. C. (2000). Orienting attention in time: Behavioural and neuroanatomical distinction between exogenous and endogenous shifts. Neuropsychologia, **38**, 808–819.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, **24**, 87–185.

d'Avossa, G., Shulman, G.L., Snyder, A.Z., Corbetta, M. (2006). *Vision Research,* **46**, 3403-3412.

Downing, E. (2000). Interactions Between Visual Working Memory and Selective Attention. *Psychological Science,* **11**, 467—473.

Driver J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, **92**, 53-78.

Droll J. A., Abbey C. K., Eckstein M. P. (2009). Learning cue validity through performance feedback. *Journal of Vision*, **9**, 1-22.

Druker, M., Anderson, B. (2010) Spatial probability aids visual stimulus discrimination. *Frontiers in Human Neuroscience*, **4** , 1–10.

Duncan, J., Humphreys, G.W. (1989) Visual search and stimulus similarity. Psychological Review, **96**, 433-458.

Eckstein M. P., Pham B. T., Shimozaki S. S. (2004). The footprints of visual attention during search with 100% valid and 100% invalid cues. *Vision Research*, **44**, 1193–1207.

Eckstein M. P., Shimozaki S. S., Abbey C. K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. Journal of Vision, **2**, 25–45,

Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. Journal of Experimental Psychology Monographs, **79,** 1-16.

Eckstein, M.P., Thomas, J.P., Palmer, J., Shimozaki, S.S. (2000) A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception and Psychophysics*, **62**, 425–451.

Eriksen C. W., Yeh Y. Y. (1985). Allocation of attention in the visual field. Journal of Experimental Psychology: Human Perception and Performance, **11**, 583–597.

Eriksen, C. W., St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics,* **40**, 225-240.

Eriksen, W. E., Hoffman, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Attention Perception and Psychophysics.* **14,** 155-160.

Estes , W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, **57**, 94-107.

Estes, W.K. Taylor, H.A. (1966). Visual detection in relation to display size and redundancy of critical elements. *Perception and Psychophysics*, **1**, 9–16.

Estes, W.K. Wessel, D.L. (1966). Reaction time in relation to display size and correctness of response in forced-choice visual signal detection. *Perception and Psychophysics*, **1**, 369–373.

Fougnie, D., and Marois, R. (2006). Distinct capacity limits for attention and working memory: Evidence from attentive tracking and visual working memory paradigms. *Psychological Science*, **17**, 526–534.

Franconeri, S. L., Alvarez, G. A., & Enns, J. T. (2007). How many locations can you select? Journal of Experimental Psychology: Human Perception & Performance, **33**, 1003-1012.

Franconeri, S. L., Jonathan, S., & Scimeca, J. M. (2010). Tracking multiple objects is limited only by object spacing, not speed, time, or capacity. Psychological Science, 21, 920-925.

Freeman, J., Pelli, D.G. (2007) An escape from crowding. *Journal of Vision*, **7**, 1-14.

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics*, **67**, 1252–1268.

Gibson, J. J. (1986). The ecological approach to visual perception Lawrence Erlbaum Associates, Hillsdale, NJ (1986) Original work Published 1979

Gibson B. S., Bryant T. A. (2005). Variation in cue duration reveals top-down modulation of involuntary orienting to uninformative symbolic cues. *Psychonomic Bulletin & Review,* **67**, 749-758.

Giordano A. M., McElree B., Carrasco M. (2009). On the automaticity and flexibility of covert attention: A speed-accuracy trade-off analysis. *Journal of Vision,* **10**, 1-10.

Gobell, J. L., Tseng, C., & Sperling, G. (2004). The spatial distribution of visual attention. *Vision Research*, **44**, 1273–1296.

Gould I. C., Wolfgang B. J., Smith P. L. (2007). Spatial uncertainty explains exogenous and endogenous attentional cuing effects in visual signal detection. *Journal of Vision*, **7**, 1-17.

Gottlob, L. R., Cheal, M. L., Lyon, D. R. (1999) Attention operating characteristics in a locationcuing task. *Journal of General Psychology*, **126**, 271–287.

Grindley, G.C. Townsend, v. (1968) Voluntary attention in peripheral vision and its effects on acuity and differential thresholds. *Quarterly Journal of Experimental Psychology*, **20**, 11–19.

Hahn, S., & Kramer, A. (1998). Further evidence for the division of attention among non-contiguous locations. *Visual Cognition*, **5**, 217–256.

He, S., Cavanagh, P., Intriligator, J. (1996). Attentional resolution and the locus of awareness. *Nature*, **383**, 334-338.

He S., Cavanagh P., Intriligator J. (1997). Attentional resolution*.* *Trends in Cognitive Sciences*, **1**, 115–121.

Hein, E., Rolke, H. E., Ulrich, R (2006). Visual attention and temporal discrimination: Differential effects of automatic and voluntary cueing. *Visual Cognition*, **13**, 29–50.

Helmholtz, H. (1867). Handbook of Physiological Optics, Leipzig.

Henderson , J. M. (1991). Stimulus discrimination following covert attentional orienting to an exogenous cue. Journal of Experimental Psychology: Human Perception and Performance, **17**, 91-106.

Hernstein R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, **4**, 267-272.

Herrnstein, R. J., Prelec, D. (1991). Melioration: A theory of distributed choice. *Journal of Economic Perspectives*, **5**, 137-156.

Hoffman J. E., Subramaniam B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, **57**, 787-795.

Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience,* **1**, 304–309.

Hollerman J. R., Tremblay L., Schultz W. (1998). Influence of reward expectation on behaviour-related neuronal activity in primate striatum. *Journal of Neurophysiology,* **80**, 947-963.

Hommel B., Pratt J., Colzato L., & Godijn R. (2001). Symbolic control of visual attention. *Psychological Science*, **12**, 360-365.

Hogg, R. V., Craig, A. T. (1995). Introduction to Mathematical Statistics, 5th ed. New York: Macmillan.

Hsu, M., Krajbich, I., Zhao, C., Camerer, C.F. (2009) Neural response to reward anticipation under risk is nonlinear in probabilities. *Journal of Neuroscience*, **29**, 2231–2237.

Intriligator, J., Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology,* **43**, *171–216.*

James, W. (1890). The principles of psychology. New York: Henry Holt. In, Carrasco, M., (2011) Visual attention: The past 25 years. Vision Research, **51**, 1484-1525.

Johnson D.N., Yantis S. (1995). Allocating visual attention: tests of a two-process model. *Journal of Experimental Psychology: Human Perception and Performance,* **21**, 1376-1390.

Jans, B., Peters, J.C., De Weerd, P. (2010). Visual spatial attention to multiple locations at once: the jury is still out. *Psychological review*, **117**, 637-682.

Joel D., Zohar O., Afek M., Hermesh H., Lerner L., Kuperman R., Gross-Isseroff R., Weizman A., Inzelberg R. *(*2005*)* Impaired procedural learning in obsessive-compulsive disorder and Parkinson's disease, but not in major depressive disorder. *Behavioral Brain Research,* **157**, 253–263.

Jonides J. (1980). Towards a model of the mind's eye's movement. *Canadian Journal of Psychology*, **2**, 103-112.

Jonides J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. B. Long, A. D. Baddeley (Eds.), *Attention and performance IX* 187–203. Hillsdale, NJ: Erlbaum.

Jonides J. (1983). Further toward a model of the mind's eye's movement. *Bulletin of the Psychonomic Society*, **21**, 247-250.

Joseph, J.S., Chun, M.M. and Nakayama, K. (1997). Attentional requirements in a "preattentive" feature search task. *Nature*, **387**, 805-807.

Kahneman , D. (1973). Attention and effort. Englewood Cliffs, N. J.: Prentice-Hall.

Kahneman D., Tversky A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica,* **47**, 263-291.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, **24**, 174–219.

Kastner S., Ungerleider L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience,* **23**, 315–341.

Kawagoe R., Takikawa Y., Hikosaka O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neuroscience,* **1**, 411–416.

Kiefer, R. J., & Siple, P. (1987). Spatial constraints on the voluntary control of attention across visual space. *Canadian Journal of Psychology*, **41**, 474–489.

Kinchla, R.A., Chen, Z., Evert, D.L. (1995) Pre-cue effects in visual search: data or resource limited?

Knill D. C., Richards W. (1996). Perception as Bayesian Inference. Cambridge: Cambridge University Press.

Kobayashi S., Lauwereyns J., Koizumi M., Sakagami M., Hikosaka O. (2002). Influence of reward expectation on visuospatial processing in macaque lateral prefrontal cortex. *Journal Neurophysiology,* **87**, 1488–1498.

Kramer, A. F., & Hahn, S. (1995). Splitting the beam: Distribution of attention over noncontiguous regions of the visual field. *Psychological Science*, **6**, 381–386.

Lauwereyns, J., Watanabe, K., Coe B., Hikosaka O. (2002). A neural correlate of response bias in monkey caudate nucleus. *Nature,* **418**, 413–417. Lawrence, B. M., Myerson, J., Abrams, R. A. (2004). Interference with spatial working memory: An eye movement is more than a shift of attention. *Psychonomic Bulletin & Review*, **11**, 488-494.

Lejarraga, T. (2009). When experience is better than description: Time delays and complexity. *Journal of Behavioral Decision Making*, **23**, 100–116.

Leon M.I., Shadlen M.N. (1999). Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron,* **24**, 415–425.

Ling, S., & Carrasco, M. (2006a). Sustained and transient covert attention enhancethe signal via different contrast response functions. Vision Research, **46**, 1210–1220.

Liu, T., Stevens, S. T., & Carrasco, M. (2007). Comparing the time course and efficacy of spatial and feature-based attention. Vision Research, **47**, 108–113.

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal* *of Neurophysiology,* **77,** 24–42.

Luck, S. J., & Hillyard, S. A. (1990a). Electrophysiological evidence for parallel and serial processing during visual search. *Perception and Psychophysics,* **48**, 603–617.

Luck , S. J. Hillyard , S. A. Mouloua , M. Woldorff , M. G. Clark , V. P. Hawkins , H. L. (1994). Effects of spatial cuing on luminance detectability: Psychophysical and electrophysiological evidence for early selection. *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 887-904.

Luck S. J., Hillyard S. A., Mouloua M., Hawkins H. L. (1996). Mechanisms of visual-spatial attention: Resource allocation or uncertainty reduction? *Journal of Experimental Psychology: Human Perception Performance,* **22**, 725–737.

Luck, S. J., Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature,* **390**, 279-281.

Madden, D. J. (1992). Selective attention and visual search: Revision of an allocation model and application to age differences. *Journal of Experimental Psychology: Human Perception and Performance,* **18**, 821–836.

Martins, A. C. R. (2005) Adaptive Probability Theory: Human Biases as an Adaptation.

Martins, A. (2006). Probability biases as Bayesian inference. *Judgment and Decision Making,* **1**, 108–117. Cogprint preprint at http://cogprints.org/4377/.

Maunsell, J. H. R. (2004) Neuronal representations of cognitive state: reward or attention? *Trends in Cognitive Sciences,* **8**, 261-265.

Maunsell, J. H., & Treue, S. (2006). Feature-based attention in visual cortex. Trends in Neuroscience, **29**, 317–322.

McCormick, P. A., Klein, R. M., & Johnston, S. (1998). Splitting versus sharing focal attention: Comment on Castiello and Umiltà (1992). *Journal of Experimental*

Psychology: Human Perception and Performance, **24**, 350–357.

Miller, J. (1988). Components of the location probability effect in visual search tasks. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 453–471.

Mirenowicz, J., Schultz, W. (1994). Importance of unpredictability for reward responses in primate dopamine neurons. *Journal of Neurophysiology*, **72**, 1024–1027.

Moray, N. (1967) Where is capacity limited? A survey and a model. In A. Sanders(.Ed), Attention and Performance, North-Holland, Amsterdam (1967)

Mordkoff, J. T., Egeth, H. E. (1993). Response time and accuracy revisited: Converging support for the interactive race model. *Journal of Experimental Psychology: Human Perception and Performance*, **19**, 981–991.

Morgan, M.J., Ward, R.M., Castet, E. (1998). Visual search for a tilted target: Tests of spatial uncertainty models. The Quarterly Journal of Experimental Psychology, **51**, 347–370.

McFadden, D. (1980). "Econometric Models of Probabilistic Choice Among Products", *Journal of Business.* **53**, 13–29.

Morse, E. B., Runquist, W. N. (1960) Probability-matching with an unscheduled random sequence. *American Journal of Psychology,* **73**, 603–7.

Muller, H. J., & Rabbitt, P. M. (1989a). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. Journal of Experimental Psychology: Human Perception and Performance, **15**, 315–330.

Nelder J. A., Mead R. (1965). Simplex method for function minimization. *The Computer Journal,* **7**, 308.

Newsome, W.T. Britten, K.R. Movshon, J.A. Shadlen, M. (1989) Single neurons and perception of visual motion D.M.-K. Lam, C.D. Gilbert (Eds.), Neural Mechanisms of

Visual Perception, Portfolio Publishing Company, The Woodlands, TX Proceedings of the Retina Research Foundation, Volume 2.

O'Doherty J., Kringelbach M. L., Rolls E. T., Hornak J., Andrews C. (2004). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience,* **4**, 95–102.

Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition*, **11**, 631–671.

Palmer, J., Ames, C.T., Lindsey, D.T. (1993) Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, **19**, 108–130

Palmer J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research*, **34**, 1703–1721.

Palmer, J., Verghese, P., Pavel, M. (2000) The psychophysics of visual search, *Vision Research,* **40**, 1227–1268.

Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **41**, 19–45.

Pavlov, I.P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. London: Oxford University Press.

Pelli, D. G. (1997) The Video Toolbox software for visual psycho-physics. Spatial Vision, **10**, 437-442.

Pestilli, F., & Carrasco, M. (2005). Attention enhances contrast sensitivity at cued and impairs it at uncued locations. Vision Research, **45**, 1867–1875.

Platt M.L., Glimcher, P.W., (1997) Responses of intraparietal neurons to saccadic targets and visual distractors. *Journal Neurophysiology*, **78**, 1574–1589.

Platt M.L., Glimcher P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature,* **400**, 233–238.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, **3**, 179–197.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology,* **32**, 3–25.

Posner, M. I., Nissen, M. J., Ogden, W. C. (1978). “Attended and unattended processing modes: The role of set for spatial location”. In *Modes of Perceiving and Processing Information*, Edited by: Pick, H. L..

Posner, M.I., Snyder, C.R.R. (1975) Attention and cognitive control. In R. Solso (Ed.), Information processing and cognition: The Loyola symposium. Hillsdale, N.J.: Erlbaum, 1975.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980) Attention and the detection of signals. *Journal of Experimental Psychology*, **109**, 160-174.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D.Bonwhuis (Eds.), Attention and performance X: Control of language processes(pp. 551–556). Hillsdale, NJ:Erlbaum.

Preston M. G., Baratta P. (1948). An experimental study of the auction-value of an uncertain outcome. *American Journal of Psychology*, **61**, 183-193.

Prinzmetal W., McCool C., Park S. (2005). Attention: Reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology: General*, **134**, 73-92.

Rabbitt, P., Cumming, G., & Vyas, S. (1977). An analysis of visual search: Entropy and sequential effects. In S. Dornic (Ed.), *Attention and performance VI* (pp. 363–386). Hillsdale, NJ: Erlbaum.

Rabbitt, P., Cumming, G., & Vyas, S. (1979). Modulation of selective attention by sequential effects in visual-search tasks. *Quarterly Journal of Experimental Psychology*, **31**, 305–317.

Ramnani N., Miall C. (2003). Instructed delay activity in the human prefrontal cortex is modulated by monetary reward expectation. *Cerebral Cortex*, **13**, 318-327.

Ratcliff, R., van Zandt, T., McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychologival Review,* **106**, 261–300.

Ratcliff, R., Carpenter, R.H.S., Reddi, B.A.J. (2001). Putting noise into neurophysiological models of simple decision making. *Nature Neuroscience,* **4**, 336–337.

Restle , F. (1961). Statistical methods for a theory of cue learning. Psychometrika, **26**, 291-306.

Remington R. W., Folk C. L. (2010). A dissociation between attention and selection. *Psychological Science*, **12**, 511-515.

Rensink, R.A. *(2000).* Seeing, sensing, and scrutinizing. *Vision Research,* **40,**1469-1487.

Reynolds, J.H., Chelazzi, L., Desimone, R. (1999) competitive mechanisms subserve attention in macaque areas v2 and v4. *Journal of Neuroscience*. **19**, 1736-1753.

Risko E. F., Blais C., Stolz J. A., Besner D. (2008). Covert orienting: A compound-cue account of the proportion cued effect. *Psychonomic Bulletin & Review*, **15**, 123-127.

Ristic J., Friesen C. K., Kingstone A. (2002). Are the eyes special? It depends on how you look at it. *Psychonomic Bulletin & Review,* **9**, 507-513.

Riggio L., Kirsner K. (1997). The relationship between central cues and peripheral cues in covert visual attention. Perception and Psychophysics, **59**, 885–899.

Romo R, Schultz W (1990) Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *Journal of Neurophysiology,* **63**, 592–606.

Santee, J. L., Egeth, H. E. (1982). Do reaction time and accuracy measure the same aspects of letter recognition? *Journal of Experimental Psychology: Human Perception and Performance*, **8**, 489–501.

Sapir A., d'Avossa G., McAvoy M., Shulman G.I., Corbetta M. (2005). BOLD signals for spatial attention predict performance in a motion discrimination task. *Proceedings of the National Academy of Sciences USA*, **102**, 17810-17815.

Schönberg T., Daw N. D., Joel D., O'Doherty J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *Journal of Neuroscience,* **27***,* 12860–12867.

Scolari, M., Kohnen, A., Barton, B., Awh, E. (2007). Spatial attention, preview, and popout: Which factors influence critical spacing in crowded displays? *Journal of Vision,* **7**, 1–23.

Sekuler, R., Ball, K. (1977). Mental set alters visibility of moving targets. *Science,* **198**, 60-62.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal,* **27**, 379-423 & 623-656.

Shaw, M. L. (1978). A capacity allocation model for reaction time. *Journal of Experimental Psychology: Human Perception & Performance*,**4**, 586–598.

Shaw, M. L. (1984).Division of attention among spatial locations: A fundamental difference between detection of letters and detection of luminance increments. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance,* 109–121 Hillsdale, NJ: Erlbaum.

Shaw, M. L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception & Performance,* **3**, 201–211.

Shiffrin, R. M., Gardner, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology*, **93**, 72-83.

Shiu, L., Pashler, H. (1994). Negligible effect of spatial precuing on identification of single digits. *Journal of Experimental Psychology: Human Perception and Performance,* **20**, 1037–1054.

Simon , H. A. (1959). Theories of decision-making in economics and behavioral science. American Economic Review, 1959, **49**, 253-283.

Smyth, M. M. and Scholey, K. A. (1994). Interference in spatial immediate memory. *Memory and Cognition*, **22**, 1–13.

Sokolov, Y.N. (1963) Perception and the conditioned reflex, Macmillan, New York (English translation, S. W. Waydenfeld, Trans.)

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, **11**, 1-29.

Sperling, G., Dosher, B. (1986). Strategy and optimization in human information processing. In Boff K., Kaufman L., Thomas J. (Eds.), *Handbook of perception and performance***, 1–85** New York: Wiley.

Sperling G., Melchner, M. J. (1978). The attention operating characteristic: Examples from visual search. *Science,* **202**, 315-318.

Sperling, G. A., Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review*, **102**, 503–532.

West, R.F., Stanovich K. E. (2003) Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Memory & Cognition,* **31**, 243–251.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology, **18**, 43–662.

Sugrue L.P, Corrado G.S, Newsome W.T. (2005) Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nature Reviews in Neuroscience*, **6**, 363–375.

Sylvester C. M., Shulman G. L., Jack A. I., Corbetta M. (2007). Asymmetry of anticipatory activity in visual cortex predicts the locus of attention and perception. *Journal of Neuroscience*, **27**, 14424–14433.

Sylvester C. M., Jack A. I., Corbetta M., Shulman G. L. (2008). Anticipatory suppression of non attended locations in visual cortex marks target location and predicts perception. *Journal of Neuroscience*, **28**, 6549-6556.

Talgar C. P., Carrasco M. (2002). Vertical meridian asymmetry in spatial resolution: Visual and attentional factors. *Psychonomic Bulletin & Review,* **9**, 714–722.

Tanner W. P., Swets, J.A. (1954) A decision-making theory of visual detection, *Psychological Review*, **61**, 401-409.

Theeuwes, J., Kramer, A. F., Irwin, D. E. (2011) Attention on our mind: The role of spatial attention in visual working memory. *Acta Psychologica,* **137**, 248-251.

Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review*, **9**, 314-318.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, **12**, 97–136.

Treisman, A., Gormican, S. (1988) Feature analysis in early vision: evidence form search asymmetries, *Psychological Review,* **95**, 15–48.

Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychological Science*, **20**, 473–479.

Van der Heijden, A. H. C. (1989). Probability matching in visual selective attention. *Canadian Journal of Psychology*, **43**, 45-52.

Van der Heijden , A. H. C., Brouwer , R. F. T., Serlie , A. W. (1992). Single-letter recognition accuracy and position information: Benefits with a symbolic cue. Psychological Review, **54**, 182-186.

Verghese, P. (2001) Visual search and attention: a signal detection theory approach. *Neuron*, **31**, 523-535.

Vossel S, Thiel CM, Fink GR (2006). Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. *Neuroimage* ,**32**, 1257–1264.

Vulkan, N. (2000) An economists perspective on probability matching. *Journal of Economic Surveys*, **14**, 101-118.

Tobler P.N., Christopoulos G.I., O'Doherty J.P., Dolan R.J., Schultz W. (2008). Neuronal distortions of reward probability without choice. *Journal of Neuroscience*, **28**, 11703-11711.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **12***,* 97–136.

Treisman , A. Souther , J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, **114**, 285-310.

van der Heijden, A. H. C. (1989). Probability matching in visual selective attention. *Canadian Journal of Psychology*, **43**, 45-52.

Wickens , C. D. (1984). Engineering psychology and human performance. Columbus, OH: Charles Merrill.

VanRullen R., Carlson T., Cavanagh P.(2007). The blinking spotlight of attention. *Proceedings of the National Academy of Sciences USA*, **104**, 19204-19209.

Vincent, B. (2011) Covert visual search: prior beliefs are optimally combined with sensory evidence, *Journal of Vision*, **11**, 1-15.

Vincent, B. (2011) Search asymmetries: parallel processing of uncertain sensory information, *Vision Research*, **51**, 1741-1750.

Vincent, B. T. (2012) How do we use the past to predict the future in oculomotor search, *Vision Research*, **74**, 93-101.

Von Wright J. M. (1970). On selection in visual immediate memory. In A. F. Sanders (Ed.), Attention and performance (Vol . III, pp. 280– 292). *Amsterdam: North Holland.*

Vossel S., Thiel C.M., Fink G.R. (2006). Cue validity modulates the neural correlates of covert endogenous orienting of attention in parietal and frontal cortex. *Neuroimage*, **32**, 1257-1264.

Walthew, C., & Gilchrist, I. D. (2006). Target location probability effects in visual search: An effect of sequential dependencies. *Journal of Experimental Psychology: Human Perception and Performance*, **32**, 1294–1301.

Weber, E. U., Shafir, S., Blais, A.R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review*, **111**, 430–445.

Williams, B. A. (1988). Reinforcement, choice and response strength. In R. C. Atkinson, R. J. Herrnstein, G. L. Lindzey, & R. D. Luce (Eds.),*Stevens' Handbook of experimental psychology* (2nd ed, 167–244). New York: Wiley.

Wright, R. D. (1994). Shifts of visual attention to multiple simultaneous location cues. *Canadian Journal of Experimental Psychology*, **48**, 205–217.

Yantis, S. (1992). Multielementvisual tracking: Attention and perceptual organization. *Cognitive psychology*, **24**, 295–340

Yantis S. (1996). Attentional capture in vision. *In Kramer A. F., Coles M.G.H., Logan G.D. (Eds.) Converging operations in the study of visual selective attention, 45-76. Washington, DC: American Psychological Association.*

Yantis, S., & Jonides, J. (1996). Attentional capture by abrupt onsets: New perceptual objects or visual masking? Journal of Experimental Psychology: Human Perception and Performance, **22**, 1505–1513.

Yantis S. (1998). Control of visual attention. *In Pashler H. (Eds.), Attention, 223-256. Hove, U.K.: Psychology Press/Erlbaum.*

Yeshurun, Y., & Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. Journal of Vision, **10**,16.

Yu A. J., Dayan P. (2005). Uncertainty, neuromodulation, and attention*. Neuron,* **46**, 681–692.

Zhang, H., Xuan, Y., Fu, X., & Pylyshyn, Z. W. (2010). Do objects in working memory compete with objects in perception? *Visual Cognition,* **18***,* 617–640.