Title

The Role of Predictive Uncertainty in Attentional Processing

Author

Witkowski, Phillip

Publication Date

2023

The Role of Predictive Uncertainty in Attentional Processing

By

PHILLIP WITKOWSKI
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Psychology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Joy J. Geng

Erie D. Boorman

Charan Ranganath

Committee in Charge

2023

Acknowledgements

I would chiefly like to express my gratitude and admiration for those who guided me on my journey to becoming a scientist. First among these are my advisors, Dr. Joy Geng and Dr. Erie Boorman, who played an enormous role in my development from a fledgling graduate student into a full independent scientist. Your guidance and support are invaluable to me, and I will always be grateful for the wisdom you have shared – about both academia and life. Equally, I would like to thank Dr. Seongmin (Alex) Park for being an incredible lab-mate, friend, and in many ways a third advisor. You were always available for questions and patient in teaching me your answers. Graduate school would not have been a success without your help. I would also like to thank the final member of my committee, Charan Ranganath, for his time and support.

In addition to my advisors, I would like to thank all the members of the Geng and Boorman labs for their help through the years. In particular, I would like to thank Xinger Yu and Zhiheng Zhou, as well as Lindsay L. Rondot for helping me collect the neuroimaging data presented in Chapter 3. Thank you to my friends in the 2017 psychology cohort who were the source of many imaginative and stimulating conversations, supported me through life’s trials and tribulations, and made my academic journey a truly enriching and memorable experience.

Finally, I would like to express my gratitude to my wonderful family. I would like to give special thanks to my wife, Clara Witkowski. You have given me constant support and encouragement throughout my graduate school journey, and I truly would not have made it this far without you. I am eternally grateful for the countless hours you spent listening to me talk about science, helping me study for exams, and proofreading my papers. Perhaps most of all, I’m grateful for your patience for all the times I trailed off in thought mid-conversation. You are my deepest love.

Abstract

The concept of target templates is a core construct in all models of attention, guiding attention during search and acting as a “template” for making target match decisions. Previous work has generally focused on how the identity of a known target is stored in the target template and the consequences of failures to maintain that representation in memory. Few, if any, have tested how the target template adapts to uncertainty in target features and makes predictions about how the target will look in the future. In this doctoral thesis, I present a series of four experiments that address how predictive uncertainty shapes representations of target templates, on both cognitive and neural levels. I show that observers readily learn statistical knowledge about targets and use this knowledge to make predictions about the likelihood of various target features. These predictions are then used to set priority for features which minimize uncertainty about the upcoming target, biasing eye-movements and reducing decision times when participants are certain about how the target will look. At the neural level, I show that feature uncertainty is coded in dorsolateral prefrontal cortex and inferior frontal junction, supporting search behavior for uncertain targets. These findings contribute to our understanding of how attentional templates adapt to statistical knowledge about uncertain target features and provide insights into theories of attention and perceptual learning.

Table of Contents

Chapter 1: Introduction

The idea of the target template is central to almost all theories of top-down attention and traditionally defined as the information held in memory which guides search (Chelazzi et al. 1998; Desimone and Duncan 1995; Duncan and Humphreys 1989; Treue and Trujillo 1999). The contents of the template are used to guide attention by modulating priority signals that bias the sensory processing of task-relevant information (Driver, 2001; Geng and Witkowski, 2019; Noah et al., 2022). For example, a target defined by its red color will result in enhanced processing of red items in the visual field. These priority signals help segment the visual scene and guide eye-movements towards likely putative targets. The target template is also the basis for critical target-match decisions during search, serving as the "template" to which incoming stimuli are compared (Geng and Witkowski, 2019; Wolfe, 2020).

Previous research has found costs in visual search performance when templates are imprecisely specified or fail to predict the target accurately (Hout and Goldinger, 2015; Malcolm and Henderson, 2009; Schmidt and Zelinsky, 2009; Yang and Zelinsky, 2009). For example, Hout and Goldinger (2015) cued subjects to a subsequent target with an exact visual copy of the object, the same object in a different state (e.g., a different configuration or viewpoint), or a different exemplar from the same category. They found that targets which were more dissimilar from the cue produced longer scan paths to the target and increased the time for target-match decisions once the target was fixated. Similarly, electrophysiological work has shown that ERP indexes of attentional selection have a larger amplitude when cued with the specific identity (e.g., the specific letter) rather than the target category (e.g., any letter), suggesting the knowing the specific identity leads to more effective search (Nako et al., 2014). Taken together, these studies argue that visual search will be more efficient when the target is specified with greater precision.

However, many of these studies have overlooked a key property of targets in real world search. Targets in the real world are dynamic. The features of objects can shift over time, creating uncertainty about the target's features and limits in how effective a precise template can be. An overly precise target template will frequently fail to capture target accurately if the features change, making search more difficult. Given this uncertainty, observers in the real world must make predictions about how the target *will* look during search, rather than how it once looked before. Further, knowing which features are likely remain consistent and prioritizing reliable information is pertinent finding objects which may appear differently each time you look for it. For example, your friend may change clothes from day to day, but knowing that your friend has a penchant for wearing Hawaiian shirts may help you to identify her in a cluttered sensory environment such as a park. Little is known about how target templates compensate for feature uncertainty, despite being critical for defining task-relevance in real world settings (Gottlieb, 2018). Understanding how the target template responds to uncertainty can inform theories of attention by showing how attentional mechanisms operate in dynamic sensory environments.

Are target templates responsive to the statistics of the environment?

If the target template defines what information is needed to identify search targets, then such a representation should be responsive to the underlying structure the search environment. A growing body of evidence suggests that contents of the target template are highly flexible, and change in accordance with task demands, following efficient coding principles with respect to how target can be identified and in what context (Geng and Witkowski, 2019; Yu et al., 2023). Categorically defined templates have been found to aid search when task demands require localizing multiple targets from the same general category, but not when specific a target must be identified (Bravo and Farid 2012, 2016). Further, in a study by Schmidt and Zelinsky (2017), the authors asked subjects to find a teddy bear target image embedded within non-teddy bear objects or different teddy bears. The number of items held in working memory during the delay, measured by contra-lateral delay activity, were greater during blocks with other teddy bears. Presumably, this is because more details are needed to differentiate intra-category items, suggesting that observers optimize the number of details within a target template with response the distractor environment.

In addition to modulating the quantity of details represented in the template, the quality of information in the template may also be constructed with respect to the distractor's observers are likely to encounter (Geng et al., 2017; Navalpakkam and Itti, 2007; Yu and Geng, 2019). In one experiment (Navalpakkam & Itti, 2007), participants were asked to identify the location of orientation targets of 55-degree tilt in two separate search arrays. In the first set of search arrays, the participants identified the target among 25 homogenous distractor objects (called T1 trials). In the second, participants reported the location of the orientation target among 8 heterogenous distractors distributed around the target (T2 trials). The key manipulation was how the distractors were oriented relative to the target. If distractors in T1 trials were consistently less tilted than the target this led subjects to report the target as the object with an orientation of 60 degrees in T2 search trials. If distractors had greater tilt in T1 trials, they were likely to report the target as a 50-degree oriented object in T2 trials. This study reveals the target template need not encode veridical representations of the target but can optimize and shift the representation of target information based on the known statistics of distractors (Geng and Witkowski, 2019; Yu et al., 2023)

Together these studies exemplify the flexibility and responsiveness of target templates to statistical knowledge about distractors and the demands of the task at hand. However, they tell us very little about if and how the contents of the target template respond to the statistics of the target itself. This is a critical gap in theories of attention because they rely on perfect knowledge of target properties and how they relate to the broader environment. Thus, while we know that the template responds to statistical knowledge about non-targets, we do not yet know how the target template makes judgements about feature relevance when the feature of the target are uncertain.

Defining the contents of the template as Information Decisions

A key question for theories of the target template is how attentional systems make judgements about what is “task-relevant”. One particularly powerful framework called active-sampling casts top-down attention as an information seeking mechanism that proactively makes “requests" for information from sources that are thought to be predictive for a given task (Gottlieb, 2018). For example, when coming to a street crossing you must decide whether it is safe to cross, and therefore must decide whether to get relevant information from in the traffic light, the clouds, or any number of other possible stimuli. In this way, attention is inherently tied to an understanding of the underlying task-structure (Gottlieb et al. 2020; S. C.-H. Yang, Lengyel, and Wolpert 2016) and critically, the value of the information source (Dayan and Yu 2003; Foley et al. 2017; Gottlieb and Balan 2010; Leong et al. 2017).

In the active sampling framework, the observer should always sample from the most informative sources. One factor that contributes to the value of an information source is its expected uncertainty, this is, how variable the information is each time its sampled (Gottlieb, and Balan, 2012; Gottlieb, 2018). Evidence that spatial sampling decisions are tied to uncertainty of information sources comes from Foley and colleagues (2016), who trained monkeys to complete a two-step task in which they used knowledge about the value of information to make directed saccades to a rewarded target. In each trial, monkeys were initially presented with two dot patterns surrounded by a colored circle indicating the “information value” of this cue - the probability that it would correctly indicate the direction of a rewarded target location. The results indicated that neurons of the lateral intraparietal sulcus showed robust saccade-related responses that scaled as a function of the difference in expected information gain between the two cues. This directly connects eye-movements and information sampling to uncertainty and characterize attention as “decisions” about which sources of information would be best to discriminate given the desired goal. However, it's still unclear whether similar mechanisms, or similar information sources, can exist in tasks that rely on feature-based processing such as visual search.

One possibility is that the visual system draws on statistical knowledge about the environment to compute the reliability of a particular feature dimension. Expectations about the sensory environment reflect previously learned knowledge about the likelihood of occurrence for incoming sensory data, which serve to modulate the processing of expected events (Clark 2013; Summerfield and Egner 2009, 2016). Recent work has shown that expectations rely on much of the same neural machinery as top-down visual attention (Iglesias et al. 2013; Kok et al. 2013; Den Ouden, Kok, and de Lange 2012). Iglesias and colleagues found that a network of sensory and frontal cortical regions including dorsolateral prefrontal cortex (DLPFC), coded precision-weighted prediction errors (pwPE) for incoming visual stimuli, suggesting that these regions serve predictive computations that represent both the predicted stimulus and the certainty of those predictions. Others have found evidence for feature specific pwPE’s in feature selective visual cortex (Stefanics et al. 2018; Stefanics, Stephan, and Heinzle 2019), suggesting that these expectations may be coded separately for independent visual dimensions. The certainty of these visual predictions could be utilized to make attentional decisions based on which features are likely to be the most information in the upcoming task. In this way, the target template might be thought of as a predictive system that draws on learned expectations to bias attentional priority for information that more certain to be identifying characteristics of targets (Gottlieb and Balan, 2010; Nobre and Stokes, 2019).

Neural Mechanisms of Integrating Uncertainty into the Target Template

Essential to understating the effect of uncertainty on attentional processing is understanding how the brain integrates estimates of uncertainty into representations of the target template.

Regions of the lateral prefrontal cortex, including DLPFC and IFJ have long been thought to be a source for maintaining selected information in protected states (Buschman, 2021; Buschman and Miller, 2007; Panichello and Buschman, 2021) and providing top-down modulation of sensory gain (Giesbrecht et al., 2003; Meyyappan et al., 2021). The Inferior frontal junction (IFJ) in particular is a likely source of feature based attention (Meyyappan et al., 2021; Zanto et al., 2010; Zhang et al., 2018), and maintains codes for target feature information throughout search (Bichot et al., 2015). Together, these studies support the idea that areas of prefrontal cortex such as IFJ and DLPFC are critical to the construction of feature-based attention templates.

More elusive is how the brain estimates the uncertainty of sensory information and integrates these estimates into target template representations. Hints have come from studies of visual working memory which is often linked with target template representations. Recent studies show that working memory representations in visual and prefrontal cortex encode both the item and uncertainty in probabilistic population codes. In a study by Li et al. (2022), participants were asked to remember and report spatial positions, and identified regions in primary visual cortex and prefrontal cortex which coded both the location and its uncertainty. This multivariate neural code corresponded to participant reports of confidence in their memory, suggesting joint encoding of location and uncertainty in the same representation.

An alternative to this hypothesis comes from the neuroscience of decision making. Previous work suggests that univariate codes in the lateral prefrontal cortex track total uncertainty in decision-making tasks, with higher activation in cases of greater uncertainty. In one study, DLPFC and IFJ activation were higher when the category of orientation stimuli was ambiguous (Summerfield et al., 2011), and in another study DLPFC was found to signal prior knowledge of possible outcomes when decisions were uncertain (Hansen et al., 2012). These studies indicate that lateral prefrontal cortex represents prior knowledge and the uncertainty of upcoming decisions in a univariate code. This is an attractive alternative because these uncertainty codes seem to be domain general and are found in the same prefrontal regions thought to be encoded by the target template. Dissociating the possible neural mechanisms for how predictive uncertainty is coded in the brain prior to visual search is an essential step in understanding how uncertainty is integrated into the target template at the neural level.

Aim of the current work

The aim of the current work is to understand how uncertainty changes the encoding of visual information in the target template. First, we test how target variability biases template representations and changes the encoding of information prior to search, elaborating the role of uncertainty in information selection and attentional biasing for features during visual search. In chapter 2, we explore how variability is learned over time and used to flexibly update priority for feature dimensions as the variability of features changes. We expect that while the precision features encoded into working memory will reflect the priority status of each feature, only the subsequent predictions about target features would be adapted to the specific variability of features. This would directly connect priority setting to predictive processes which account for the inherent uncertainty of dynamic target objects.

Our second aim is to identify the neural mechanisms that encode expected variability for task-relevant features. In chapter 4, we use functional imaging to identify regions of the brain that encode expected variability for visual features and explore how this information is integrated into representations of target templates. We expect that DLPFC, IFJ and visual cortex will carry target representations along with the expected variability of features prior to search. This would give insight into how the brain constructs representations of uncertain targets and will illuminate the neural mechanisms of feature-based attentional processing for dynamic target objects.

Chapter 2: Learned feature variance is encoded in the target template and drives visual search

Introduction

Visual attention helps us identify objects that are relevant for goal-directed behavior, but the world we live in is highly dynamic and the features of target objects can vary over time. For example, imagine going to meet your friend in a park on a summer afternoon. As you search, you may recall the way she looked the last time you saw her, but this will not perfectly predict her appearance today. She may be wearing different clothes, have a different hairstyle or accessories, and you may not know what she will be doing (e.g., walking or sitting). This means that your attentional template, or the representation of the target held in memory (Desimone and Duncan, 1995; Duncan and Humphreys, 1989; Geng and Witkowski, 2019; Olivers et al., 2006), may be imprecise, impairing your ability to locate your friend. The attentional system must deal with this target variability during visual search in natural environments, but how it does so is poorly understood.

Here, we examine how top-down knowledge about the variance of feature dimensions might change the construction of templates to ameliorate the effects of uncertainty. For example, as you search for your friend you may remember that some "dimensions" of your friend are less variable than others and knowing which dimensions are reliable over time can help you find her. If you know your friend normally wears bright red glasses, but is constantly changing her style of clothes, you can focus on information about your friend that is likely to be similar to the last time you saw her (i.e. the glasses) and down-weight information that is likely to change (i.e. her clothes) (Olivers et al., 2011).

Previous research has found costs in visual search performance when templates are less precisely specified or fail to exactly match the target (Malcolm and Henderson, 2010; Schmidt and Zelinsky, 2009; Yang and Zelinsky, 2009). For example, Hout & Goldinger (2014) cued subjects to a subsequent target with an exact visual copy of the object, the same object in a different state (e.g., configuration or viewpoint), or a different exemplar from the same category. They found that cues that were visually more dissimilar from the target produced longer response times, scan paths to find the target, and dwell time on the target (see also, Malcolm & Henderson, 2009; Nako, Wu, Smith, & Eimer, 2014; Reeder & Peelen, 2013).

This is presumed to happen because the target information held within the template is used to set sensory gain and decide if an object is the target (Chelazzi et al., 1993, 1998; Desimone and Duncan, 1995; Kastner and Ungerleider, 2000; Moore and Fallah, 2001). For example, when searching for a red colored object, mechanisms of gain enhancement will increase the response of sensory neurons encoding “red”, resulting in increased firing rate, larger amplitude ERPs, and increased BOLD activation to all red stimuli in the visual field (Andersen et al., 2008; Giesbrecht et al., 2003; Liu et al., 2007; Ruff et al., 2006; Serences et al., 2009; Treue and Trujillo, 1999; Zhang and Luck, 2009). These changes lead to better target selection. However, when the template-to-target correspondence is poor, then gain enahncement will be inaccurate and lead to longer search and decision times.

While there is good evidence for why there are performance costs when the template-to-target match is inaccurate, it remains unclear how the attentional system copes when there is uncertainty about target features. How are priority signals set when the target features themselves are variable? Dimension weighting provides evidence that it is possible to selectively enhance specific feature dimensions (e.g. color, orientaion, size), even when the exact features of the target are unknown (Found and Müller, 1996; Müller and Krummenacher, 2006; Müller et al., 1995, 2003; Pollmann et al., 2006; Reeder et al., 2017). These studies find that information from an entire dimension can be "weighted" if that dimension reliably distinguishes the target from distractors. For example, in Müller et al. (1995) subjects searched for a singleton target in one dimension (e.g., color) amongst distractors from another dimension (e.g., orientation, size). When the dimension that distinguished targets from distractors could be predicted, subjects were equally good at finding the target as control subjects who knew the exact target feature. Similar to single feature enhancement, dimension weighting appears to increase sensory gain for task-relevant dimensions (Gramann et al., 2007; Pollmann et al., 2006; Reeder et al., 2017; Töllner et al., 2010; Wei et al., 2018).

Evidence from visual search (Kerzel and Witzel, 2019) and retro-cueing paradigms (Hollingworth and Hwang, 2013; Niklaus et al., 2017; Rajsic and Woodman, 2019; Rajsic et al., 2017) have shown that task-relevent dimensions are also represented with greater precision compared to non-relevant dimensions in working memory. For example, in Niklaus, Nobre, & Van Ede (2017), showed participants an array of three colored arrows that were to be held in memory until subjects were probed to report the color or orientation of one arrow. During the interval between viewing the array and the probe response, subjects were cued with 75% validity as to which dimension would be probed. Critically, the cue did not indicate a specific arrow, and consequently, subjects did not know exactly which feature value would be probed. The results showed greater precision for the retro-cued dimension compared to the un-cued dimension. This demonstrates that multiple features within a dimension can be sharpend in working memory when they are believed to be task-relevant.

In the current studies, we examine how template specificity is shaped by top-down knowledge of the variance in different feature dimensions. To test this hypothesis, we use a novel paradigm in which target feature values change between the target cue and the actual target within the visual search display. Importantly, the changes in the two feature dimensions are governed by distributions such that the variance of one dimension is smaller than the other. We hypothesized that the expected variance of each feature would be 1) reflected in the contents of the target template because it is task-relevant and 2) be used to weight attentional priority towards the feature with less variance during visual search. Further, we directly compared the effects of template-to-target similarity on visual search performance against other known metrics of search difficulty including target-distractor similarity and distractor-distractor similarity (Chetverikov, Campana, & Kristjánsson, 2017; Duncan & Humphreys, 1989).

Experiment 1

The purpose of this first experiment was to test the hypothesis that the target template will reflect task-relevant expectations of the cue-to-target feature variance and predict the effect of each feature on visual search. We do this by using an experimental procedure in which participants are shown a cue stimulus on every trial composed of color and motion. The cue is followed by one of two tasks, interleaved across trials: One is a probe task that assesses the memory precision of each of the two cued features using color and direction wheels; the second is a visual search task to measure how cue-to-target variance in each feature dimensions affects RT and decision times. Use of the two tasks is critical to acquiring separate measurements of the contents of the target template created from the cue stimulus, and the use of that information during visual search with distractor competition.

Methods

Participants

Forty students (females = 28, mean age = 20.35 years) from the University of California, Davis were recruited to participate in the study. This sample size was calculated from pilot data with G*power (http://www.gpower.hhu.de/) showing that it would give us at least 80% power in an ANOVA comparing template precision, at a partial eta-squared of .015. All participants had normal or corrected-to-normal vision, normal color vision, and received course credit for participating in the experiment. Thirty-Nine subjects were right-handed.

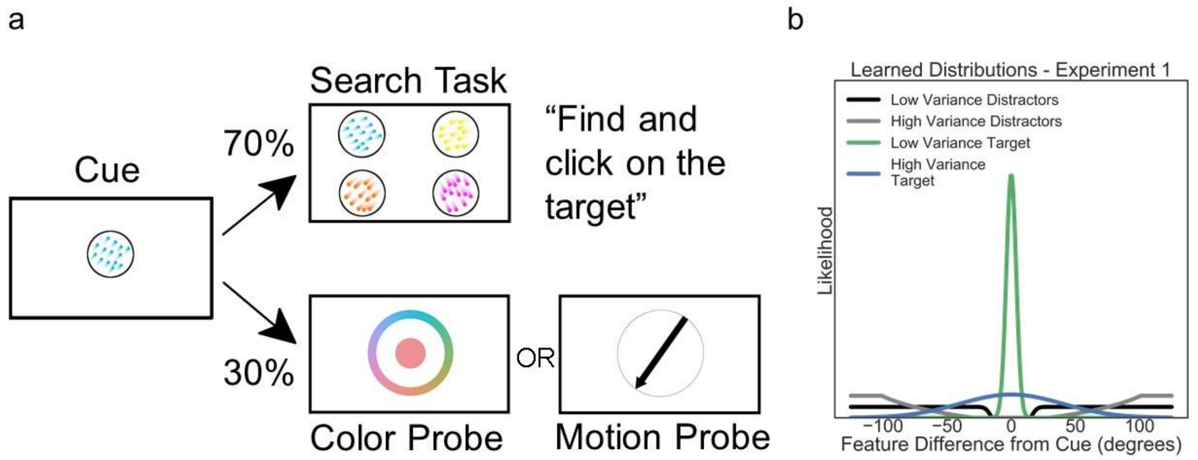

Figure 2.1: a) Schematic of task. Subjects were given a cue object at the beginning of each trial defined by a randomly selected color and direction of motion. On 70% of trials, this was followed by a visual search display. On 30% of trials, they were asked to report the remembered color or motion of the cue object. b) Distributions of target and distractor features for Experiment 1. Zero represents no difference from the cue. Half of the subjects saw motion follow the "low-variance target" distribution and color follow the "high-variance target" distribution. The other half of subjects experienced the reverse.

Procedure

Trials began with the presentation of a centrally presented cue, a random-dot-kinematic (RDK) of colored dots moving in a specific direction with 100% coherence, for 1000ms (Fig

2.1a). This was followed by a jittered ISI between 2000-3000ms and then one of two tasks. On 70% of trials, the cue was followed by a visual search display in which subjects we instructed to report the location of the target that most closely matched the cue by clicking on it with a mouse. Subjects were told to respond as rapidly as possible without sacrificing accuracy. The stimuli were RDKs viewed within an outline of a circle aperture 3.5° of visual angle in radius. Each RDK was 3.88° of visual angle from the central fixation cross on both the horizontal and vertical axis, and one stimulus was presented in each quadrant of the screen. All RDKs disappeared after 1000ms, but there was no limit on the time to respond. The circle outline bounding the RDKs remained visible on screen until the response.

Critically, on visual search trials, the target features changed from the cue systematically, creating variance in cue-to-target similarity between trials. Cue-to-target similarity was defined as the difference between the feature values of the cue and the target for each dimension independently. Each target feature was drawn from one of two Gaussian distributions, that were centered at zero **(**Fig 2.1b**)**. However, the low-variance distribution had a standard deviation of 3.8^o, bounded at -15 to 15 degrees from the cue. The high-variance distribution had a standard deviation of 38 degrees, bounded at -70 and 70 degrees from the cue. Half of all subjects had motion as the low-variance dimension and color as the high-variance dimension (i.e., the "motion" group), while the other half had color as the low-variance dimension and motion as the high-variance dimension (i.e., the "color" group). Distractors were pulled from a uniform distribution between -135 and 135 degrees, although all distractors were required to be at least 30 degrees further from the cue than the target feature in the same dimension. This was done so that the target features in each dimension were closer to the cue than any another object within the visual search array. Thus, either dimension could be used to find the target, but the task-relevance of each dimension was defined by its expected variability because weighting the less-variable dimension could lead to more efficient target selection.

On the other 30% of trials, the cue stimulus was followed by a color wheel or a motion-direction wheel on which subjects indicated the remembered color or motion-direction of the cue. These trials served as a direct measurement of the template representation of the two cued target features in working memory.

The experiment progressed in two parts. The first part consisted of 220 training trials, in which subjects were “trained” on the distributions through repeated exposure to search displays. Probe Trials were included here to prevent subjects from noticing any changes when the test session began. This was followed by 220 test trials, which were identical to the training trials, but now included “switch” trials on 15% of trials. On switch trials, each target feature was drawn from the opposite distribution (i.e., the high-variance feature was now the low-variance feature). This increased the observed variance of the low-variance distribution and decreased the observed variance of the high-variance dimension, allowing us to measure the effect of unexpected cue-to-target similarity values. Inclusion of these trials did not change the center of the distributions (both remained at zero) but increased the overall observed variance of the low-variance dimension to a standard deviation of 12.88 and the high-variance distribution to a standard deviation of to 22.5. Only data from the test trials were analyzed. In total, subjects completed 420 trials with a short break every 60 trials. Once subjects made a response, there was a 2000-3000ms inter-trial-interval before the next trial began.

Apparatus and Stimuli

Stimuli were presented on an ASUS MG279Q monitor with a screen resolution of 1920 x 1200 pixels and a refresh rate of 60Hz. A computer running PsychoPy (Peirce, 2007) controlled all stimulus presentation and the recording of response variables. All participants were seated 60cm from the screen.

Stimuli were colored RDKs subtending 3.5 degrees of visual angle, with coherence set to one. Each RDK consisted of 20 dots that were .21 degrees of visual angle in diameter and moved at a speed of 2.4 degrees of visual angle per second. Motion and color were both randomly selected on each trial from a uniform distribution of motion directions in a 360-degree circle, and colors from CIE space (Bae, Olkkonen, Allred, & Flombaum, 2015) with one-degree increments.

Baseline experiment

In order to establish any inherent differences in representing the two feature dimensions (color, motion), a “baseline experiment” was run in which subjects only performed the memory probe task for a cued stimulus (Fig 2.1a), without interleaved visual search trials. The baseline experiment was used to establish the default precision of representing each feature dimension, irrespective of our manipulation of cue-to-target variance.

Twenty different students (females = 11, mean age = 21.37 years) from the University of California, Davis were recruited to participate in the study. Seventeen of these subjects were right-handed. The results showed that motion probe responses (M=11.05, SD=4.69) were significantly more precise that color probe (M=17.25, SD=7.05) responses ($t(19)=-6.19$, $p<.001$, $BF_{10}>300$). This suggest that representation of motion direction in our stimulus space was more precisely represented at baseline.

Results

We first looked at the response error on probe trials as a measure of the precision of template representations for the target. Subjects’ mean absolute response errors were normalized with respect to baseline response errors. We did this by calculating the deviation of each subjects mean response from baseline responses, weighted by the variance in baseline responses. We then recoded the probes so that it reflected whether the dimension was the low-variance or high-variance dimension for each subject and compared the response errors in a 2 (group) x 2 (dimension) ANOVA (Fig 2.2). The results showed no main effect of group (*F*(1,38)=.51, *p*=.48, BF_{10}=.36) nor an interaction between group and dimension (*F*(1,38)=1.79, *p*=.19, BF_{10}=*.6*).

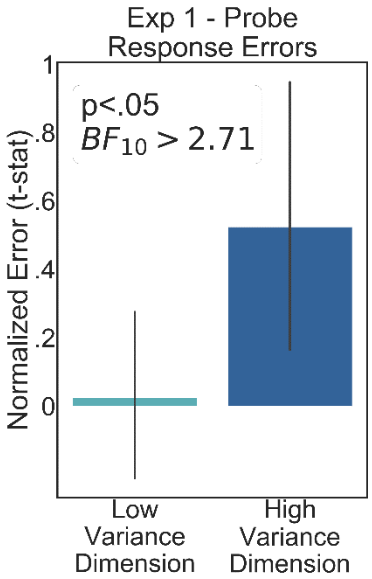

Figure 2.2: Response errors from probe trials. Errors were normalized to baseline values and compared. The representation of the low-variance dimension was more precise than that of the high-variance dimension.

Importantly, there was a significant effect of dimension such that the low-variance dimension was represented with lower error (i.e., more precisely) than the high-variance dimension (*F*(1,38)=5.46, *p*<.05, BF_{10}=*2.8*). This confirms our hypothesis that the variance of a feature dimension was encoded in the precision of the template for each cue feature.

Overall rates of correct responses on visual search trials were high for the motion group (M=.95, SD=0.48) and the color group (M=.93, SD=0.36), and no difference was detected between groups (*t*(35.18)=-1.04, *p*=.35, BF_{10}<.45). Due to near ceiling performance on accuracy and our a priori hypothesis that these effects would be more prominent in response times (RTs), we focused all subsequent analyses on RTs from trials where subject responded correctly. Response times were measured as the difference between stimulus onset and mouse click. Mean RTs were overall shorter when expectations of cue-to-target variance were met (Mean= 1.08 seconds, SD=.37) compared to when they were violated (Mean= 1.44 seconds, SD=.21), (*t*(39)=7.05, p < .001, BF_{10} *> 100)*. More importantly, we analyzed RTs as a function of the relationships between stimuli during search in each trial. We hypothesized that trial by trial RT would increase as the similarity between the cue and the target features decreased. Further, cue-to-target similarity in the low-variance dimension would have a greater effect on RT than the high-variance dimension because subjects will expect the target be more similar to the cue in this dimension and rely on it more. To test these predictions, we fit a hierarchical linear model with gamma-distributed residuals (Lo and Andrews, 2015) to subject response data with six fixed effects. We used two separate regressors for the cue-to-target similarity of features in the low-variance and high-variance dimensions in each trial (lvCT and hvCT). We also included regressors for target-distractor similarity (TD) and distractor-distractor relationships (DD). More specifically, for target-distractor similarity we took the absolute value of the differences between the target and each distractor and averaged them together for the low-variance and high-variance dimensions separately (lvTD and hvTD). For distractor-distractor similarity, we calculated the difference between each distractor and averaged those values for

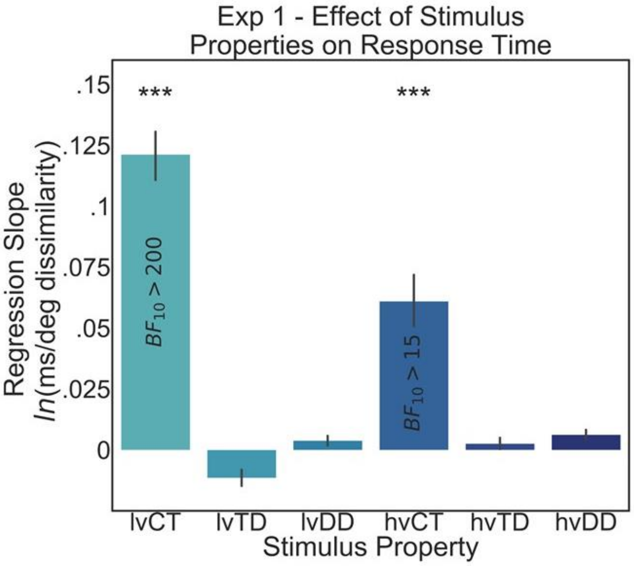

Figure 2.3: Regression slopes from a multilevel model measuring the effect of stimulus properties on response time. Bar heights show the size of the fixed effect while error-bars show the variance of the subject specific random effects. The model tested the effects of cue-to-target similarity (CT), target-distractor similarity (TD) and distractor-distractor similarity (DD) for the low-variance (lv) and high-variance (hv) dimensions separately. The model showed significant effects of lvCT and hvCT only, with more of the variance in RT explained by lvCT.

each dimension independently (lvDD and hvDD). All regressors were normalized before model-fitting due to large differences in the scale of each metric and we included all corresponding random-effects for each subject.

The results of this analysis (Fig 2.3) showed no interactions between group and any of the regressors ($\chi^2(1) < 2.08$, all *p*'s >.14, all BF_{10}*'s* <.05), and no main effect of group ($\chi^2(1)=.46$, *p*=.5, BF_{10}<.05), suggesting that any effects on RT did not differ between groups. The model showed a significant fixed-effect of cue-to-target similarity in both the low-variance (β =.121, $\chi^2(1)=28.21$, *p*<.001, BF_{10} > 200) and high-variance dimensions (β=.06, $\chi^2(1)=10.43$, *p* <.001, BF_{10}>15). All other main effects were non-significant and in favor of the null ($\chi^2(1)$'s > 3.72, all *p*'s > .05, all BF_{10} < .01).

We tested for a difference between these coefficients by fitting a separate model where low-variance and high-variance cue-to-target similarity were fixed to have the same effect size ($H_0\!: lvCT = hvCT$), and compared it to the original model where the parameters were free to differ ($H_1\!: lvCT \neq hvCT$). If the constrained model was an equally good or better fit for these data, this would suggest no difference in the magnitude of the coefficients. The results showed significant evidence in favor of the original model ($\chi^2(1)=11.652$, *p* < .001, BF >3), indicating that low-variance cue-to-target similarity coefficient was significantly larger than that for high-variance cue-to-target similarity. The pattern was confirmed by a Wald-test using the pooled variance of the regression estimates (*Z*=3.69, *p*<.001). These results show that RTs varied as a function of the cue-to-target similarity, but the strength was modulated by the dimension's expected variance: the low-variance feature was a stronger predictor of RT than the high-variance feature, consistent with the probe results showing a greater precision for the feature with lower expected variance.

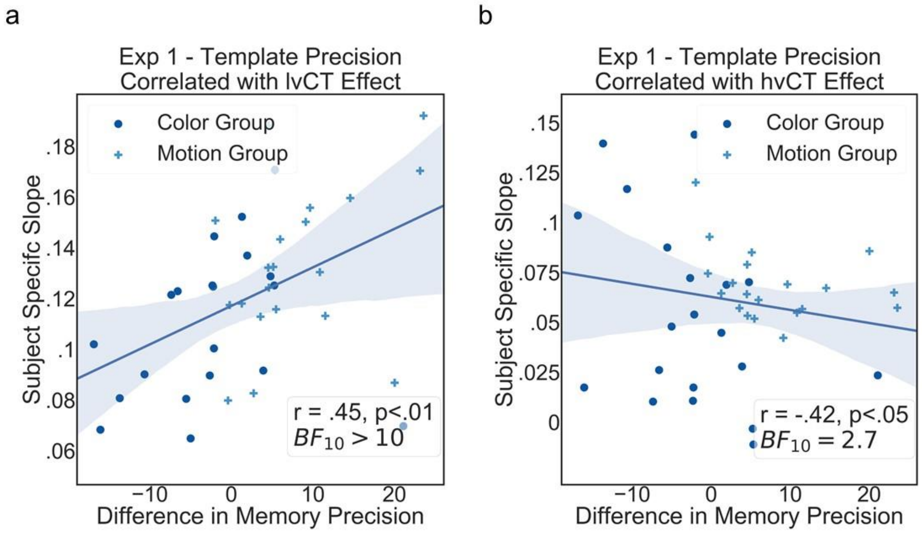

Figure 2.4: Correlations between individual coefficients from the RT regression models with the difference in memory precision between feature dimensions. The difference in memory precision score was calculated for each subject by subtracting the mean response error of the low-variance dimension from that in the high-variance dimension. High values of the difference in memory precision indicate subjects had a more precise template for the low-variance dimension relative to the high-variance dimension. (a) The difference in memory precision positively correlated with lvCT and (b) negatively with hvCT, suggesting that subject with templates biased towards the low-variance dimension used that dimension more during visual search.

To further test the relationship between cue-target similarity and visual search RT, we created a “difference in memory precision” score for each subject, calculated as the difference in probe precision between the high-variance dimension and the low-variance dimension (high-variance minus low-variance), which was then regressed against each individual’s corresponding coefficient for cue-to-target similarity from the response time regression models (Fig 2.4). This was done separately for the low- and high-variance dimensions separately. For the low-variance dimension, we found no main effect of group (β =.33 $t(38)$=.964, p=.34, BF_{10} <.3) nor an interaction between the difference in memory precision and group (β =.25, $t(38)$=.997, p=.49,

BF_{10} <.25). The results did show an effect of the difference in memory precision (β =.45, t(38)=3.07, *p*<.01, BF_{10}>10), indicating that individuals with larger precision indices also showed a larger effect of cue-to-target similarity in the low-variance dimension. This suggests that subjects who held a sharper template for the low-variance dimension were also more affected by this dimension during search when it mismatched the cue. An identical model fit for coefficients from the high-variance dimension showed the opposite effect. These results showed no interaction between the difference in memory precision and group (β =.38, *t*(38)=1.07, *p* >.29). However, we did find a main effect of group (β =.88, *t*(38)=2.52, *p*<.05, BF >3.75) and an effect of the difference in memory precision (β =-.42, *t*(38)=-2.38, *p*<.05, BF=2.74). This shows that subjects who maintained more precise representations of the low-variance dimensions prior to search weighted that information more during search and down-weighted the less precise, high-variance dimension making it less useful during search.

Eye-tracking data

To obtain a finer grained understanding of how cue-to-target similarity influenced the scan time (i.e., time to first fixate the target) and decision time (i.e., duration of looking at the target before response) of visual search, we turned to eye-tracking data. One subject was removed from these analyses due to technical issues that occurred during data collection.

First, we fit a hierarchical regression model with normally distributed residuals predicting scan time (i.e., time to first fixate the target) using cue-to-target variance, target-distractor similarity, and distractor-distractor similarity in each feature dimension as predictors (Fig 2.5a). Group was not included as a predictor in this model to reduce model complexity given that it had no effect on response times for search. We found that both the low-variance feature dimension (β=30.47, $\chi^2(1)$=22.04, p <.001, BF_{10}>150) and the high-variance feature dimension

(β=24.37, $\chi^2(1)$=11.05, p <.001, BF_{10}>3) predicted scan times. Again, we tested the difference

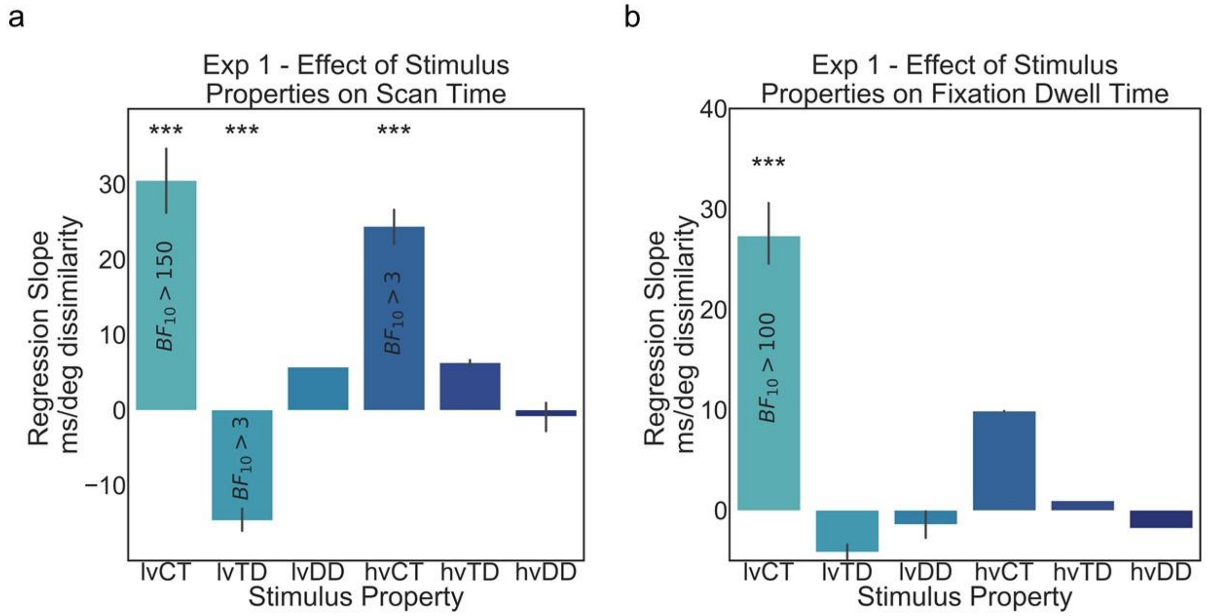

Figure 2.5: Regressions slopes measuring the effect of stimulus properties on (a) scan time and (**b)** fixation dwell time on the target, from a multilevel regression model. Bar heights show the size of the fixed effect while error-bars show the variance of the subject specific random effects. The model tested the effects of cue-to-target similarity (CT), target-distractor similarity (TD) and distractor-distractor similarity (DD) for the low-variance (lv) and high-variance (hv) dimensions separately. Results show that while there was a small effect of lvTD, lvCT and hvCT had greater effects on scan time, indicating that the cue-to-target similarity dominated scan times. However, only lvCT influenced fixation dwell times, suggesting that the similarity of the cue-to-target in the low-variance dimension played a unique role in deciding if the stimulus matched the target.

between these coefficients by fitting another model in which they were artificially constrained to be equal and found evidence in favor of the constrained model, suggesting that there was no significant difference in the strength of these two coefficients on scan time ($\chi^2(1)$ =.61 p =.435, BF_{10}<.2; Wald-test: Z=.785, p=.216). We also found an effect of low-variance target-distractor similarity (β=-14.63, $\chi^2(1)$ =11.02, p <.001, BF_{10}>3). To test whether this was different from either the cue-to-target similarity effects we flipped the sign of the coefficient (to have an unbiased estimate of absolute magnitude) then fit two independent models where the low-variance target-distractor coefficient was set to be equal to the low-variance or high-variance cue-to-target similarity coefficient. This allowed us to test whether the magnitude of the coefficient for the low-variance target-distractor similarity – independent of the sign – was different from the two cue-target effects. The model comparisons revealed a difference with the low-variance cue-target similarity ($\chi^2(1)$=6.35 p <.05; Wald-test: Z=1.96, p<.05), although the Bayes factor was ambiguous (BF_{10} = .37), and no difference with high-variance cue-to-target similarity ($\chi^2(1)$=1.45 p =.23, BF_{10} < .01; Wald-test: Z=1.18, p=.12). This suggests scan times increased similarly when the target differed more from the cue in either feature dimension and decreased when target and distractor features in low-variance dimension were dissimilar. However, cue-to-target similarity in low-variance dimension had a larger effect on scan times than target-distractor similarity.

A second model with identical regressors was fit to predict decision time (i.e., total dwell time on the target before response) (Fig 2.5b). The results were significant for cue-target similarity in the low-variance feature (β=27.31, $\chi^2(1)$=20.43, p <.001, BF_{10}>20), but not the high-variance feature (β=9.87, $\chi^2(1)$=2.56, p=.1, BF_{10}<.05). All other regressors had a null effect (all $\chi^2(1)$'s < 1.4, all p's>.24, all BF_{10}<.05). These results extend the RT and scan time analyses showing that cue-target similarity in the low-variance dimension was the primary source of information used to search for and decide if an object is the target. This suggests that statistical regularities in each feature dimension contributed to information processing at different stages of visual search. Higher attentional priority was given to information from low-variance dimension resulting in more precise encoding of that dimension and more reliance on the dimension guide search and make target decisions.

Discussion

The purpose of Experiment 1 was to test the hypothesis that statistical regularities governing cue-to-target similarity in two feature dimensions would be reflected within the target template and have a concomitant effect on visual search efficiency. The data were consistent with these hypotheses, showing that on probe trials there was a smaller response error for the low-variance feature compared to the high-variance feature; on visual search trials, RT was better predicted by the cue-to-target similarity on the low-variance than the high-variance feature dimension. Moreover, individuals with a greater precision for the low-variance dimension were more affected when the low-variance target feature was distant from the cue. This suggests a direct relationship between the precision of the template representation on each dimension and the effect of cue-to-target similarity in a feature dimension during visual search.

In addition to overall RT, the eye data demonstrated that the effect of cue-target variance on visual search could be broken down into two components: search duration, indicated by scan time and decision duration, indicated by the fixation time on the target. Interestingly, scan time was influenced by both feature dimensions, but decision time was predicted only by the low-variance feature. This suggests that the dimension with lower variance was the primary information source for deciding if an object is the target, but both dimensions affected the time to locate the target. Together, these results demonstrate that subjects maintain a more precise representation of low-variance dimension and correspondingly relied on information from low-variance during visual search.

Experiment 2

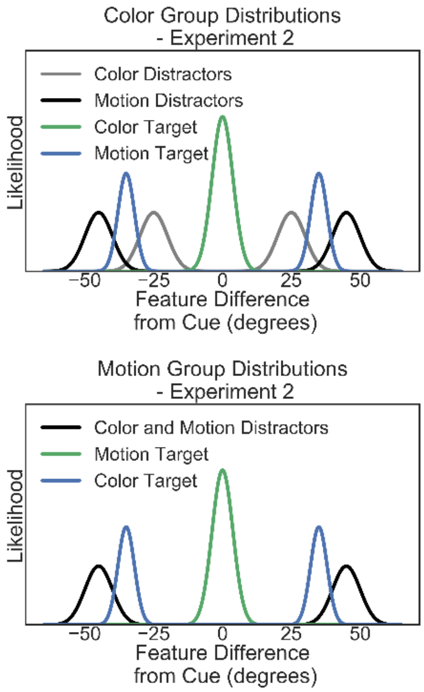

Figure 2.6:Distributions of target and distractor features for Experiment 2 for the “color group” and the “motion group”. Zero represents no difference from the cue. The distractor distributions were generated around category boundaries to increase distractor competition.

The purpose of Experiment 2 was to provide a conceptual replication of Experiment 1 and test if stronger distractor competition would reduce the effect of cue-to-target variance on visual search. Distractor competition is hypothesized to be a core precondition for attentional selection (Desimone and Duncan, 1995; Duncan and Humphreys, 1989). It was therefore somewhat surprising in Experiment 1 to find that target-distractor similarity and distractor-distractor similarity did not predict RT. One reason for this may be that distractors were sufficiently distant from targets in Experiment 1 (i.e., at least 30 degrees from the target, sampled from a uniform distribution) that they produced very little competition for attention (Duncan, 1989; Wolfe et al. 1989; Wolfe, 1992). In this experiment, we increased distractor similarity, and therefore competition for attention, by focusing distractor distributions much closer to the target distributions. We expected the increase in search difficulty to increase RTs overall. The critical question was whether doing so would also lead to the emergence of target-distractor similarity and distractor-distractor similarity as predictors of visual search RT and scan times. However, we predict that the manipulation will not change the effect of cue-to-target similarity decision times, given that decisions are more purely based on matching the current stimulus to the target template.

Methods

Participants

Forty students (Females = 24, Age = 20.9) from the University of California, Davis were recruited to participate in the study. All participants had normal or corrected-to-normal vision, normal color vision, and received course credit for participating in the experiment. Thirty-six subjects were right-handed.

Procedure

All methods were identical to Experiment 1, except for the feature distributions used. New distributions were used to increase the likelihood that distractors would be highly similar to a given target object and create more competition. To do this, both motion and color distractors were drawn from bimodal distributions centered on -45 and 45 degrees, with SD's of 5 degrees (Fig 2.6). However, when color was the low-variance dimension (in the color group) color distractors were drawn from distributions centered on -25 and 25 degrees with SD's of 5 degrees – this was necessary to ensure competition given that 25 degrees is the average distance between the category boundary and category center for all colors (Bae, Olkkonen, Allred, & Flombaum, 2015).

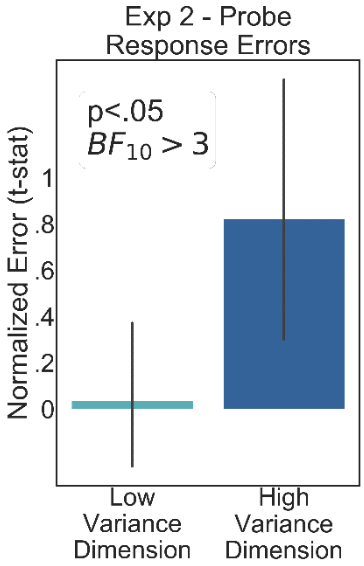

Figure 2.7: Response errors from probe trials normalized with the baseline measures (see Methods, Experiment 1). These data replicate Experiment 1 showing representations of the low-variance dimension were more precise than those of the high-variance dimension.

The target features from the high-variance dimension were drawn from a bimodal distribution with centers at -35 and 35, and SD's of 3.8 (Fig 2.6). This ensured that the distractor features were always near target features, while still maximizing the distance of the high-variance target feature from the cue. The low-variance target distribution was the same as in Experiment 1. Because the high-variance target distributions were non-overlapping with the low-variance target distributions, target features in each dimension were always inversely correlated with each other – when one feature was sampled from the low-variance distribution, the other was sampled from the distributions centered 35 degrees away, and vice versa on switch trials. This meant that the overall distribution for the low-variance features had standard deviation of 12 degrees and the high-variance distribution had a standard deviation of 33 degrees. Because of this covariation in the two target features, we used a single metric of feature similarity in our analyses that combined both feature dimensions together in this experiment. The combined regressor represents the impact of the expectation that the low-variance dimension will have feature values more similar to the cue than the high-variance dimension.

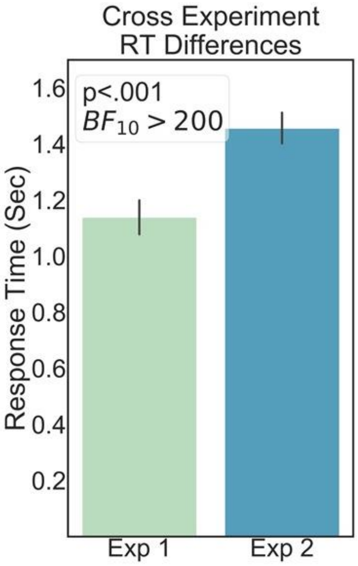

Figure 2.8: Comparison of RT from Experiments 1 and 2. These data show that correct RTs in Experiment 2 were significantly longer than in Experiment 1, suggesting the overall difficulty of Experiment 2 was successfully increased by distractor competition.

Results

We again first looked at the error in probe responses as a measure of the precision of the template representation for the cue. The response errors were first normalized with respect to baseline error data and then compared in a 2 (group) x 2 (dimension) ANOVA. Replicating results from Experiment 1, we found no effect of group ($F(1,38)=.114$, $p<.737$, $BF_{10}=.3$) nor an interaction between group and dimension ($F(1,38)=.878$, $p=.35$, $BF_{10}=.47$). Only main effect of dimension ($F(1,38)=6.38$, $p< .01$, $BF_{10} > 3$) was significant, showing that the low-variance dimension was more precisely represented compared to the high-variance dimension. This result replicates the probe data from Experiment 1.

Overall accuracy rates were lower for both motion (Mean = .69, SD=.07) and color groups (Mean = .65, SD=.08), compared to Experiment 1, but no differences were found between groups (*t*(39)=1.73, *p*=.09, BF_{10}=1). The change in accuracy reflects our success in increasing the difficulty of search through increased distractor competition. Similarly, mean RT's were longer in Experiment 1 (M=1.46s, SD=.18) than Experiment 2 (M=1.14s, SD=.2), (*t*(76.43)=7.43, *p*<.001, BF_{10}>200), suggesting that increasing distractor competition did indeed make the visual search task more difficult (Fig 2.8).

Next, we fit regression models in the same way as Experiment 1, with the exception that now the cue-to-target similarity in both dimensions were combined into a single regressor representing the overall cue-to-target similarity. This was done by subtracting the high-variance cue-to-target similarity from the low-variance cue-to-target similarity. Negative values mean that the low-variance target feature was closer to the cue on that trial; positive values mean that the low-variance target feature was more distant from the cue on that trial (which occurred infrequently). We coded it this way so that larger positive coefficients would correspond to longer RTs in response to target stimuli that mismatched the anticipated target values.

Mean RTs were again shorter when expectations of cue-to-target variance were met (Mean= 1.41 seconds, SD=.18) compared to when these expectations were violated (Mean= 1.67 seconds, SD=.27), (*t*(39)=6.89, p < .001, BF_{10} *> 100).* The results showed a positive effect of cue-to-target similarity (β =.048, $\chi^2(1) = 7.89$, *p*<.001, BF_{10} > 100). This finding replicates

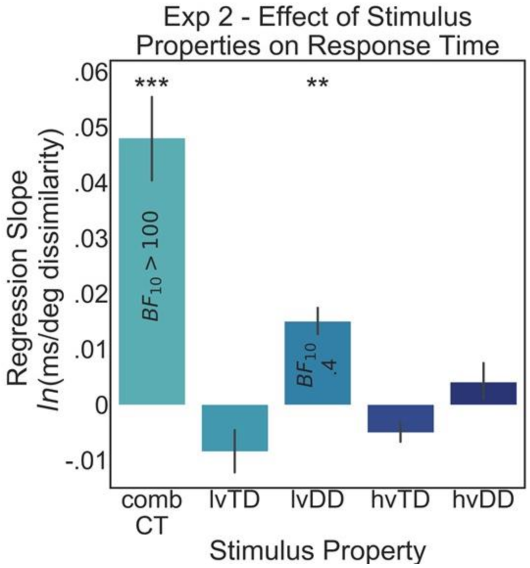

Figure 2.9: Slopes measuring the effect of stimulus properties on response time from a multilevel regression model. Bar heights show the size of the fixed effect while error-bars show the variance of the subject specific random effects. The model tested the effects of cue-to-target similarity (CT) combined from both dimension, and target-distractor similarity (TD) and distractor-distractor similarity (DD) for the low-variance (lv) and high-variance (hv) dimensions separately. Results show that while there was a small effect of lvDD, the combined effect of CT – the difference between the target and the way it was expected to appear – was again the primary predictor of response time variation.

Experiment 1 by showing that violations of the expected cue-to-target similarity increased RT (Fig 2.9) and suggesting that subjects relied on task-relevant expectations of target feature variance to guide attention during visual search. There was also an effect of distractor-distractor similarity in the low-variance dimension (β=.016, $\chi^2(1)$=6.45, *p* <.05), although this result should be interpreted with caution as the Bayes Factor indicated marginal evidence in favor of the null (BF_{10} < .4). Further, the magnitude of the coefficient was significantly smaller compared to the combined coefficient for cue-to-target similarity ($\chi^2(1)$ = 7.21, *p*<.01, BF_{10} =1.6; Wald-test: *Z*=4.98, *p*<.001), suggesting that expectations about the target features had a larger effect on RTs than distractor similarity. There were no significant main effects associated with group ($\chi^2(1)$ =.46, *p*=.5, BF_{10} <.05) or any other conditions ($\chi^2(1)$'s < 3.72, all *p*'s > .05, all BF_{10} < .01). There were also no interactions between group and any regressor ($\chi^2(1)$'s < 2.08, all *p*'s >.14, all BF_{10}*'s* <.05), except for distractor-distractor differences ($\chi^2(1)$ = 6.2, *p*<.05), suggesting a larger effect of distractor-distractor similarity for the color group. However, analysis of the Bayes Factor indicated evidence in favor of a null effect ($BF_{10} < .2$).

Finally, we correlated the subject specific slopes from the regression model to each subject's difference in memory precision, calculated by subtracting the probe error in the low-variance dimension from the high-variance dimension (Fig 2.10). We again found a positive correlation between these variables ($\beta = .46$, $t(38)=2.55$, $p<.05$, BF>10), replicating Experiment 1. This shows that the bias in the precision of the subject's template representations in memory towards the low-variance dimension directly relates to visual search performance.

Eye-Tracking Data

Next, we analyzed eye-tracking data by fitting a hierarchical model with normally distributed residuals to data predicting scan time using the same regressors as in Experiment 1 (Fig 2. 11a). A single subject was removed from these analyses due to issues during data collection. We found a null effect of cue-to-target similarity ($\beta = .04$, $\chi2(1)=0.027$, $p=.87$, $BF_{10} <.05$), and null effects for target-distractor similarity ($\chi2(1)$'s < 1, all p's >.4, all BF_{10}'s<.05) . However, we found commensurate effects of distractor-distractor differences in both the low-variance ($\beta = 16.23$, $\chi2(1)=10.79$, $p<.01$, $BF_{10} >3$), and the high-variance dimensions ($\beta = 14.73$, $\chi2(1)=8.65$, $p<.01$, $BF_{10} = 1.2$). In addition, we found no

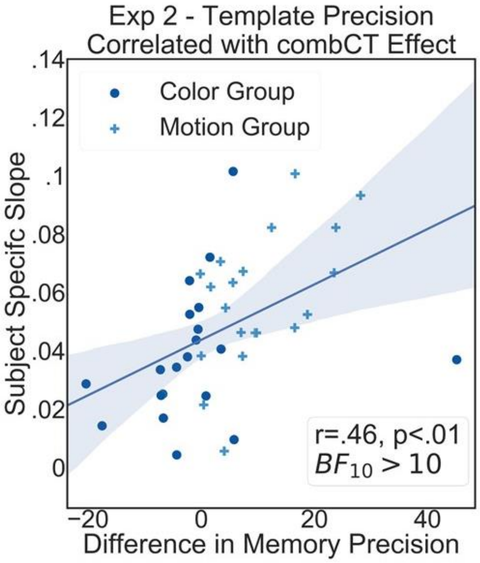

Figure 2.10: Subject specific coefficients from the regression models were correlated with each subject's difference in memory precision.

difference in the magnitude of these coefficients ($\chi2(1) = .049$, $p=.82$, $BF_{10} <.02$; Wald-test: $Z=.22$, $p=.41$). These suggest that in this task, attention was guided via perceptual grouping and rejection of distractors. However, cue-to-target similarity continued to be the only significant predictor of dwell time on the target ($\beta = 25.69$, $\chi2(1)=10.67$, $p<.01$, $BF_{10} >3$; Fig 2.11b). All other effects in both models were null (all $\chi2(1)$'s < 1, all p's >.4, all BF_{10}'s<.05). These data suggest that cue-to-target similarity continued to have a unique effect on the time subjects needed to decide whether the target matched their template, but that the effect on scan time was reduced when distractor competition was increased.

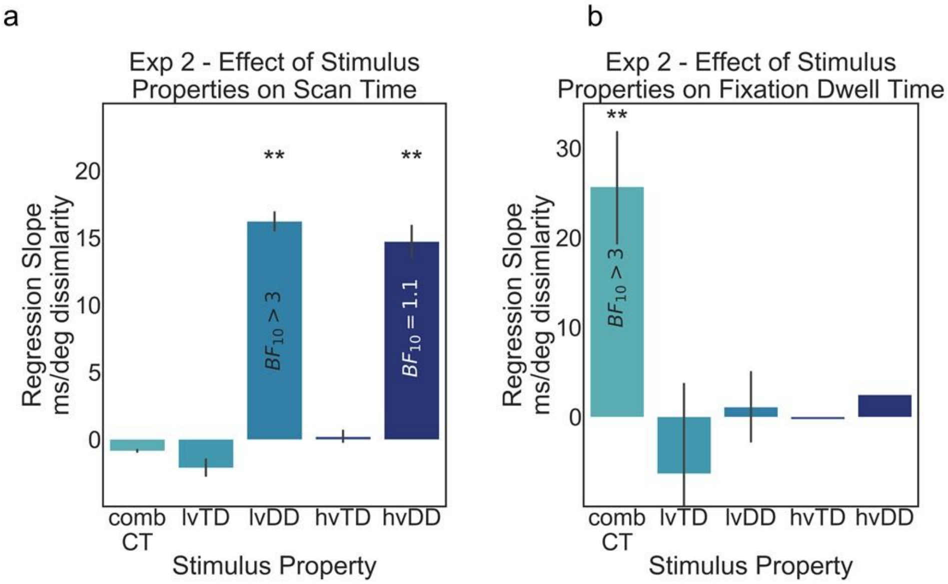

Figure 2.11: Slopes measuring the effect of stimulus properties on scan time (a**)** and fixation dwell time on the target (b**)**, from a multilevel regression model. Bar heights show the size of the fixed effect while error-bars show the variance of the subject specific random effects. The models show that while variation in scan time was primarily a function of distractor-distractor differences, CT was still the only regressor that explained variation in dwell time.

Discussion

Experiment 2 provides a conceptual replication of Experiment 1, again showing that expectations of cue-to-target similarity are the strongest predictor of variation in RT even when distractor competition has made the task overall more difficult. In this experiment we also increased distractor competition and show that under these conditions overall response times are longer, with increases in the effect of distractor-distractor similarity on search and scan time. This suggests that even though scan times may be influenced more by distractor properties when distractor competition is high, decisions about the target depend more purely on how well the stimulus matches the target template.

General Discussion

The purpose of these experiments was to test hypotheses of how expected variability in target features shape the target template and visual search performance. We introduced a new paradigm which uses systematic variation between the cue and target stimuli to ask whether the level of variation would predict the quality of the target template and visual search performance. The contents of the target template (i.e., what is remembered about the cue stimulus) was measured through probe trials that asked subjects to recall each cue feature using a color or motion wheel. These trials did not require attentional processes engaged in selection and therefore served as a direct measurement of the target representation, uncontaminated by competition for attention. Target selection was measured through visual search trials that included distractors that differed in target similarity (Experiments 1 and 2). Eye-tracking data during the search display allowed us to dissociate the consequence of expected variability on the scan time to locate the target and the decision time to determine that an object is the target once it was fixated.

In the first experiment we manipulated the cue-to-target variance using two distributions, both centered on the cue feature value. They differed in that the “low-variance dimension” target feature was typically sampled from a narrower distribution than the “high-variance dimension” feature. We found that features of the low-variance dimension were represented more precisely in the target template and cue-to-target similarity along this dimension was significantly better predictor of visual search RT. Moreover, the low-variance feature predicted both scan times and decision times, whereas the high-variance feature only predicted scan times

This pattern of results indicates both target features were used to locate target, but only the more precise one was used to make the final decision regarding whether the selected object was indeed the target. Subjects clearly used the task-relevant expectations of cue-to-target variance to bias the target template towards the more reliable feature, but this bias was not all-or-none: the high-variance feature was still maintained and used to constrain search for the target, as indicated by the significant coefficient for high-variance cue-to-target similarity. This latter finding is in line with those from Foerster & Schneider (2018) showing that the task-irrelevant color of a cue object biased the direction of first saccades during search. More generally, others have found evidence that templates in working memory are maintained as bounded features (Luck and Vogel, 1997; Vogel et al., 2001), but attentional biases toward each feature are weighted by their task-relevance (Müller et al., 1995; Reeder et al., 2017). Thus, our data are consistent with previous work showing attentional guidance from task-irrelevant features and go further to show that learned variance is used to define the task-relevant dimension even both dimensions could be used to identify the target. Specifically, these data show that when target features are uncertain, the low-variance dimension is used to make target-match decisions.

In Experiment 2 we increased distractor competition in order to see if doing so would reduce the importance of cue-to-target variance on search processes. To do so, selection of distractor features was constrained to distributions that increased distractor similarity to the target. Following the literature (Duncan and Humphreys, 1989), this manipulation made search harder, resulting in overall longer visual search RTs and lower accuracy compared to Experiment 1. Despite this change, cue-to-target variance was still the strongest predictor of behavior overall. However, the eye-tracking data now showed that distractor-distractor similarity predicted scan times. It took longer to fixate the target when distractor heterogeneity was greater. This suggests that scan times were influenced by distractor properties when the distractors were overall more similar to the target. Interestingly, cue-to-target similarity was still the only predictor of decision times once the target was fixated, replicating Experiment 1. The eye-tracking data suggests that different components of visual search (scan time, decision time) are informed by different stimulus properties (Malcolm and Henderson, 2010).

Together, these experiments show that target feature variability is reflected in the attentional template and has direct consequences on visual search. Task-relevant visual dimensions that are expected to be less variable are held in working memory with greater precision compared to those that are expected to be highly variable. The efficiency of search is similarly determined by expected variance of the task-relevant dimensions, presumably because search times are determined by how well the template matches targets. It is worth noting that in our studies, learned expectations were for task-relevant stimulus properties and therefore we cannot disentangle contributions from stimulus expectations devoid of relevance in our data. In recent years, stimulus expectations have been shown to be separable from task-relevance (Summerfield and Egner, 2009), but in the current paradigm, the expectation of the variance of the distribution defines the task-relevance of each target feature. It is possible that subjects are estimating the perceived relevance of each feature dimension (i.e. how likely is it that they can use that dimension to identify the target) as opposed to estimating the expected feature values of the target, *per se* (i.e. the likelihood that the target feature will have some value). Future work is needed to determine how expectations of stimulus values vs. their utility for target selection differ in guiding attention.

Taken together, these results illustrate that attention is sensitive to the reliability of different target feature dimensions and will adaptively encode less-variable features and use them to guide search. When subjects expect a feature dimension will be a reliable source of information about how the target will appear, they represent this information in the attentional template with greater precision and rely on this information more during search. In contrast, when a target feature is more likely to be dissimilar to the cue, it is likely to have a less precise representation and have less of an effect on search. In addition to illuminating the importance of expected variance of different target features on attention, the data suggest that visual search involves multiple stages of processing that differentially weight sources of information. In sum, these studies confirm the hypothesis that the attentional templates are sensitive to variation in specific target features and actively weights information from low-variance dimensions compared to high-variance dimensions during visual search but also suggest that information from both dimensions are preserved during search for the target.

References

Andersen, S.K., Hillyard, S.A., and Müller, M.M. (2008). Attention Facilitates Multiple Stimulus Features in Parallel in Human Visual Cortex. Curr. Biol. 18, 1006–1009.

Badre, D., Doll, B.B., Long, N.M., and Frank, M.J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. Neuron 73, 595–607.

Bae, G.Y., Olkkonen, M., Allred, S.R., and Flombaum, J.I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. J. Exp. Psychol. Gen. 144, 744–763.

Becker, S.I., Folk, C.L., and Remington, R.W. (2010). The Role of Relational Information in Contingent Capture. J. Exp. Psychol. Hum. Percept. Perform. 36, 1460–1476.

Van Bergen, R.S., and Jehee, J.F.M. (2019). Probabilistic Representation in Human Visual Cortex Reflects Uncertainty in Serial Decisions. J. Neurosci. 39, 8164–8176.

Bichot, N.P., Heard, M.T., DeGennaro, E.M., and Desimone, R. (2015). A Source for Feature-Based Attention in the Prefrontal Cortex. Neuron 88, 832–844.

Boettcher, S.E.P., van Ede, F., and Nobre, A.C. (2020). Functional biases in attentional templates from associative memory. J. Vis. 20, 7.

Boettcher, S.E.P., Gresch, D., Nobre, A.C., and van Ede, F. (2021). Output planning at the input stage in visual working memory. Sci. Adv. 7, eabe8212.

Bravo, M.J., and Farid, H. (2012). Task demands determine the specificity of the search template. Attention, Perception, Psychophys. 74, 124–131.

Bravo, M.J., and Farid, H. (2016). Observers change their target template based on expected context. Attention, Perception, Psychophys. 78, 829–837.

Buschman, T.J. (2021). Balancing Flexibility and Interference in Working Memory. Annu. Rev. OfVision Sci. 7, 1–22.

Buschman, T.J., and Miller, E.K. (2007). Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and. Science (80-.). 315, 1860–1863.

Carandini, M., and Heeger, D.J. (2012). Normalization as a canonical neural computation. Nat. Rev. Neurosci. 13, 51–62.

Chelazzi, L., Miller, E.K., Duncan, J., and Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. Nature 363, 345–347.

Chelazzi, L., Duncan, J., Miller, E.K., and Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. J. Neurophysiol. 80, 2918–2940.

Chetverikov, A., Campana, G., and Kristjánsson, Á. (2017a). Learning features in a complex and changing environment: A distribution-based framework for visual attention and vision in general. Prog. Brain Res. 236, 97–120.

Chetverikov, A., Campana, G., and Kristjánsson, Á. (2017b). Representing Color Ensembles. Psychol. Sci. 28, 1510–1517.

Chetverikov, A., Campana, G., and Kristjánsson, Á. (2020). Probabilistic rejection templates in visual working memory. Cognition 196, 104075.

Chun, M.M., and Jiang, Y. (1998). Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention. Cogn. Psychol. 36, 28–71.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. 36, 181–204.

Desimone, R., and Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. Annu. Rev. Neurosci. 18, 193–222.

Driver, J. (2001). A selective review of selective attention research from the past century. Br. J. Psychol. 92, 53–78.

Duncan, J. (1989). Boundary condition on parallel in human vision. Perception 18, 457–469.

Duncan, J., and Humphreys, G.W. (1989). Visual Search and Stimulus Similarity. Psychol. Rev. 96, 433–458.

van Ede, F., Board, A.G., and Nobre, A.C. (2020). Goal-directed and stimulus-driven selection of internal representations. Proc. Natl. Acad. Sci. 202013432.

van Ede, F., Deden, J., and Nobre, A.C. (2021). Looking ahead in working memory to guide sequential behaviour. Curr. Biol. 31, R779–R780.

Fetsch, C.R., Pouget, A., Deangelis, G.C., and Angelaki, D.E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. Nat. Neurosci. 15, 146–154.

Foerster, R.M., and Schneider, W.X. (2018). Involuntary top-down control by search-irrelevant features: Visual working memory biases attention in an object-based manner. Cognition 172, 37–45.

Foley, N.C., Kelly, S.P., Mhatre, H., Lopes, M., and Gottlieb, J. (2017). Parietal neurons encode expected gains in instrumental information. Proc. Natl. Acad. Sci. 114, E3315–E3323.

Found, A., and Müller, H.J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a "dimension-weighting" account. Percept. Psychophys. 58, 88–101.

Friston, K. (2010). The free-energy principle: A unified brain theory? Nat. Rev. Neurosci. 11, 127–138.

Geng, J.J., and Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. Percept. Psychophys. 67, 1252–1268.

Geng, J.J., and Witkowski, P. (2019). Template-to-distractor distinctiveness regulates visual search efficiency. Curr. Opin. Psychol. 29, 119–125.

Geng, J.J., Di Quattro, N.E., and Helm, J. (2017). Distractor probability changes the shape of the attentional template. J. Exp. Psychol. Hum. Percept. Perform. 43, 1993–2007.

Geng, J.J., Won, B.Y., and Carlisle, N.B. (2019). Distractor Ignoring: Strategies, Learning, and Passive Filtering. Curr. Dir. Psychol. Sci. 28, 600–606.

Gershman, S.J. (2019). Uncertainty and exploration. Decision 6, 277–286.

Giesbrecht, B., Woldorff, M.G., Song, A.W., and Mangun, G.R. (2003). Neural mechanisms of top-down control during spatial and feature attention. Neuroimage 19, 496–512.

Gottlieb, J. (2018). Understanding active sampling strategies: Empirical approaches and implications for attention and decision research. Cortex 102, 150–160.

Gottlieb, J., and Balan, P. (2010). Attention as a decision in information space. Trends Cogn. Sci. 14, 240–248.

Gottlieb, J., Oudeyer, P.Y., Lopes, M., and Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. Trends Cogn. Sci. 17, 585–593.

Gramann, K., Toellner, T., Krummenacher, J., Eimer, M., and Müller, H.J. (2007). Brain electrical correlates of dimensional weighting: An ERP study. Psychophysiology 44, 277–292.

Grubert, A., and Eimer, M. (2013). Qualitative differences in the guidance of attention during single-color and multiple-color visual search: Behavioral and electrophysiological evidence. J. Exp. Psychol. Hum. Percept. Perform. 39, 1433–1442.

Hénaff, O.J., Boundy-singer, Z.M., Meding, K., and Ziemba, C.M. (2020). Representation of visual uncertainty through neural gain variability. Nat. Commun. 1–12.

Henderson, J.M., and Hayes, T.R. (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. Nat. Hum. Behav. 1, 743–747.

Hillis, J.M., Watt, S.J., Landy, M.S., and Banks, M.S. (2004). Slant from texture and disparity cues: Optimal cue combination. J. Vis. 4, 967–992.

Hollingworth, A., and Hwang, S. (2013). The relationship between visual working memory and attention: Retention of precise colour information in the absence of effects on perceptual selection. Philos. Trans. R. Soc. B Biol. Sci. 368.

Horan, M., Daddaoua, N., and Gottlieb, J. (2019). Parietal neurons encode information sampling based on decision uncertainty. Nat. Neurosci. 22, 1327–1335.

Hout, M.C., and Goldinger, S.D. (2015). Target templates: the precision of mental representations affects attentional guidance and decision-making in visual search. Attention, Perception, Psychophys. 77, 128–149.

Huang, C., Vilotijević, A., Theeuwes, J., and Donk, M. (2021). Proactive distractor suppression elicited by statistical regularities in visual search. Psychon. Bull. Rev. 28, 918–927.

Iglesias, S., Mathys, C., Brodersen, K.H., Kasper, L., Piccirelli, M., denOuden, H.E.M., and Stephan, K.E. (2013). Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning. Neuron 80, 519–530.

Kastner, S., and Ungerleider, L.G. (2000). Mechanisms of Visual Attention in the Human Cortex. Annu. Rev. Neurosci. 23, 315–341.

Kerzel, D., and Huynh Cong, S. (2021). Attentional templates are sharpened through differential signal enhancement, not differential allocation of attention. J. Cogn. Neurosci. 33, 594–610.

Kerzel, D., and Witzel, C. (2019). The allocation of resources in visual working memory and multiple attentional templates. J. Exp. Psychol. Hum. Percept. Perform. 45, 645–658.

Knill, D.C., and Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. Trends Neurosci. 27, 712–719.

Kok, P., Jehee, J.F.M., and de Lange, F.P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. Neuron 75, 265–270.

Kozyrev, V., Daliri, M.R., Schwedhelm, P., and Treue, S. (2019). Strategic deployment of feature-based attentional gain in primate visual cortex. PLoS Biol. 17, 1–28.

Lee, J., and Geng, J.J. (2020). Flexible weighting of target features based on distractor context. Attention, Perception, Psychophys. 82, 739–751.

Li, H.-H., Sprague, T.C., Yoo, A.H., Ma, W.J., and Curtis, C.E. (2021). Joint representation of working memory and uncertainty in human cortex. Neuron 2021.04.05.438511.

Liu, T., Larsson, J., and Carrasco, M. (2007). Feature-Based Attention Modulates Orientation-Selective Responses in Human Visual Cortex. Neuron 55, 313–323.

Lo, S., and Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. Front. Psychol. 6, 1171–1187.

Louie, K., Khaw, M.W., and Glimcher, P.W. (2013). Normalization is a general neural mechanism for context-dependent decision making. Proc. Natl. Acad. Sci. U. S. A. 110, 6139–6144.

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

Luke, S.G. (2017). Evaluating significance in linear mixed-effects models in R. Behav. Res. Methods 49, 1494–1502.

Ma, W.J., and Jazayeri, M. (2014). Neural Coding of Uncertainty and Probability. Annu. Rev. Neurosci. 37, 205–220.

Malcolm, G.L., and Henderson, J.M. (2009). The effects of target template specificity on visual search in real-world scenes : Evidence from eye movements. J. Vis. 9.

Malcolm, G.L., and Henderson, J.M. (2010). Combining top-down processes to guide eye movements during real-world scene search. J. Vis. 10.

Mehrpour, V., Martinez-Trujillo, J.C., and Treue, S. (2020). Attention amplifies neural representations of changes in sensory input at the expense of perceptual accuracy. Nat. Commun. 11.

Meyyappan, S., Rajan, A., Mangun, G.R., and Ding, M. (2021). Role of inferior frontal junction (ifj) in the control of feature versus spatial attention. J. Neurosci. 41, 8065–8074.

Michael, E., De Gardelle, V., Nevado-Holgado, A., and Summerfield, C. (2015). Unreliable evidence: 2 sources of uncertainty during perceptual choice. Cereb. Cortex 25, 937–947.

Moore, T., and Fallah, M. (2001). Control of eye movements and spatial attention. Proc. Natl. Acad. Sci. 98, 1273–1276.

Müller, H.J., and Krummenacher, J. (2006). Locus of dimension weighting: Preattentive or postselective? Vis. Cogn. 14, 490–513.

Müller, H.J., Heller, D., and Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. Percept. Psychophys. 57, 1–17.

Müller, H.J., Reimann, B., and Krummenacher, J. (2003). Visual Search for Singleton Feature Targets Across Dimensions: Stimulus- and Expectancy-Driven Effects in Dimensional Weighting. J. Exp. Psychol. Hum. Percept. Perform. 29, 1021–1035.

Nah, J.C., and Geng, J.J. (2022). Thematic object pairs produce stronger and faster grouping than taxonomic pairs. J. Exp. Psychol. Hum. Percept. Perform. 48, 1325–1335.

Nako, R., Wu, R., Smith, T.J., and Eimer, M. (2014). Item and Category-Based Attentional Control During Search for Real-World Objects : Can You Find the Pants Among the Pans ? J. Exp. Psychol. Hum. Percept. Preformance 40, 1283–1288.

Navalpakkam, V., and Itti, L. (2007). Search Goal Tunes Visual Features Optimally. Neuron 53, 605–617.

Niklaus, M., Nobre, A.C., and Van Ede, F. (2017). Feature-based attentional weighting and spreading in visual working memory. Sci. Rep. 7.

Niv, Y., Daniel, R., Geana, A., Gershman, S.J., Leong, Y.C., Radulescu, A., and Wilson, R.C. (2015). Reinforcement Learning in Multidimensional Environments Relies on Attention Mechanisms. J. Neurosci. 35, 8145–8157.

Noah, S., Meyyappan, S., Ding, M., and Mangun, G.R. (2022). Anticipatory attention is a stable state induced by transient control mechanisms. Front. Hum. Neurosci. 16, 1–12.

Nobre, A.C., and Stokes, M.G. (2019). Premembering Experience: A Hierarchy of Time-Scales for Proactive Attention. Neuron 104, 132–146.

O’Reilly, J.X., Schüffelgen, U., Cuell, S.F., Behrens, T.E.J., Mars, R.B., and Rushworth, M.F.S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. Proc. Natl. Acad. Sci. U. S. A. 110.

Olivers, C.N.L., Meijer, F., and Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. J. Exp. Psychol. Hum. Percept. Perform. 32, 1243–1265.

Olivers, C.N.L., Peters, J., Houtkamp, R., and Roelfsema, P.R. (2011). Different states in visual working memory: When it guides attention and when it does not. Trends Cogn. Sci. 15, 327–334.

Panichello, M.F., and Buschman, T.J. (2021). Shared mechanisms underlie the control of working memory and attention. Nature 4.

Peirce, J.W. (2007). PsychoPy-Psychophysics software in Python. J. Neurosci. Methods 162, 8–13.

Pinsk, M.A., Doniger, G.M., and Kastner, S. (2004). Push-pull mechanism of selective attention in human extrastriate cortex. J. Neurophysiol. 92, 622–629.

Pollmann, S., Weidner, R., Müller, H.J., and von Cramon, D.Y. (2006). Neural correlates of visual dimension weighting. Vis. Cogn. 14, 877–897.

Press, C., Kok, P., and Yon, D. (2019). The Perceptual Prediction Paradox. Trends Cogn. Sci. 24, 12–24.

Press, C., Kok, P., and Yon, D. (2020). Learning to Perceive and Perceiving to Learn. Trends Cogn. Sci. 24, 260–261.

Rajsic, J., and Woodman, G.F. (2019). Do we remember templates better so that we can reject distractors better ? Attention, Perception, Psychophys.

Rajsic, J., Ouslis, N.E., Wilson, D.E., and Pratt, J. (2017). Looking sharp: Becoming a search template boosts precision and stability in visual working memory. Attention, Perception, Psychophys. 79, 1643–1651.

Rajsic, J., Carlisle, N.B., and Woodman, G.F. (2020). What not to look for: Electrophysiological evidence that searchers prefer positive templates. Neuropsychologia 140, 107376.

Rao, R.P., and Ballard, D.H. (1999). Hierarchical Predictive Coding Model Hierarchical Predictive Coding of Natural Images. Nat. Neurosci. 2, 79–87.

Reeder, R.R., and Peelen, M. (2013). The contents of the search template for category-level search in natural scenes. J. Vis. 13, 1–13.

Reeder, R.R., Hanke, M., and Pollmann, S. (2017). Task relevance modulates the cortical representation of feature conjunctions in the target template. Sci. Rep. 7.

Ruff, C.C., Blankenburg, F., Bjoertomt, O., Bestmann, S., Freeman, E., Haynes, J.D., Rees, G., Josephs, O., Deichmann, R., and Driver, J. (2006). Concurrent TMS-fMRI and Psychophysics Reveal Frontal Influences on Human Retinotopic Visual Cortex. Curr. Biol. 16, 1479–1488.

Scheibehenne, B., Jamil, T., and Wagenmakers, E.-J. (2016). Bayesian Evidence Synthesis Can Reconcile Seemingly Inconsistent Results. Psychol. Sci. 27, 1043–1046.

Schmidt, J., and Zelinsky, G.J. (2009). Search guidance is proportional to the categorical specificity of a target cue. Q. J. Exp. Psychol. 62, 1904–1914.

Serences, J.T., Saproo, S., Scolari, M., Ho, T., and Muftuler, L.T. (2009). Estimating the influence of attention on population codes in human visual cortex using voxel-based tuning functions. Neuroimage 44, 223–231.

Summerfield, C., and Egner, T. (2009). Expectation (and attention) in visual cognition. Trends Cogn. Sci. 13, 403–409.

Summerfield, C., Behrens, T.E., and Koechlin, E. (2011). Perceptual classification in a rapidly-changing environment. Neuron 71, 725–736.

Töllner, T., Zehetleitner, M., Gramann, K., and Müller, H.J. (2010). Top-down weighting of visual dimensions: Behavioral and electrophysiological evidence. Vision Res. 50, 1372–1381.

Tomov, M.S., Truong, V.Q., Hundia, R.A., and Gershman, S.J. (2020). Dissociable neural correlates of uncertainty underlie different exploration strategies. Nat. Commun. 11.

Trapp, S., Parr, T., Friston, K., and Schröger, E. (2021). The Predictive Brain Must Have a Limitation in Short-Term Memory Capacity. Curr. Dir. Psychol. Sci.

Treue, S., and Trujillo, J.C.M. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. Nature 399, 575–579.

Vickery, T.J., King, L.W., and Jiang, Y. (2005). Setting up the target template in visual search. J. Vis. 5, 81–92.

Vogel, E.K., Woodman, G.F., and Luck, S.J. (2001). Storage of features, conjunctions, and objects in visual working memory. J. Exp. Psychol. Hum. Percept. Perform. 27, 92–114.

Vossel, S., Thiel, C.M., and 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.

Wei, P., Yu, H., Müller, H.J., Pollmann, S., and Zhou, X. (2018). Differential brain mechanisms for processing distracting information in task-relevant and -irrelevant dimensions in visual search. Hum. Brain Mapp. 40, 110–124.

Witkowski, P., and Geng, J.J. (2019). Learned feature variance is encoded in the target template and drives visual search. Vis. Cogn. 1–15.

Witkowski, P.P., Park, S.A., and Boorman, E.D. (2022). Neural mechanisms of credit assignment for inferred relationships in a structured world. Neuron 2021.12.22.473879.

Wolfe, J.M. (2020). Visual Search: How Do We Find What We Are Looking For? Annu. Rev. Vis. Sci. 6, 1–24.

Wolfe, J.M. (2021). Guided Search 6.0: An updated model of visual search (Psychonomic Bulletin & Review).

Wolfe, J.M., and Horowitz, T.S. (2017). Five factors that guide attention in visual search. Nat. Hum. Behav. 1, 1–8.

Wolfe, J.M., Cave, K.R., and Franzel, S.L. (1989). Guided Search: An Alternative to the Feature Integration Model for Visual Search. J. Exp. Psychol. Hum. Percept. Perform. 15, 419–433.

Wolfe, J.M., Friedman-Hill, S.R., Stewart, M.I., and O'Connell, K.M. (1992). The Role of Categorization in Visual Search for Orientation. J. Exp. Psychol. Hum. Percept. Perform. 18, 34–49.

Won, B.-Y., Venkatesh, A., Witkowski, P.P., Banh, T., and Geng, J.J. (2021). Memory precision for salient distractors decreases with learned suppression. Psychon. Bull. Rev.

Woodman, G.F., and Vogel, E.K. (2008). Selective storage and maintenance of an object's features in visual working memory. Psychon. Bull. Rev. 15, 223–229.

Woodman, G.F., Carlisle, N.B., and Reinhart, R.M.G. (2013). Where do we store the memory representations that guide attention? J. Vis. 13, 1–17.

Yang, H., and Zelinsky, G.J. (2009). Visual search is guided to categorically-defined targets. Vision Res. 49, 2095–2103.

Yu, X., and Geng, J.J. (2019). The Attentional Template is Shifted and Asymmetrically Sharpened by Distractor Context. J. Exp. Psychol. Hum. Percept. Perform. 45, 1–37.

Yu, X., Hanks, T.D., and Geng, J.J. (2022). Attentional Guidance and Match Decisions Rely on Different Template Information During Visual Search. Psychol. Sci. 33, 105–120.

Yu, X., Zhou, Z., Becker, S.I., Boettcher, S.E.P., and Geng, J.J. (2023). Good-enough attentional guidance. Trends Cogn. Sci. xx, 1–13.

Zanto, T.P., Rubens, M.T., Bollinger, J., and Gazzaley, A. (2010). Top-down modulation of visual feature processing: The role of the inferior frontal junction. Neuroimage 53, 736–745.

Zhang, W., and Luck, S.J. (2009). Feature-based attention modulates feedforward visual processing. Nat. Neurosci. 12, 24–25.

Zhang, X., Mlynaryk, N., Ahmed, S., Japee, S., and Ungerleider, L.G. (2018). The role of inferior frontal junction in controlling the spatially global effect of feature-based attention in human visual areas. PLoS Biol. 16, 1–28.

Zhou, Z., and Geng, J.J. (2023). Learned associations serve as target proxies during difficult but not easy 1 visual search. PsyArXiv.

Zivony, A., and Lamy, D. (2016). The role of dimension relevance in features' access to response-selection mechanisms. J. Exp. Psychol. Hum. Percept. Perform. 42, 1873–1885.

Chapter 3: Attentional priority is determined by predicted feature distributions

Introduction

Observers have access to an immense amount of sensory information in natural environments but only a subset is relevant for behavior. Optimizing behavior in such environments requires knowledge of which information to prioritize and which to ignore (Gottlieb, 2018; Niv et al., 2015; Nobre and Stokes, 2019; Wolfe, 2021; Wolfe and Horowitz, 2017). However, which information to prioritize is often not immediately obvious. For example, when searching for a friend in the park you are unlikely to know exactly how she will look that day, or the features that will distinguish her from others in the crowd because her clothes vary from day to day. One way to address this uncertainty is to make predictions about the distribution of features your friend *could have* rather than trying to match your friend to any single instance of her that you have previously seen (i.e., a single memory). If she normally wears warm colored shirts, prioritizing only a specific shirt you have seen her wear before (e.g., a red t-shirt) is likely to be incorrect since her shirt color might be any color between yellow and pink. However, maintaining a predictive distribution across a set of colors she is likely to wear can help determine attentional priority in situations where a single precise memory is unlikely to be correct. Such scenarios are commonplace in daily life. Yet, previous studies have only measured attentional priority as a function of memory precision for known target features (e.g., a single red t-shirt), rather than predictions about the distribution of ways the target *could look* (e.g., all warm colored shirts). The current experiments test if the inherent uncertainty of target objects informs attentional priority, and whether this uncertainty is encoded in predictions about possible upcoming target features.

Theories of attention use the concept of the target template to describe how the weighting of task-relevant information is implemented in service of attentional priority (Chelazzi et al., 1993; Duncan and Humphreys, 1989; Mehrpour et al., 2020; Olivers et al., 2006; Woodman et al., 2013; Zhang and Luck, 2009). Target templates have historically been characterized as storing a single target's remembered identity (Olivers et al., 2011; Vickery et al., 2005). Many studies have found that more precise memories produce faster and more accurate search (Hout and Goldinger, 2015; Malcolm and Henderson, 2009; Schmidt and Zelinsky, 2009; Yang and Zelinsky, 2009). For example, Hout and Goldinger (2015) gave participants a single target that was either identical to the cue, a rotated version of the cue, or a new exemplar of the same category. The authors found that search was most efficient when the cue was an exact match for the target. However, most real-world targets have variable features that change over time, creating uncertainty about their exact features. Such targets may be missed by highly precise templates because the actual features of the target will often be a poor match to the target template.

One way attentional systems might overcome this limitation and mitigate the effects of uncertainty is by representing a predicted distribution of target features within the target template (Knill and Pouget, 2004; Ma and Jazayeri, 2014); that is, representing a distribution of expected features rather than any single feature. Theories of perception consider such probabilistic representations of sensory data critical to cognition (Clark, 2013; Friston, 2010; Trapp et al., 2021), and show that perception is often linked to a prior belief about what we are likely to see in the world (Kok et al., 2012; Press et al., 2019, 2020). Indeed, probabilistic expectations are thought to interact with attentional mechanisms to guide spatial attention to likely target locations (Chun and Jiang, 1998; Foley et al., 2017; Geng and Behrmann, 2005; Vossel et al., 2006). Similarly, others have shown that probabilistic expectations about distractor features are used to bias attention away from possible distractors in an upcoming task (Chetverikov et al., 2017a, 2020; Huang et al., 2021; Lee and Geng, 2020; Won et al., 2021). These studies show that expectations about the upcoming search environment increase the efficiency of attentional processing. Despite this, few studies have asked whether the target template encodes predicted target feature distributions to guide attention when the features of the target are uncertain.

One benefit of maintaining predictive representations of target feature distributions is that this information can be used to reduce uncertainty by setting attentional priority for features with lower variability. For example, if the friend you are searching for wears only warm colored shirts, but pants of any color, prioritizing the low-variability distribution of shirt colors over the high-variability distribution of pant colors will maximize predictive information in the target template (i.e., minimize target uncertainty). Target features encoded into working memory are known to be remarkably sensitive to the task-relevance of information (Becker et al., 2010; Boettcher et al., 2020; Geng and Witkowski, 2019; Nobre and Stokes, 2019; Rajsic et al., 2020), and it has been shown that object features are flexibly removed or compressed so that only the most informative features are used to guide attention (Bravo and Farid, 2012, 2016; Kerzel and Huynh Cong, 2021; Reeder et al., 2017; Woodman and Vogel, 2008; Zivony and Lamy, 2016). Previous work has also shown that when sensory information differs in reliability, perception is optimally weighted toward the most reliable cue (Fetsch et al., 2012; Hillis et al., 2004), resulting in a “push-pull” dynamic for attentional priority (Foley et al., 2017; Kozyrev et al., 2019; Mehrpour et al., 2020; Pinsk et al., 2004). However, it is still unknown whether feature-based attentional biases depend on probabilistic representations of the expected target, or merely on the precision of memories for target features. Linking predictions about target feature distributions to attentional priority would show that attentional biases are inherently connected to predictive processes that minimize our uncertainty about the upcoming task, rather than being only a consequence of the quality of memories from previously seen targets.

In the current studies we investigated the role of predictive representations in the target template when the identity of upcoming targets is uncertain. We did this by using a visual search paradigm in which target features were stochastically determined, creating uncertainty in the target's appearance. Our main hypothesis was that observers would use predictions about the variability of upcoming target features to set attentional priority and encode these predictions into the target template prior to search. To anticipate, we found that attentional priority is set by explicit predictions about the variability of upcoming target features, and not by the precision of memory for specific cues. Our results suggest that predictive representations are a key element in determining attentional priority and illuminate how attentional systems mitigate feature uncertainty inherent to the natural world.

Experiment 1

In Experiment 1, we tested if predictions about the distribution of possible target features underlie changes in attentional priority. We used a visual search task in which target features were drawn from distributions that change in standard deviation (STDEV) over time without notice, requiring participants to actively maintain predictions about the distribution of possible features to know which feature dimension should be prioritized on each trial. If participants use these predictions to set attentional priority, we should see that response times (RT) to select targets is modulated by the STDEV of underlying target feature distributions.

Methods

We recruited a sample of 60 participants (self-reported females=47, males=13; mean age =19.05) from the University of California, Davis SONA system. This sample size was calculated to have 90% power to detect the desired RT effect based on simulations from pilot data. All procedures were approved by the University of California, Davis Internal Review Board (IRB). The experiment was completed online using the Testable platform (www.testable.org), and all participants were required to calibrate their screens before beginning the experiment so that search stimuli were a constant size (approximately 2.5 degrees of visual angle). Participants did so by holding a standard credit card up to the screen and adjusting a line so that it matched the size of the card.

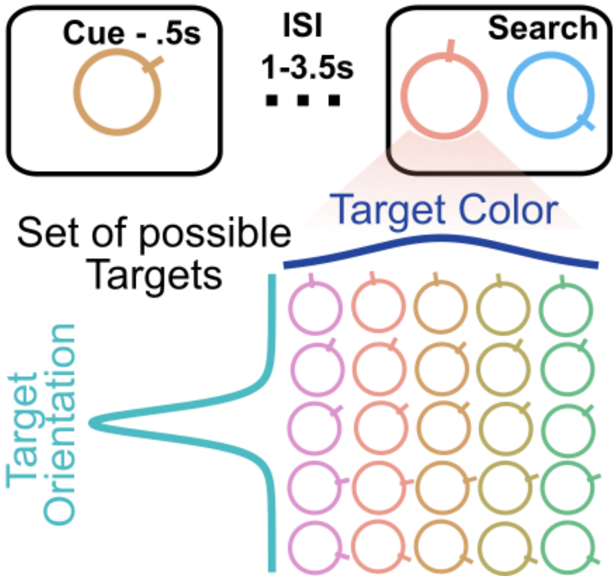

Figure 3.1) Task schematic for Experiment 1 with example trial where orientation is the low-variability feature. Each trial began with the presentation of a cue, which lasted 500ms. After a short delay, participants were asked to search for the target most similar to the cue as quickly as possible and report its location. Stimuli were visible for 500ms. The set of targets illustrate possible target features for one example cue with particular color and orientation features. The actual distribution of features was continuous in 1-degree units over each feature space. The “Target Color” and “Target Orientation” distribution represent the relative likelihood with which a target object would have each feature in a case where color variability is high, and orientation variability is low, but these distributions changed overtime (see methods).

At the beginning of each trial participants were given a “cue” object composed of a colored circle with a bar that indicated orientation (Fig 3.1). On each trial, the cue was randomly selected to have one of three colors (RGB values: (194.9, 156.4, 226.8), (210.2, 160.9, 104.0), (56.9, 190.3, 180.7)) and one of three orientations (35, 155, 275-degrees) for a total of 9 possible combinations. The same cues were used throughout the experiment, irrespective of the current STDEV. Cues were shown for 500ms and followed by a 1000-3500ms interstimulus-interval (ISI). The ISI was followed by a visual search display during which participants were asked to report the location of the target as quickly and accurately as possible by pressing “j” if it was on the left and “k” if it was on the right. The target was defined as the item that was *most similar* to the cue. Calibration for color on individual monitors was not possible. The search display was visible for 500ms, after which the stimuli disappeared, but participants had an unlimited time to respond. Each trial ended with feedback about whether or not they correctly identified the target (“Correct”, “Incorrect”), followed by a randomly jittered 500-2500ms intertrial-interval. Each participant completed 160 trials in total during the experiment.

Both target orientation and target color were drawn from t-distributions centered on the cued feature (Mean=0) on each trial, but the variability of the distribution randomly changed without warning throughout the experiment (Fig 3.2a). Thus, while there were only nine possible cue values, there were over 180 possible targets. However, each feature had a particular variability that was consistent across cues and could be utilized to reduce the uncertainty of the upcoming target features. Target distributions were bounded such that the cue-to-target feature difference could never be greater than 90-degrees to minimize correlations between target and distractor values. Since distractors were required to be more dissimilar from the cue than the target, trials with large cue-to-target differences were also more likely to have small target-to-distractor differences, making it difficult to dissociate their influence on the search process. However, bounding the target distribution minimized these correlations (Pearson r=.05) allowing us to capture the unique influence of target variability and distractor features on search.

At seven predetermined “change points” during the experiment, the variability of orientation and color abruptly shifted. These “change points” in the STDEV of both target features occurred every 25 to 50 trials and followed a fixed schedule: STDEV=25,10,65,45,40,25 for orientation and STDEV=40,10,20,15,45,60 for color. Values of STDEV for each feature distribution were chosen such that they produced noticeable shifts in feature variability over time based on simulations and independent pilot experiments. All participants experienced the same transitions in STDEV. Notably, participants were not given any prior information about the variability of each feature, how the variability would change, nor which feature would be most similar to the cue on the upcoming trial. Thus, participants had to predict the distribution of possible target features in the upcoming trial based on recent experiences in order to determine which feature would be the most useful to prioritize in order to maximize the efficiency of target selection.

Finally, distractors in each trial were randomly selected from a uniform distribution between -179 degrees and 180 degrees away from the cue. The only constraint on the distractor features were that distractors must be at least 20 degrees more distant from the cue than the target. All stimuli that appeared on the screen had colors taken from a CIE color space color wheel, which contained 360 colors of equal luminance separated by one degree units (Bae, Olkkonen, Allred, & Flombaum, 2015).

Bayesian Modeling of Predicted Variability

To understand if participants used predictions about the distribution of target features to set attentional priority, we created a Bayesian learning model that allowed us to estimate feature predictions based on previously seen targets, without assuming perfect knowledge of the new distribution at each change point. We modeled predictions about the distribution of target features as beliefs about the STDEV of the target distribution for the upcoming trial, given the history of observed targets. We used a normative Bayesian learning model (O'Reilly et al., 2013; Summerfield et al., 2011). On each trial *t*, the probability of observing a specific target feature (for orientation and color separately) *y* given the variance of the feature distribution was written as:

$$(y_t | \sigma_t) \sim N(\mu = 0, \sigma_t) \tag{1}$$

where μ is the mean and σ is the STDEV of the feature distribution. The range of σ was bounded between STDEV=2 and STDEV=80. Each new target revealed information that participants could use to update their beliefs about the STDEV of each feature for the next trial. The true STDEV of each distribution (Fig 3.2a, gray lines) stayed constant until a change point occurred. Thus, prior beliefs in the STDEV of the underlying distribution were modeled as the following:

$$p(\sigma_t | \sigma_{t-1}, J_t) = \begin{cases} \delta(\sigma_t - \sigma_{t-1}) & J_t = 0 \\ U(2, 80) & J_t = 1 \end{cases} \tag{2}$$

This equation states that the estimated standard deviation of the feature distribution (σ) on trial *t* stayed the same as the previous trial if a change point was not estimated to have occurred (J_t= 0) or switched to any value between 2 and 80 with uniform probability if a change point was estimated to have occurred (J_t=1). This allows the model to down-weight the influence of trials after a change point occurs. Change points were estimated to have occurred by ν, based on the likelihood that the current feature value was more likely to have been sampled from a distribution with a different standard deviation:

$$v = p(J_t = 1) \tag{3}$$

Where $p(J_t = 1)$ is the marginalized probability that σ has changed, over all values of σ. The conditional prior for σ_t can be written as:

$$p(\sigma_t | \sigma_{t-1}, \nu) = (1 - \nu)[\ \delta(\sigma_t - \sigma_{t-1})] + \nu U(2{,}80) \tag{4}$$

Importantly, integrating the change probability, v, allows the model to quickly switch beliefs about the STDEV of each feature distribution when the evidence changes dramatically. At these same points, participants would be expected to rapidly change their beliefs given new observations. The generative model can then be inverted using Bayes rule and written as:

$$p(\sigma_t, v\ | y_{1:t}) = \frac{p(y_t|\sigma_t)p(\sigma_t, \nu | y_{t-1})}{\iint p(y_t|\sigma_t)p(\sigma_t, \nu | y_{t-1})\ d\sigma_t d\nu} \tag{5}$$

The model’s predicted variability estimates for color and orientation (Fig 3.2a) reflect an “optimal observer’s” prior belief in the STDEV of each feature distribution *on the next trial* (*t+1*). That is, each estimate of the feature distribution contains information learned from previous trials and are applied *prior* to the next trial before the true target features have been observed.

We then used these estimates to compute the *difference in predicted variability* of each feature. The *difference in predicted variability* is positive when color is more predictive, and it is negative when orientation is more predictive. If participant RT follows the *difference in predicted variability*, it would mean that they use predictions about the distribution of possible target features to weight attentional priority towards the feature with lower variability on the upcoming trial.

Regression Modeling of Response Time Data

To estimate the effect of feature predictions on attentional priority, we used a linear regression model to test how *the difference in predicted variability*, derived from the model, moderates participant RT search costs for targets over-and-above cue-to-target similarity alone. Specifically, we hypothesized that the effect of cue-to-target similarity in each feature dimension would be modulated by expectations about which of the two features was currently less variable (and therefore more reliable). For example, the expected RT cost for a target with an orientation that is 45-degrees from the cue is expected to be large when the *difference in predicted variability* indicates that orientation is the less variable feature (i.e., orientation is more reliable); in contrast, the RT cost for the same target orientation should be smaller when the *difference in predicted variability* is positive, indicating that color is the less variable dimension. This outcome would lead to a statistical interaction in RT between the *cue-to-target difference for each feature dimension and the Bayesian model generated estimates of the difference in predicted variability*. Such an interaction would suggest that participants maintain learned predictions about the distribution of each feature and use those predications to weight attentional priority during search.

We tested this hypothesis by fitting RT data from trials with correct responses to a gamma-distributed hierarchical regression model (Lo and Andrews, 2015), using the lme4 package in R. Trials with incorrect responses or RTs greater than 3 times the interquartile range

(IQR) from the median RT were excluded from the analysis (Mean=14 trials total per participant). The regression model had one regressor for *orientation cue-to-target difference* and one for *color cue-to-target difference*. Each regressor reflects the number of milliseconds added to RT for each degree of difference between the cue and the target (ms/degree). In other words, each regressor reflects the slope of RT increase as the target becomes more dissimilar from the cue along one feature dimension. Importantly, there was also a regressor for the *difference in predicted variability* from the Bayesian model (see above). This regressor represents a prediction of how much more reliable orientation is than color, or vice versa. If observer's are sensitive to this difference in predicted variability of the upcoming target, attentional priority should be biased towards the feature with lower variability (Fetsch et al., 2012; Foley et al., 2017; Hillis et al., 2004). We therefore hypothesized that the *difference in predicted variability* would interact with *orientation cue-to-target difference* and with *color cue-to-target difference.* Because of the way in which the *difference in predicted variability* is calculated (see above), the sign of the interaction will be negative for orientation and positive for color. Nevertheless, both interactions reflect change in RT (ms) per degree of *cue-to-target* similarity as a function of the degree of *difference in predicted variability* (ms/deg/deg). This will only occur if participants maintain a representation of the predicted distribution of both target features. Finally, we included two regressors for the target-to-distractor difference in each feature dimension, *orientation target-to-distractor difference* and *color target-to-distractor difference* (Duncan and Humphreys, 1989; Wolfe and Horowitz, 2017). The random effects structure included random intercepts for each participant and random by-subject slopes for each of the fixed effects explained above.

All significance testing was done using likelihood-ratio tests between the full model and models with the relevant fixed effect removed, as is appropriate for hierarchical models (Luke, 2017). For example, the significance of the interaction between *orientation cue-to-target difference* and *difference in predicted variability* was tested by fitting a model with and without the fixed-effect of the interaction, then comparing the difference in likelihoods. The HDI (Bayesian confidence interval) of each parameter was determined by fitting the full model in PyMC3 with uninformative priors over parameter values.

Results

Mean accuracy was high across all participants (Mean=.935, SD=.045). We observed the hypothesized interaction between *orientation cue-to-target difference* and *difference in predicted variability* (β =-.024ms/deg/deg, $\chi2(1)=11.76$, $p<.001$, HDI 95% [-.024,-.001]) (Fig 3.2C). Greater *orientation cue-to-target* differences led to longer RTs as expected, but this effect proportionately decreased as target orientation became more variable relative to target color. Similarly, we found a positive interaction between *color cue-to-target difference* and *difference in predicted variability* (β =.040ms/deg/deg difference, $\chi2(1)=15.35$, $p<.001$, HDI 95% [.026, .058]). This shows that participants responded slower with increasing *color cue-to-target differences* when color was less variable than orientation. Together, these results indicate that predictions about the distribution of possible target features were used to bias attentional priority. When the predicted variability in one dimension (e.g., orientation) was higher than the other (e.g., color), dissimilar targets in that dimension (e.g., orientation) did not slow RT as much as when predicted variability is lower than the other dimension (e.g., color) and the dissimilar target feature was more surprising.

In addition to looking at cue-to-target similarity, we also looked at the effect of target-to-distractor similarity on search RT. There was a significant interaction between *difference in predicted variability* and *orientation target-to-distractor difference* (β =.006ms/deg/deg difference, χ2(1)= 11.59, p<.001, HDI 95% [.004,.011]), and a weaker opposite interaction with *color target-to-distractor difference* (β =-.004ms/deg, χ2(1)= 6.25, p=.012, HDI 95% [-.006, .003]). RTs were longer when distractors had features close to the target, but this effect was bigger in the dimension with lower predicted variability. That is, distractor features interfered with search more when they were similar to the less variable target feature, presumably because that dimension was prioritized for finding the target.

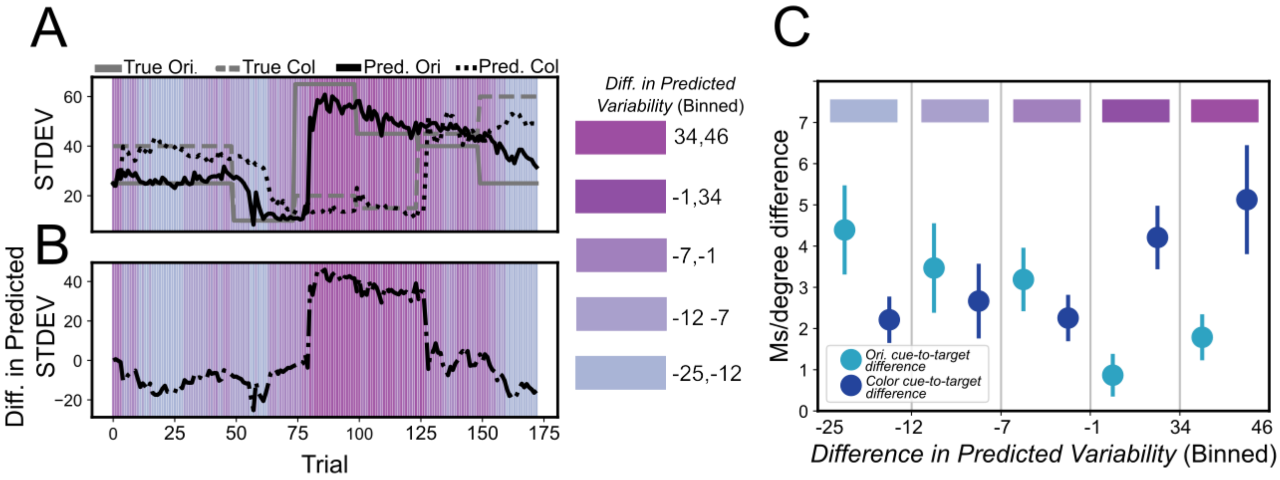

Figure 3.2: A) Predicted variability estimates for each feature were derived from a Bayesian learning model that tracked changes in the STDEV of each target feature distribution given previously seen targets (Black lines). These estimates of STDEV were the same across participants. Estimated STDEV per trial are overlayed on the true STDEV of each feature distribution (Grey lines). Background hue (blue to purple) indicates the difference in predicted variability. Blue colors indicate trials when orientation STDEV is lower than color (i.e., orientation has lower predicted variability); purple colors indicate trials when color STDEV is lower (i.e., color has lower predicted variability). B) Same as A but shows the *difference in predicted variability* across trials used in the regression model. C) Illustration of the interaction between cue-to-target difference and difference in predicted variability for orientation (dark blue) and color (teal). Color bars at the top indicate the trials in A from which the data were taken. Bins were created such that there were approximately equal trial numbers in each bin per participant. The Y-axis

indexes the change in RT (in milliseconds) per degree of cue-to-target dissimilarity for each feature. Error bars represent a 95% confidence interval.

Experiments 2a and 2b

Experiment 1 found that observers readily used predictions about the distribution of possible target features to bias attentional priority towards the low-variability feature. However, it is still unclear how these predicted distributions are encoded by the target template. One possibility is that the target template holds only a single memory of the most likely target and adapts the precision of this working memory representation to reflect the predicted STDEV (Grubert and Eimer, 2013; Hout and Goldinger, 2015; Olivers et al., 2011). For example, the memory representation of the current cue may degrade if the participants do not expect the target to match perfectly, leading to less precise memories for the cue as target variability increases. If true, this would mean that the expected variability of a distribution is not encoded directly but interacts with working memory precision to produce behavior. Another possibility is that observers directly encode a distribution of predicted likelihoods for each feature that is separate from the memory of the cue (Knill and Pouget, 2004; Ma and Jazayeri, 2014). If true, the memory representation may remain precise but there would be additional information regarding variability in the predicted target. Such a result would be evidence for the hypothesis that predictions are distinct from memories of the cue and that attentional priority is more related to predictions about target variability than memory precision. We test these alternative possibilities in Experiments 2a and 2b using probe trials interleaved amongst visual search trials.

Methods

We recruited two independent samples of 240 participants for Experiments 2a (self-reported females=185, self-reported males=55; mean age=20.0 years) and 2b (self-reported females=184, self-reported males=56, mean age=19.67 years) from the University of California, Davis undergraduate SONA system. This sample size was calculated to have 90% power to detect the desired RT effect in each experiment alone, based on simulations from pilot data. All procedures were approved by the University of California, Davis IRB prior to the study.

All methods were the same as in Experiment 1 with a few key differences. To simplify the design, we only manipulated the variability of orientation, while color variability was kept constant. Participants in each experiment were split into three *predicted variability* conditions (N=60 per condition per experiment) defined by the STDEV of the target orientation distribution (Fig 3.3AB). As in Experiment 1, all cues had a randomly selected combination of color (RGB values: (194.9, 156.4, 226.8), (210.2, 160.9, 104.0), or (56.9, 190.3, 180.7)) and orientation (35, 155, or 275-degrees) and this was true for all groups. Target orientations were drawn from t-distributions centered on the cue (mean=0) and had a standard deviation of 10, 25 or 40 degrees. These corresponded to low-variability (STDEV=10), medium-variability (STDEV=25) and high-variability (STDEV=40) conditions, respectively. Splitting participants into groups with a stable STDEV allowed us to probe representations of specific learned distributions more accurately. In all variability conditions, color was drawn from a high-variability distribution (STDEV=55). Finally, all distractor features were sampled from a uniform distribution but were restricted to values that were at least 20 degrees more distant from the cue than the target.

An additional 60 participants per experiment were assigned to a *control condition*, in which both color and orientation were drawn from identical low-variability distributions (STDEV=10). This established a “baseline” in which expectations of the feature distributions were matched and any differences in memory or prediction probe responses could only reflect pre-existing differences in each dimension. The data from this control group were used to normalize probe responses from the other conditions following the procedures in Witkowski and Geng (2019). Specifically, we calculated the difference of each participants mean responses from the control group responses, divided by the standard deviation in control responses. This gave us a normalized measure of the change in working memory precision or predictions in response to the variability conditions.

In Experiment 2a, 14 memory probe trials (7 orientation/7 color) were interleaved between 40 search trials, following 54 initial “training” visual search trials. On memory probe trials, participants were shown either a color wheel or an orientation wheel. Each wheel contained 24 feature segments of different colors or orientations, respectively, split into 15-degree bins. Each feature segment was associated with a number, and participants were asked to report the number next to the remembered cue feature. In Experiment 2b, the 14 interleaved prediction probe trials asked participants to rate the likelihood that a subsequent object would be the target. Participants made their responses on a 1 to 9 scale with 1 being “very unlikely” and 9 being “very likely”.

Probe objects spanned the distribution of all possible cue-to-target differences (-90 to 90 degrees) in 15-degree intervals, with color and orientation being randomly combined such that there was no correlation between features. Importantly, although a single participant only saw 14 probes, all the possible cue-to-target differences for color and orientation were probed across participants. Both probe types were then fit to Gaussian functions using maximum likelihood to make these data commensurate with the stimuli generated in search. For prediction probes, we allowed amplitude of the Gaussian function (i.e., the height of the peak) to range from 1 to 9, so that ratings about the likelihood of target features could be directly mapped to predictions about the distribution of possible features.

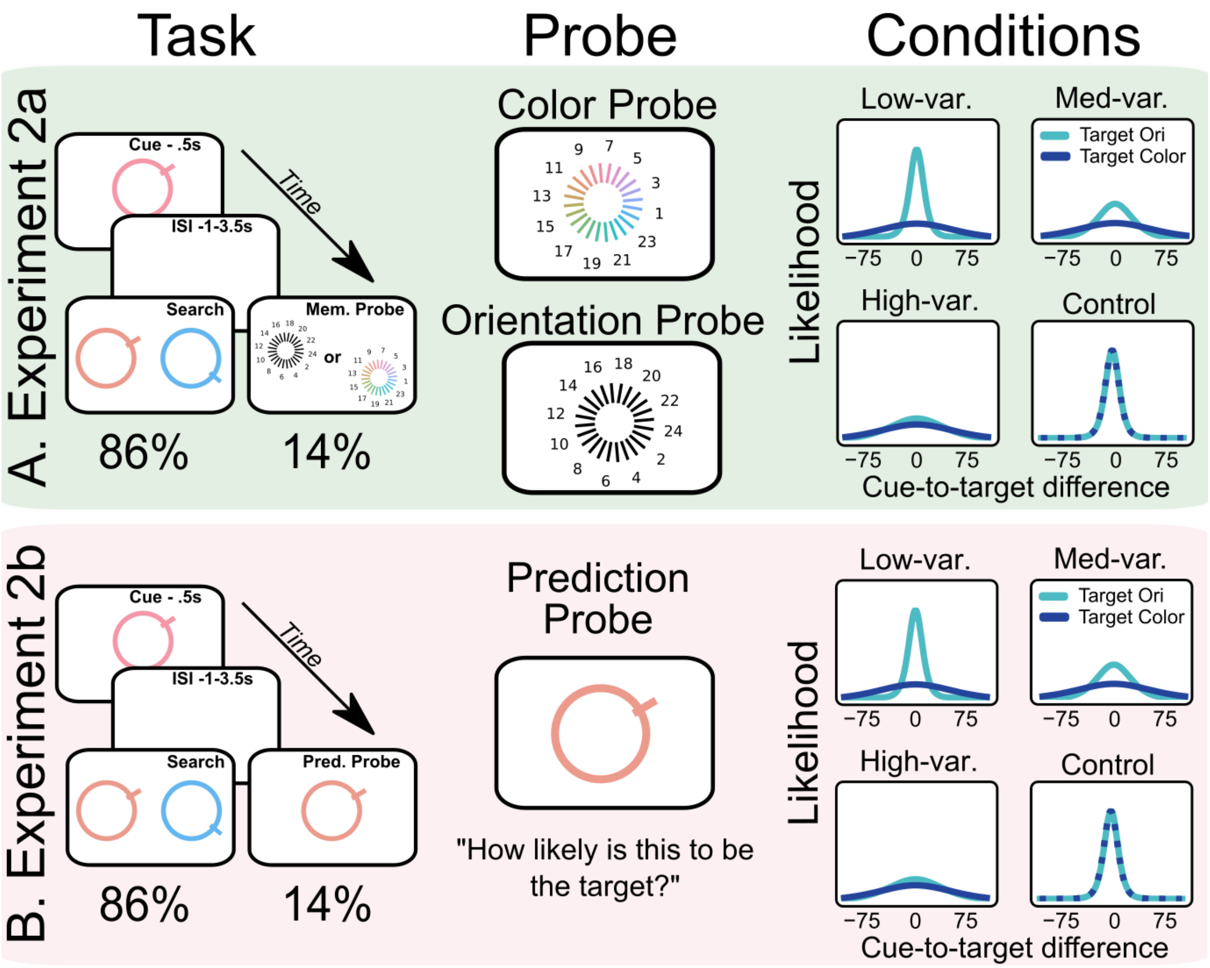

Figure 3.3: A) Task schematic for Experiment 2a. Each trial began with the presentation of a cue, which lasted 500ms. On each trial, participants were asked to either search for the target object (86% of trials) or complete a memory probe task (14%) in which they reported the remembered features of the cue item by pressing the number associated with the remembered feature. Participants were split into 3 *predicted variability* conditions and 1 *control condition* defined by the distribution of target features (see Figure 3.1 for illustration of search target variability). B) Same as A but for Experiment 2b. In this experiment, participants completed a "prediction probe trial" by reporting the predicted likelihood of a possible stimulus being the target.

Results

Search behavior is determined by the predicted variability of target features

We used the same regression model to analyze RT data in Experiments 2a and 2b as in Experiment 1, except that the three *predicted variability* conditions were manipulated between-subjects. We expected the RT data to be identical given that in these experiments the visual search trials were identical. Accuracy was overall high in both experiments (Experiment 2a Mean=0.918, SD=0.062; Experiment 2b Mean=0.921, SD=0.060), but significantly decreased as orientation variability increased (Exp2a: $\chi 2(1)=37.11$, $p<.001$ HDI 95% [-0.507, -0.288]; Exp2b: $\chi 2(1)=14.36$, $p<.001$ HDI 95% [-0.405,-0.127]); this was consistent with the idea that the task became more difficult as orientation variability increased (and color variability did not change). Again, we excluded trials where responses were incorrect or greater than 3 IQRs from the median RT (mean=9 trials per subject in both experiments).

Consistent with Experiment 1, the RT results from Experiment 2a showed a negative interaction between *orientation cue-to-target difference* and *predicted variability* (β =-.98ms/deg/group, $\chi 2(1)=7.49$, p=0.006, HDI 95% [-1.924 ,-0.769], Fig 3.4B). The interaction was due to *orientation cue-to-target difference* having a weaker effect on RT as the variability of the feature distribution increased. We found no significant interaction between *color cue-to-target difference* and *predicted variability* (beta=.18ms/deg, $\chi 2(1)=1.65$, p=0.20, HDI 95% [-0.064,0.31]), although the sign of the interaction was the same as in experiments 1 and 2b (see below). Finally, the model showed a significant effect of *orientation target-to-distractor difference* on search time (β =-.11ms/deg, $\chi 2(1)=4.82$, p=0.028, HDI 95% [-0.222 ,-0.056]), but not for color *target-to-distractor difference* (β =.09ms/deg, $\chi 2(1)=3.19$, p=0.074, HDI 95% [-0.058,0.111]). The effect of distractor similarity is consistent with the overall attentional prioritization of target orientation over color given that orientation was always less variable than color.

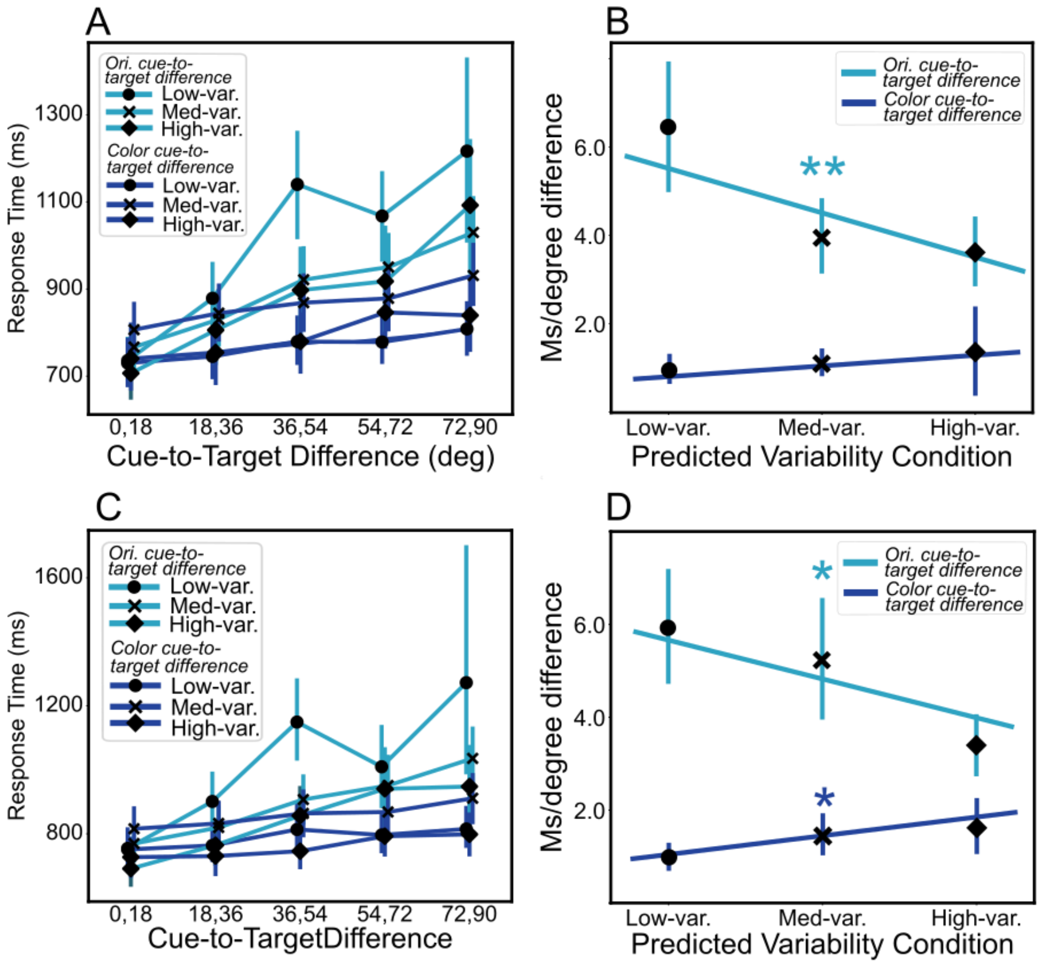

Figure 3.4: A) Mean response time for binned values of *cue-to-target difference* in Experiment 2a for color and orientation separately. Bins were created by splitting the data into equally portioned quantiles. Each line represents the increase in response times due to a specific cue-to-target difference for orientation (light blue) or color (dark blue) in each *predicted variability* condition (denoted by marker points). B) Points represent the slopes from A for each *predicted variability* condition with error bars representing the 95% confidence interval. C) same as A but for Experiment 2b. D) Same as B but for Experiment 2b. *=p<.05, ** = p<.01

Analysis of RT in Experiment 2b replicated these findings. There was a significant negative interaction between *orientation cue-to-target difference* and *predicted variability* (β=-.89ms/deg/grp, χ2(1)=5.26, p=0.022, HDI 95%[-1.93,-0.724], Fig 3.4D). There was also a significant interaction between *color cue-to-target difference* and *predicted variability*, but in the opposite direction (β =.35ms/deg/grp, χ2(1)=4.23, p=0.040, HDI 95% [0.017,0.437]), showing that as orientation variability increased, participants were more sensitive to the cue-to-target difference for color. Finally, the results replicated the significant effect of *orientation target-to-distractor difference* (β =-.16ms/deg, χ2(1)=7.83, p=0.005, HDI 95% [-0.242,-0.067]), and a null effect of color target-to-distractor difference (β =.01ms/deg, χ2(1)=.17, p=.718, HDI 95% [-0.123,0.051]). These results provide further evidence that predictions about target feature distributions are used to set attentional priority; specifically, attentional priority for orientation decreased when it was predicted to be more variable, and priority for color increased.

Predictions, but not memory for the cue, reflect knowledge about the learned target feature distribution

We next analyzed differences in memory precision (Experiment 2a) and predictions about target features (Experiment 2b) across variability conditions. First, we fit each participant's memory probe responses to a Gaussian function using maximum likelihood, separately for orientation and color. The resulting estimates of the standard deviation of memory responses were log-transformed to fit assumptions of a linear regression model. The model included regressors for *predicted variability* (low-variability, medium-variability, high-variability) and *feature dimension* (orientation, color), and their interaction. If the information about feature variability that is used to compute attentional priority is contained in the precision of cue memory, we should see an interaction between *predicted variability* and *feature dimension* similar to the patten of results on search trials. Specifically, the precision of memory for orientation features should decrease with variability, and the precision of memory for the color feature should increase or stay the same.

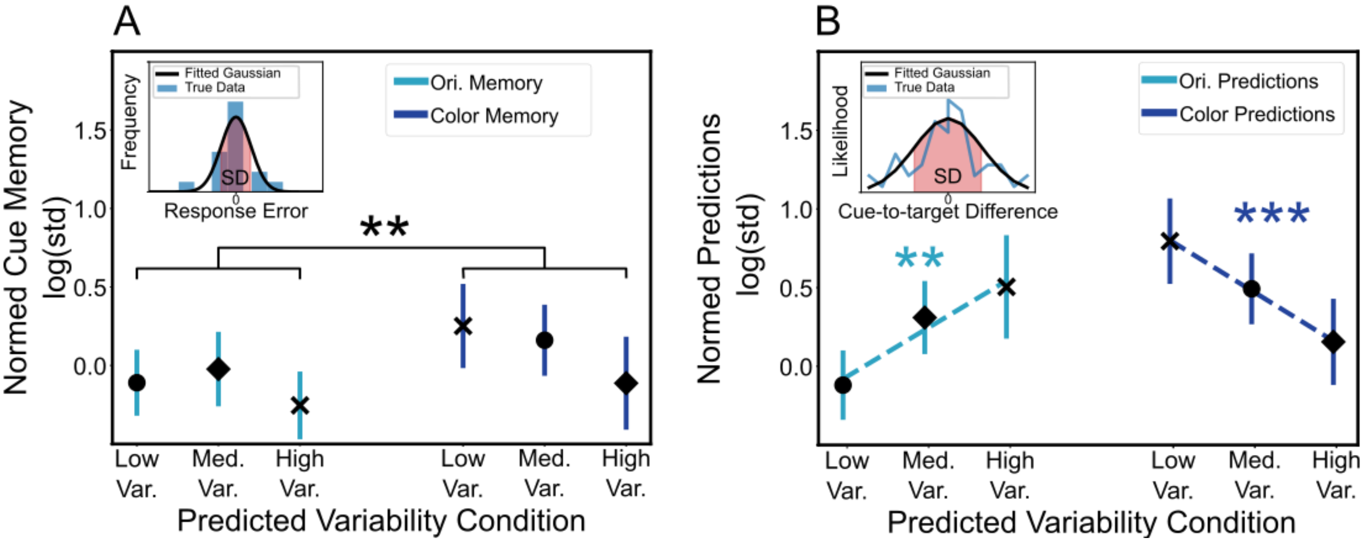

Figure 3.5: A) Fitted standard deviation of memory probe responses in each *predicted variability* condition for Experiment 2a. Error bars represent the 95% confidence interval around the mean of each group. Y-axis represents the difference in variability relative to the *control condition* B) Same as A but for prediction probe responses in Experiment 2b. The colored lines show the best-fit line for differences between variability conditions. ** = p<.01, *** = p<.001

The results revealed a marginal main effect of increasing memory precision across predicted variability conditions (χ2(1)=3.67, p=0.055, HDI 95% [-0.263, -0.006]), suggesting that memory for both features slightly increased in precision when both features were highly variable, and the search task was overall more difficult. There was also a main effect of feature dimension (χ2(1)=6.04, p=0.014, HDI 95% [0.027, 0.402]), suggesting that orientation was held more precisely compared to color (Fig 3.5A) across all three variability conditions. However, we found no interaction between variability condition and feature dimension (χ2(1) = .94, p=.33, HDI 95% [-0.313, 0.122], BF_{01}=14.28), suggesting no link in the precision of memory to the pattern of attentional priority during search. The main effect replicates previous findings, showing that more reliable (i.e., low variability) features are encoded in working memory more precisely compared to low reliability features (Witkowski and Geng, 2019), but goes further to suggest that this overall benefit is not sensitive to the difference in the predicted variability of each target feature.

In contrast to the memory probes, in Experiment 2b we observed changes in the distribution of responses to the prediction probes commensurate with the specific standard deviation in each *predicted variability* condition. The same regression for prediction probe responses produced a significant interaction between *predicted variability* and *feature dimension* ($\chi2(1) = 23.56$, $p<.001$, HDI 95% [-0.864, -0.384]). Follow up comparisons with simple regression showed the distribution of prediction responses for orientation features became wider as a function of increasing predicted variability condition ($\beta =.31$, $\chi2(1)=10.67$, $p=0.001$, HDI 95% [0.135, 0.503]), but became narrower for color ($\beta =-.32$, $\chi2(1)=11.84$, $p<.001$, HDI 95% [-0.494, -0.130]). We then tested whether this pattern was significantly different from Experiment 2a by including both sets of data into a single model and adding an interaction term between experiments. The results showed that the pattern in Experiment 2b was significantly different from the pattern shown in Experiment 2a ($\chi2(1)=9.41$, $p=.002$, HDI 95% [0.792, 0.221]) .

This pattern mirrors the pattern of attentional priority seen during search (Fig 3.5B), showing that manipulations of the underlying feature distribution change both predictions about the likely target feature and attentional priority for each feature during search. Even though the variability of color did not change across conditions, participants estimated the predicted variability as lower, presumably because perceived estimates of color variability were yoked to estimates of the variability of orientation, paralleling attentional priority. This result indicates that predictions about the likelihood of target features, but not memories of the cue, are used to set attentional priority during visual search.

Discussion

In the current experiments, we test whether observers learn the feature distributions of uncertain targets and we measure the consequences of this knowledge on attentional priority. We used a novel visual search paradigm in which target features were drawn from a distribution and participants learned the feature variability from observing targets over time. Search behavior showed that attentional priority for each feature tracked the predicted distribution of the feature distributions over time, such that low-variability features were prioritized over high-variability features in a push-pull manner. We then probed memory precision for cued features and explicit predictions about target features. Our results showed that the target-template dynamically encoded the variability as explicit predictions about target features, while memory for the cue remained static. These data suggest that observers use predictions about the likelihood of possible features to determine attentional priority when target features are uncertain.

The current studies provide novel evidence that attentional priority is tied to predictions about upcoming targets, and not memories of a cue. Previous research has generally focused on how the precision of memories determines attentional priority (Hout and Goldinger, 2015; Olivers et al., 2011; Schmidt and Zelinsky, 2009; Vickery et al., 2005) but have not examined how attentional priority is set when the target features are variable. Our results show that when faced with feature uncertainty, observers actively generate predictions about the distribution of possible features before the true target features can be observed. These predictions are then used to mitigate uncertainty by prioritizing features with the least variability. This predictive process is also highly flexible. When observers learn that the variability of the target has changed, they rapidly update priority settings for each feature. These findings directly connect feature distribution learning to information biases in the target-template and show how predictions are used by attentional systems to reduce uncertainty.

The results observed in search behavior also shed light on how target variability is translated into attention priority. Push-pull dynamics have been observed throughout the literature, showing that attending to one place or feature results in decreased encoding of an unattended feature (Kozyrev et al., 2019; Mehrpour et al., 2020). In Experiment 1, we observed that the attentional priority for both orientation and color were oppositely modulated by the *difference in the predicted variability* of each feature. Experiment 2b confirmed this finding by showing that attentional priority for color increased in response to manipulations of the orientation feature distribution, even though the color distribution remained the same. Although this effect did not reach significance in experiment 2a, we found the data across experiments 2a and 2b show substantially more evidence for this effect compared to the null (BF_{10}=6.3, Scheibehenne et al., 2016). This pattern of results is exactly what is expected if predicted variability was translated into attentional priority in a push-pull manner (Carandini and Heeger, 2012; Louie et al., 2013; Pinsk et al., 2004). These findings suggest that target-templates help reduce uncertainty by optimally weighting features based on the relative predicted utility for identifying the target, rather than removing or adding features in an “all-or-none” fashion.

Knowledge about the variability of feature distributions also moderated distractor interference. Distractor similarity to targets is known to decrease the efficiency of search due to the competition that arises when distractors and targets have similar features (Driver, 2001; Duncan and Humphreys, 1989; Wolfe and Horowitz, 2017). RT analyses in all three experiments revealed that the effect of target-to-distractor difference on search was greater for prioritized feature dimensions (i.e., the low-variance feature). This shows that distractors only affect the search process when they share features used to identify the target.

Finally, in Experiments 2a and 2b we tested two possibilities for how predicted variability is encoded by the target template. One possibility was that the predicted distribution would only be reflected in the precision of a single feature encoded into working memory (i.e. the cue) (Grubert and Eimer, 2013; Olivers et al., 2011). The other possibility was that predictions were encoded as a likelihood distribution over possible target features - separate from cue memory. While our data showed that memory was more precise for low-variance features compared to high variance features (Boettcher et al., 2020; Kerzel and Witzel, 2019; Niklaus et al., 2017), we also found that only prediction probes were sensitive to changes in the actual distribution of target orientation. This suggests that separate representations are held for memories of the cue and for predictions about what the target will look like, with the latter being used to set attentional priority.

Interestingly, predictions about the target color, which did not change in variability in Experiment 2, became more precise as predictions about target orientation became more variable. This could be due to the fact that predictions about the reliability of target features are coded relative to each other (Foley et al., 2017; Horan et al., 2019). This would result in color seeming “less variable” because all target orientations became more variable. However, further experiments will be needed to confirm this interpretation. Nonetheless, the changes in both color and orientation revealed that feature predictions showed the same pattern of changes as visual search performance, suggesting that predictions about the target distribution, not specific memories of the average target, are linked to attentional priority. These results highlight the fact that when target features are uncertain, observers rely on predictions about possible target features rather than a single memory trace.

One outstanding question about the representation of target feature variability concerns whether the knowledge of variability is specific to targets or represented across targets. For example, it is difficult to tell with these data if there is a unique representation of variability for each target, which appear similar because all targets follow similar statistics. Or in contrast, if feature variability is coded in a common representation across targets. Future studies may address this question by using multiple targets which share features but differ in variability to test for target specific knowledge of the variability.

Furthermore, while our data suggest that predictions influence feature-based attentional priority, it is difficult to tell whether this change in priority is specific to changes in guidance or target decisions. Previous work has shown that the initial guidance of attention is a distinct process from target-match decisions, and that these two processes may rely different information about target objects (Rajsic and Woodman, 2019; Wolfe, 2021; Yu et al., 2022). However, this study cannot determine how predictions influence attentional priority at each stage of visual search. Thus, we are agnostic to whether these results describe a change in the information being used for guidance or target-match decisions within a visual search cycle.

In conclusion, we have shown that the target template encodes detailed predictive knowledge that is vital to the search process and determines how attention is allocated in feature-based attention. Our work joins a growing literature that reframes the target template as an inherently prospective construct that seeks to maximize predictive information for the upcoming task and ignore information with little predictive value (Nobre and Stokes, 2019). These findings inform theories of attention by showing how attentional systems overcome the uncertainty of dynamic stimuli during goal directed processing, and by showing that predictive processes play a key role in setting priority for uncertain targets.

References

Andersen, S.K., Hillyard, S.A., and Müller, M.M. (2008). Attention Facilitates Multiple Stimulus Features in Parallel in Human Visual Cortex. Curr. Biol. *18*, 1006–1009.

Badre, D., Doll, B.B., Long, N.M., and Frank, M.J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. Neuron *73*, 595–607.

Bae, G.Y., Olkkonen, M., Allred, S.R., and Flombaum, J.I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. J. Exp. Psychol. Gen. *144*, 744–763.

Becker, S.I., Folk, C.L., and Remington, R.W. (2010). The Role of Relational Information in Contingent Capture. J. Exp. Psychol. Hum. Percept. Perform. *36*, 1460–1476.

Van Bergen, R.S., and Jehee, J.F.M. (2019). Probabilistic Representation in Human Visual Cortex Reflects Uncertainty in Serial Decisions. J. Neurosci. *39*, 8164–8176.

Bichot, N.P., Heard, M.T., DeGennaro, E.M., and Desimone, R. (2015). A Source for Feature-Based Attention in the Prefrontal Cortex. Neuron *88*, 832–844.

Boettcher, S.E.P., van Ede, F., and Nobre, A.C. (2020). Functional biases in attentional templates from associative memory. J. Vis. *20*, 7.

Boettcher, S.E.P., Gresch, D., Nobre, A.C., and van Ede, F. (2021). Output planning at the input stage in visual working memory. Sci. Adv. *7*, eabe8212.

Bravo, M.J., and Farid, H. (2012). Task demands determine the specificity of the search template. Attention, Perception, Psychophys. *74*, 124–131.

Bravo, M.J., and Farid, H. (2016). Observers change their target template based on expected context. Attention, Perception, Psychophys. *78*, 829–837.

Buschman, T.J. (2021). Balancing Flexibility and Interference in Working Memory. Annu. Rev. OfVision Sci. *7*, 1–22.

Buschman, T.J., and Miller, E.K. (2007). Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and. Science (80-.). *315*, 1860–1863.

Carandini, M., and Heeger, D.J. (2012). Normalization as a canonical neural computation. Nat. Rev. Neurosci. *13*, 51–62.

Chelazzi, L., Miller, E.K., Duncan, J., and Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. Nature *363*, 345–347.

Chelazzi, L., Duncan, J., Miller, E.K., and Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. J. Neurophysiol. *80*, 2918–2940.

Chetverikov, A., Campana, G., and Kristjánsson, Á. (2017a). Learning features in a complex and changing environment: A distribution-based framework for visual attention and vision in general. Prog. Brain Res. *236*, 97–120.

Chetverikov, A., Campana, G., and Kristjánsson, Á. (2017b). Representing Color Ensembles. Psychol. Sci. *28*, 1510–1517.

Chetverikov, A., Campana, G., and Kristjánsson, Á. (2020). Probabilistic rejection templates in visual working memory. Cognition *196*, 104075.

Chun, M.M., and Jiang, Y. (1998). Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention. Cogn. Psychol. *36*, 28–71.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. *36*, 181–204.

Desimone, R., and Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. Annu. Rev. Neurosci. *18*, 193–222.

Driver, J. (2001). A selective review of selective attention research from the past century. Br. J. Psychol. *92*, 53–78.

Duncan, J. (1989). Boundary condition on parallel in human vision. Perception *18*, 457–469.

Duncan, J., and Humphreys, G.W. (1989). Visual Search and Stimulus Similarity. Psychol. Rev. *96*, 433–458.

van Ede, F., Board, A.G., and Nobre, A.C. (2020). Goal-directed and stimulus-driven selection of internal representations. Proc. Natl. Acad. Sci. 202013432.

van Ede, F., Deden, J., and Nobre, A.C. (2021). Looking ahead in working memory to guide sequential behaviour. Curr. Biol. *31*, R779–R780.

Fetsch, C.R., Pouget, A., Deangelis, G.C., and Angelaki, D.E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. Nat. Neurosci. *15*, 146–154.

Foerster, R.M., and Schneider, W.X. (2018). Involuntary top-down control by search-irrelevant features: Visual working memory biases attention in an object-based manner. Cognition *172*, 37–45.

Foley, N.C., Kelly, S.P., Mhatre, H., Lopes, M., and Gottlieb, J. (2017). Parietal neurons encode expected gains in instrumental information. Proc. Natl. Acad. Sci. *114*, E3315–E3323.

Found, A., and Müller, H.J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a “dimension-weighting” account. Percept. Psychophys. *58*, 88–101.

Friston, K. (2010). The free-energy principle: A unified brain theory? Nat. Rev. Neurosci. *11*, 127–138.

Geng, J.J., and Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. Percept. Psychophys. *67*, 1252–1268.

Geng, J.J., and Witkowski, P. (2019). Template-to-distractor distinctiveness regulates visual search efficiency. Curr. Opin. Psychol. *29*, 119–125.

Geng, J.J., Di Quattro, N.E., and Helm, J. (2017). Distractor probability changes the shape of the attentional template. J. Exp. Psychol. Hum. Percept. Perform. *43,* 1993–2007.

Geng, J.J., Won, B.Y., and Carlisle, N.B. (2019). Distractor Ignoring: Strategies, Learning, and Passive Filtering. Curr. Dir. Psychol. Sci. *28,* 600–606.

Gershman, S.J. (2019). Uncertainty and exploration. Decision *6*, 277–286.

Giesbrecht, B., Woldorff, M.G., Song, A.W., and Mangun, G.R. (2003). Neural mechanisms of top-down control during spatial and feature attention. Neuroimage *19*, 496–512.

Gottlieb, J. (2018). Understanding active sampling strategies: Empirical approaches and implications for attention and decision research. Cortex *102*, 150–160.

Gottlieb, J., and Balan, P. (2010). Attention as a decision in information space. Trends Cogn. Sci. *14,* 240–248.

Gottlieb, J., Oudeyer, P.Y., Lopes, M., and Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. Trends Cogn. Sci. *17,* 585–593.

Gramann, K., Toellner, T., Krummenacher, J., Eimer, M., and Müller, H.J. (2007). Brain electrical correlates of dimensional weighting: An ERP study. Psychophysiology *44*, 277–292.

Grubert, A., and Eimer, M. (2013). Qualitative differences in the guidance of attention during single-color and multiple-color visual search: Behavioral and electrophysiological evidence. J. Exp. Psychol. Hum. Percept. Perform. *39,* 1433–1442.

Hénaff, O.J., Boundy-singer, Z.M., Meding, K., and Ziemba, C.M. (2020). Representation of visual uncertainty through neural gain variability. Nat. Commun. 1–12.

Henderson, J.M., and Hayes, T.R. (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. Nat. Hum. Behav. *1,* 743–747.

Hillis, J.M., Watt, S.J., Landy, M.S., and Banks, M.S. (2004). Slant from texture and disparity cues: Optimal cue combination. J. Vis. *4,* 967–992.

Hollingworth, A., and Hwang, S. (2013). The relationship between visual working memory and attention: Retention of precise colour information in the absence of effects on perceptual selection. Philos. Trans. R. Soc. B Biol. Sci. *368.*

Horan, M., Daddaoua, N., and Gottlieb, J. (2019). Parietal neurons encode information sampling based on decision uncertainty. Nat. Neurosci. *22,* 1327–1335.

Hout, M.C., and Goldinger, S.D. (2015). Target templates: the precision of mental representations affects attentional guidance and decision-making in visual search. Attention, Perception, Psychophys. *77*, 128–149.

Huang, C., Vilotijević, A., Theeuwes, J., and Donk, M. (2021). Proactive distractor suppression elicited by statistical regularities in visual search. Psychon. Bull. Rev. *28,* 918–927.

Iglesias, S., Mathys, C., Brodersen, K.H., Kasper, L., Piccirelli, M., denOuden, H.E.M., and Stephan, K.E. (2013). Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning. Neuron *80,* 519–530.

Kastner, S., and Ungerleider, L.G. (2000). Mechanisms of Visual Attention in the Human Cortex. Annu. Rev. Neurosci. *23*, 315–341.

Kerzel, D., and Huynh Cong, S. (2021). Attentional templates are sharpened through differential signal enhancement, not differential allocation of attention. J. Cogn. Neurosci. *33*, 594–610.

Kerzel, D., and Witzel, C. (2019). The allocation of resources in visual working memory and multiple attentional templates. J. Exp. Psychol. Hum. Percept. Perform. *45*, 645–658.

Knill, D.C., and Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. Trends Neurosci. *27*, 712–719.

Kok, P., Jehee, J.F.M., and de Lange, F.P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. Neuron *75*, 265–270.

Kozyrev, V., Daliri, M.R., Schwedhelm, P., and Treue, S. (2019). Strategic deployment of feature-based attentional gain in primate visual cortex. PLoS Biol. *17*, 1–28.

Lee, J., and Geng, J.J. (2020). Flexible weighting of target features based on distractor context. Attention, Perception, Psychophys. *82*, 739–751.

Li, H.-H., Sprague, T.C., Yoo, A.H., Ma, W.J., and Curtis, C.E. (2021). Joint representation of working memory and uncertainty in human cortex. Neuron 2021.04.05.438511.

Liu, T., Larsson, J., and Carrasco, M. (2007). Feature-Based Attention Modulates Orientation-Selective Responses in Human Visual Cortex. Neuron *55*, 313–323.

Lo, S., and Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. Front. Psychol. *6*, 1171–1187.

Louie, K., Khaw, M.W., and Glimcher, P.W. (2013). Normalization is a general neural mechanism for context-dependent decision making. Proc. Natl. Acad. Sci. U. S. A. *110*, 6139–6144.

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

Luke, S.G. (2017). Evaluating significance in linear mixed-effects models in R. Behav. Res. Methods *49*, 1494–1502.

Ma, W.J., and Jazayeri, M. (2014). Neural Coding of Uncertainty and Probability. Annu. Rev. Neurosci. *37*, 205–220.

Malcolm, G.L., and Henderson, J.M. (2009). The effects of target template specificity on visual search in real-world scenes : Evidence from eye movements. J. Vis. *9*.

Malcolm, G.L., and Henderson, J.M. (2010). Combining top-down processes to guide eye movements during real-world scene search. J. Vis. *10*.

Mehrpour, V., Martinez-Trujillo, J.C., and Treue, S. (2020). Attention amplifies neural representations of changes in sensory input at the expense of perceptual accuracy. Nat. Commun. *11*.

Meyyappan, S., Rajan, A., Mangun, G.R., and Ding, M. (2021). Role of inferior frontal junction (ifj) in the control of feature versus spatial attention. J. Neurosci. *41*, 8065–8074.

Michael, E., De Gardelle, V., Nevado-Holgado, A., and Summerfield, C. (2015). Unreliable evidence: 2 sources of uncertainty during perceptual choice. Cereb. Cortex *25*, 937–947.

Moore, T., and Fallah, M. (2001). Control of eye movements and spatial attention. Proc. Natl. Acad. Sci. *98*, 1273–1276.

Müller, H.J., and Krummenacher, J. (2006). Locus of dimension weighting: Preattentive or postselective? Vis. Cogn. *14*, 490–513.

Müller, H.J., Heller, D., and Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. Percept. Psychophys. *57*, 1–17.

Müller, H.J., Reimann, B., and Krummenacher, J. (2003). Visual Search for Singleton Feature Targets Across Dimensions: Stimulus- and Expectancy-Driven Effects in Dimensional Weighting. J. Exp. Psychol. Hum. Percept. Perform. *29*, 1021–1035.

Nah, J.C., and Geng, J.J. (2022). Thematic object pairs produce stronger and faster grouping than taxonomic pairs. J. Exp. Psychol. Hum. Percept. Perform. *48*, 1325–1335.

Nako, R., Wu, R., Smith, T.J., and Eimer, M. (2014). Item and Category-Based Attentional Control During Search for Real-World Objects : Can You Find the Pants Among the Pans ? J. Exp. Psychol. Hum. Percept. Preformance *40*, 1283–1288.

Navalpakkam, V., and Itti, L. (2007). Search Goal Tunes Visual Features Optimally. Neuron *53*, 605–617.

Niklaus, M., Nobre, A.C., and Van Ede, F. (2017). Feature-based attentional weighting and spreading in visual working memory. Sci. Rep. *7*.

Niv, Y., Daniel, R., Geana, A., Gershman, S.J., Leong, Y.C., Radulescu, A., and Wilson, R.C. (2015). Reinforcement Learning in Multidimensional Environments Relies on Attention Mechanisms. J. Neurosci. *35*, 8145–8157.

Noah, S., Meyyappan, S., Ding, M., and Mangun, G.R. (2022). Anticipatory attention is a stable state induced by transient control mechanisms. Front. Hum. Neurosci. *16*, 1–12.

Nobre, A.C., and Stokes, M.G. (2019). Premembering Experience: A Hierarchy of Time-Scales for Proactive Attention. Neuron *104*, 132–146.

O'Reilly, J.X., Schüffelgen, U., Cuell, S.F., Behrens, T.E.J., Mars, R.B., and Rushworth, M.F.S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. Proc. Natl. Acad. Sci. U. S. A. *110*.

Olivers, C.N.L., Meijer, F., and Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. J. Exp. Psychol. Hum. Percept. Perform. *32*, 1243–1265.

Olivers, C.N.L., Peters, J., Houtkamp, R., and Roelfsema, P.R. (2011). Different states in visual working memory: When it guides attention and when it does not. Trends Cogn. Sci. *15*, 327–334.

Panichello, M.F., and Buschman, T.J. (2021). Shared mechanisms underlie the control of working memory and attention. Nature *4*.

Peirce, J.W. (2007). PsychoPy-Psychophysics software in Python. J. Neurosci. Methods *162*, 8–13.

Pinsk, M.A., Doniger, G.M., and Kastner, S. (2004). Push-pull mechanism of selective attention in human extrastriate cortex. J. Neurophysiol. *92*, 622–629.

Pollmann, S., Weidner, R., Müller, H.J., and von Cramon, D.Y. (2006). Neural correlates of visual dimension weighting. Vis. Cogn. *14*, 877–897.

Press, C., Kok, P., and Yon, D. (2019). The Perceptual Prediction Paradox. Trends Cogn. Sci. *24*, 12–24.

Press, C., Kok, P., and Yon, D. (2020). Learning to Perceive and Perceiving to Learn. Trends Cogn. Sci. *24*, 260–261.

Rajsic, J., and Woodman, G.F. (2019). Do we remember templates better so that we can reject distractors better ? Attention, Perception, Psychophys.

Rajsic, J., Ouslis, N.E., Wilson, D.E., and Pratt, J. (2017). Looking sharp: Becoming a search template boosts precision and stability in visual working memory. Attention, Perception, Psychophys. *79*, 1643–1651.

Rajsic, J., Carlisle, N.B., and Woodman, G.F. (2020). What not to look for: Electrophysiological evidence that searchers prefer positive templates. Neuropsychologia *140*, 107376.

Rao, R.P., and Ballard, D.H. (1999). Hierarchical Predictive Coding Model Hierarchical Predictive Coding of Natural Images. Nat. Neurosci. *2*, 79–87.

Reeder, R.R., and Peelen, M. (2013). The contents of the search template for category-level search in natural scenes. J. Vis. *13*, 1–13.

Reeder, R.R., Hanke, M., and Pollmann, S. (2017). Task relevance modulates the cortical representation of feature conjunctions in the target template. Sci. Rep. *7*.

Ruff, C.C., Blankenburg, F., Bjoertomt, O., Bestmann, S., Freeman, E., Haynes, J.D., Rees, G., Josephs, O., Deichmann, R., and Driver, J. (2006). Concurrent TMS-fMRI and Psychophysics Reveal Frontal Influences on Human Retinotopic Visual Cortex. Curr. Biol. *16*, 1479–1488.

Scheibehenne, B., Jamil, T., and Wagenmakers, E.-J. (2016). Bayesian Evidence Synthesis Can Reconcile Seemingly Inconsistent Results. Psychol. Sci. *27*, 1043–1046.

Schmidt, J., and Zelinsky, G.J. (2009). Search guidance is proportional to the categorical specificity of a target cue. Q. J. Exp. Psychol. *62*, 1904–1914.

Serences, J.T., Saproo, S., Scolari, M., Ho, T., and Muftuler, L.T. (2009). Estimating the influence of attention on population codes in human visual cortex using voxel-based tuning functions. Neuroimage *44*, 223–231.

Summerfield, C., and Egner, T. (2009). Expectation (and attention) in visual cognition. Trends Cogn. Sci. *13*, 403–409.

Summerfield, C., Behrens, T.E., and Koechlin, E. (2011). Perceptual classification in a rapidly-changing environment. Neuron *71*, 725–736.

Töllner, T., Zehetleitner, M., Gramann, K., and Müller, H.J. (2010). Top-down weighting of visual dimensions: Behavioral and electrophysiological evidence. Vision Res. *50*, 1372–1381.

Tomov, M.S., Truong, V.Q., Hundia, R.A., and Gershman, S.J. (2020). Dissociable neural correlates of uncertainty underlie different exploration strategies. Nat. Commun. *11*.

Trapp, S., Parr, T., Friston, K., and Schröger, E. (2021). The Predictive Brain Must Have a Limitation in Short-Term Memory Capacity. Curr. Dir. Psychol. Sci.

Treue, S., and Trujillo, J.C.M. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. Nature *399*, 575–579.

Vickery, T.J., King, L.W., and Jiang, Y. (2005). Setting up the target template in visual search. J. Vis. *5*, 81–92.

Vogel, E.K., Woodman, G.F., and Luck, S.J. (2001). Storage of features, conjunctions, and objects in visual working memory. J. Exp. Psychol. Hum. Percept. Perform. *27*, 92–114.

Vossel, S., Thiel, C.M., and 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.

Wei, P., Yu, H., Müller, H.J., Pollmann, S., and Zhou, X. (2018). Differential brain mechanisms for processing distracting information in task-relevant and -irrelevant dimensions in visual search. Hum. Brain Mapp. *40*, 110–124.

Witkowski, P., and Geng, J.J. (2019). Learned feature variance is encoded in the target template and drives visual search. Vis. Cogn. 1–15.

Witkowski, P.P., Park, S.A., and Boorman, E.D. (2022). Neural mechanisms of credit assignment for inferred relationships in a structured world. Neuron 2021.12.22.473879.

Wolfe, J.M. (2020). Visual Search: How Do We Find What We Are Looking For? Annu. Rev. Vis. Sci. *6*, 1–24.

Wolfe, J.M. (2021). Guided Search 6.0: An updated model of visual search (Psychonomic Bulletin & Review).

Wolfe, J.M., and Horowitz, T.S. (2017). Five factors that guide attention in visual search. Nat. Hum. Behav. *1*, 1–8.

Wolfe, J.M., Cave, K.R., and Franzel, S.L. (1989). Guided Search: An Alternative to the Feature Integration Model for Visual Search. J. Exp. Psychol. Hum. Percept. Perform. *15*, 419–433.

Wolfe, J.M., Friedman-Hill, S.R., Stewart, M.I., and O'Connell, K.M. (1992). The Role of Categorization in Visual Search for Orientation. J. Exp. Psychol. Hum. Percept. Perform. *18*, 34–49.

Won, B.-Y., Venkatesh, A., Witkowski, P.P., Banh, T., and Geng, J.J. (2021). Memory precision for salient distractors decreases with learned suppression. Psychon. Bull. Rev.

Woodman, G.F., and Vogel, E.K. (2008). Selective storage and maintenance of an object's features in visual working memory. Psychon. Bull. Rev. *15*, 223–229.

Woodman, G.F., Carlisle, N.B., and Reinhart, R.M.G. (2013). Where do we store the memory representations that guide attention? J. Vis. *13*, 1–17.

Yang, H., and Zelinsky, G.J. (2009). Visual search is guided to categorically-defined targets. Vision Res. *49*, 2095–2103.

Yu, X., and Geng, J.J. (2019). The Attentional Template is Shifted and Asymmetrically Sharpened by Distractor Context. J. Exp. Psychol. Hum. Percept. Perform. *45*, 1–37.

Yu, X., Hanks, T.D., and Geng, J.J. (2022). Attentional Guidance and Match Decisions Rely on Different Template Information During Visual Search. Psychol. Sci. *33*, 105–120.

Yu, X., Zhou, Z., Becker, S.I., Boettcher, S.E.P., and Geng, J.J. (2023). Good-enough attentional guidance. Trends Cogn. Sci. *xx*, 1–13.

Zanto, T.P., Rubens, M.T., Bollinger, J., and Gazzaley, A. (2010). Top-down modulation of visual feature processing: The role of the inferior frontal junction. Neuroimage *53*, 736–745.

Zhang, W., and Luck, S.J. (2009). Feature-based attention modulates feedforward visual processing. Nat. Neurosci. *12*, 24–25.

Zhang, X., Mlynaryk, N., Ahmed, S., Japee, S., and Ungerleider, L.G. (2018). The role of inferior frontal junction in controlling the spatially global effect of feature-based attention in human visual areas. PLoS Biol. *16*, 1–28.

Zhou, Z., and Geng, J.J. (2023). Learned associations serve as target proxies during difficult but not easy 1 visual search. PsyArXiv.

Zivony, A., and Lamy, D. (2016). The role of dimension relevance in features’ access to response-selection mechanisms. J. Exp. Psychol. Hum. Percept. Perform. *42*, 1873–1885.

Chapter 4: Representations of feature identity and uncertainty in prefrontal cortex

Introduction

Uncertainty is the only guarantee in life. Our sensory environment is rich with visual variability, creating uncertainty in the way even categorically similar objects may appear. For example, imagine going to look for an apple in a friend's kitchen. With no *a priori* knowledge of its *exact* features, you will have to account for some perceptual uncertainty. Will the apple be red, green, yellow, or somewhere in between? Will its shape be perfectly matched to a canonical apple? If not, you will need to account for these possible differences when deciding what objects in the visual field are a possible "apples", or else risk dramatic delays in search times or failure to locate the target (Bravo & Farid, 2016; Geng & Witkowski, 2019). Situations like this abound in the real world, everywhere from developmental transitions in appearance over time, to friends who must be recognized even though they switch outfits from day to day. Despite this ubiquity, previous studies of visual search have generally focused on understanding the neural mechanisms underlying specific target features, and very little on how the brain integrates uncertainty into predictions about how the target *might look* when we see it. The current experiments seek to explain the neural mechanisms the underlie how uncertainty is combined with predictive knowledge of search targets to guide attention during visual search.

Theories of attention use the concept of the target template to describe the representation of visual search targets that is used to set sensory gain and make target match decisions (Duncan & Humphreys, 1989; Noah et al., 2022; Wolfe, 2021; Woodman et al., 2013). Target templates have previously been characterized as storing a single target's remembered identity (Olivers et al., 2011; Vickery et al., 2005), and many studies have found that more precise memories produce faster and more accurate search (Hout & Goldinger, 2015; Malcolm & Henderson, 2009; Schmidt & Zelinsky, 2009; Yang & Zelinsky, 2009). For example, Hout and Goldinger (2015) gave participants a single cue that was either identical to the target, a rotated version of the target, or an exemplar from the same category. The authors found that search was most efficient when the cue was an exact match for the target. However, many features of real-world targets are variable, creating uncertainty in the exact properties the target prior to search. This means that targets may be missed by highly precise templates because the actual features of the target will often be a poor match to the target template.

One way in which attentional systems might overcome this limitation is by integrating information about the uncertainty of features into the target template. Theories of perception hold such probabilistic representations as critical to cognition, linking perception to internal “beliefs” about what stimuli we will encounter in the world around us (Clark, 2013; Kok et al., 2012; Press et al., 2019, 2020). Previous work has revealed that statistical information about the distribution of target features is encoded by target templates and helps to determine which features are prioritized during visual search (Hansmann-Roth et al., 2022; Witkowski & Geng, 2019, 2022). For example, in Witkowski and Geng (2022) the authors systematically manipulated the variability of two target features in a visual search task. Participants in this task not only tracked the variability of each feature over time but used this knowledge to generate target predictions set attentional priority prior to search. While these studies revealed that the uncertainty of target features is intimately related to the construction of the target template, the neural mechanisms which underlie the integration of uncertainty into target template representations are largely unknown.

Regions of the lateral prefrontal cortex, including DLPFC and IFJ have long been thought to critical for supporting target template representations. Prior work has identified these regions as a source for maintaining task-relevant information in protected states (Buschman & Miller, 2007; Ester et al., 2015; Miller et al., 1996) and providing top-down modulation of activity in regions coding sensory information such as visual cortex (D'Esposito et al., 2000; Giesbrecht et al., 2003; Jackson et al., 2021; Meyyappan et al., 2021). In one compelling study, Bichot and colleagues trained rhesus monkeys to perform a cued visual search task, and recorded neural activity in the ventral pre-arcuate region, thought to be a nonhuman primate homologue of IFJ (Bedini & Baldauf, 2021). The authors found that neurons in this region were selective to cued search targets and maintained this selectivity throughout the visual search process (Bichot et al., 2015). Indeed, the IFJ has been characterized as probable a probable source for feature-based attention about target objects in humans (Baldauf & Desimone, 2014; Meyyappan et al., 2021; Zanto et al., 2010; Zhang et al., 2018), alongside representations of task-relevant mnemonic information in DLPFC (Ester et al., 2015). Together, these studies support the idea that DLPFC and IFJ encode information critical to the construction of feature-based attention templates but say little about the uncertainty of target features may be coded in these regions.

One possibility is that representations of feature uncertainty in prefrontal cortex rely on similar mechanisms as representations of decision uncertainty. When making decisions between various choices, univariate codes in prefrontal cortex track the total uncertainty of possible outcomes in a task, increasing when outcome- or categorical uncertainty increases (Badre et al., 2012; S. Li & Yang, 2012; Michael et al., 2015; Summerfield et al., 2011; Tomov et al., 2020). For example, Summerfield et al. (2011) found that when participants were asked to make judgments about the learned category of orientation stimuli, but the category was ambiguous, DLPFC and IFJ showed higher activation profiles compared to when it was unambiguous. Studies from Hansen and colleagues (2012) compliment this work by showing that DLPFC plays a role in signaling prior knowledge about the distribution of possible outcomes when current decision is uncertain. When a stimulus is viewed, these regions also show error responses proportionate to the likelihood of the observed features (Iglesias et al., 2013). Together, these studies suggest that activity DLPFC and IFJ contain statistical knowledge about the uncertainty of visual stimuli and may be a plausible source generating predictions about search target which integrate both the identity and uncertainty of target features.

In the current paper, we test the idea that DLPFC and IFJ represents target features in conjunction with statistical knowledge about the uncertainty of these features to support feature-based attention. We do so by employing a visual search task in which target features follow distributions with differing variability, requiring participants to know both the predicted target and its associated uncertainty. We then test whether these regions encode both the identity and uncertainty of the predicted target and whether this uncertainty moderate's activity in these regions during search. We find that IFJ, DLPFC and visual cortex form an interconnected network of regions that support coding of the target identity, but critically, we find that the prefrontal regions of this network code the uncertainty of target feature prediction prior to search. These results help to illuminate neural mechanisms that underlie attentional processing by showing how uncertainty is integrated with information about the target in advance of visual search.

Methods

Participants

We recruited a sample of 20 participants (self-reported females=13, males=7; mean age =20.01) from the University of California, Davis SONA system. This sample size was calculated to have 90% power to detect the desired RT effect based on simulations of from pilot data, and matched the typical sample size of similar studies (Summerfield et al., 2011). All participants had normal or corrected-to-normal vision and no impairment in color vision. Participants received a combination of payment and course credit for participation in the experiment. All procedures were approved by the University of California, Davis Internal Review Board (IRB). One block from one subject was removed due to excessive motion during scanning (head movement > 3mm) but removing this block did not qualitatively change any results.

Behavioral Task

Each trial began with an auditory tone that cued the participant to the target color. A high tone (750hz) indicated that the target color would come from a range of colors centered around an orange-pink value (RGB: 0.92,0.58,0.53); a low tone (375hz) indicated the target would come from a range of colors centered at a blue-green value (RGB: 0.18,0.74,0.82). The central color of in each range was always the most likely to be target color, and we call these colors the "predicted target". Tones were presented bilaterally for 250ms and followed by an 8000ms interstimulus-interval (ISI) of silence. The ISI was followed by a visual search display during which participants rapidly reported the location of the object that most closely matched the

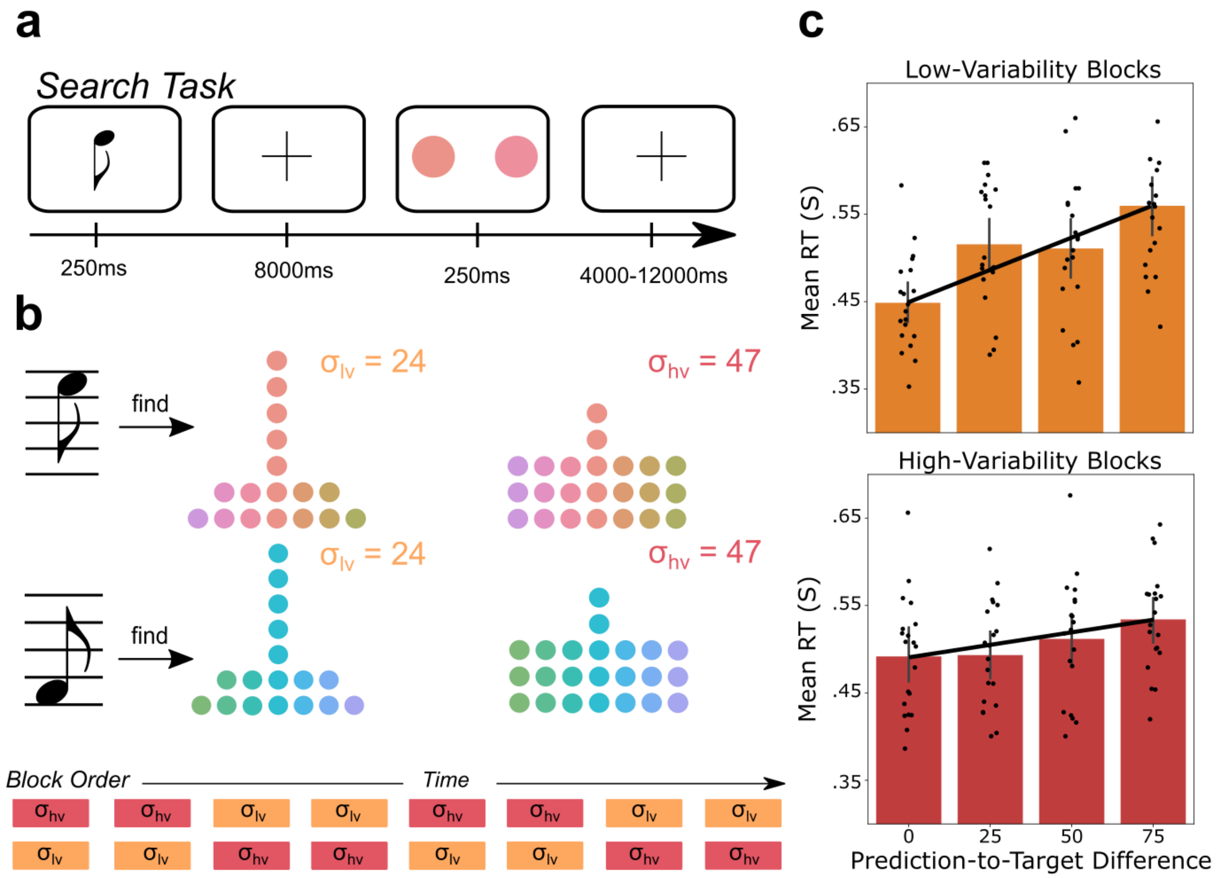

Figure 4.1 Search task schematic and behavioral results

a) Each trial began with the presentation of a cue-tone which informed participants the relevant to-be-matched color. High-tones (750hz) indicated a pink-orange central color (RGB 0.92,0.58,0.53) while low-tones (350hz) indicated a blue-green color (RGB 0.18,0.74,0.82). After the interstimulus interval (8000ms) two colors, one target and one distractor, were displayed on the screen for 250ms but responses for the target location (left/right) were allowed for up to one second. b) blocks were divided into two conditions based on the variability of the target distribution. In low-variability condition, the target distributions had a low standard deviation (24 degrees) centered around the to-be-matched color. In high-variability blocks, the target distribution had a high standard deviation (47 degrees). c) Response time during the search period followed the expected distribution of target stimuli. Response times to exact matching stimuli were faster for exactly matching stimuli in low-variability blocks. However, response times increased as a function of prediction-to-target difference. High-variability blocks showed a significantly shallower slope.

predicted target color. We call this the “target”. Participants reported the location of the target by pressing “1” if it was on the left and “2” if it was on the right. The search display was visible for 250ms, after which the stimuli disappeared, but participants were allowed 1000ms to respond. Each trial ended with feedback about whether the participant correctly identified the target in time (green circle or a red “X” otherwise). This was followed by a randomly jittered 4000-12000ms intertrial-interval, which followed a Gamma distribution with a mean of 6600ms. Each participant completed 160 trials in total during the experiment.

The target color on each trial was drawn from a predetermined distribution of 7 possible colors, ranging between -75- to 75 degrees from the mean target color in 25-degree steps. Participants were told beforehand that the mean of each target distribution was most likely target indicated by the cue, and that the distribution was symmetric. The variance of the target distributions differed across two conditions. In the “low-variability” condition, the target was the mean color value on 80% of trials, -/+25 degrees away from the mean target color on 5% of trials, and -/+50 degrees away on 10% of trials. Finally, in 5% of trials the target was -/+75 degrees away from the mean target color. We upsampled targets that were 50 degrees away to increase the number of observations in the tails of the distribution while still maintaining a Gaussian-like profile. In the “high-variability” condition only 25% of target colors were from the mean of the distribution, and each other possible target color occurred on 12.5% of trials. The manipulation created more certainty about the upcoming target’s color in the low-variability condition compared to the high-variability condition.

Finally, distractors in each trial were randomly selected from a uniform distribution between -179 degrees and 180 degrees from the target distribution's mean color value. The only constraint on selecting the distractor color for a single trial was that it must be at least 20 degrees more distant from the mean value of the target distribution than the actual target. All stimuli had colors taken from a CIE color space color wheel (Bae, Olkkonen, Allred, & Flombaum, 2015). Both the target and distractor had a radius of 2 degrees of visual angle and were displayed at 7.5 degrees of eccentricity.

Statistical Analysis of Behavior

To estimate the effect of variability on response times we used a linear regression model to test how *variability condition* moderated participant RT for targets over-and-above the trial's *prediction-to-target difference* alone. *Prediction-to-target difference* was calculated as the circular difference between the predicted target (mean of the target distribution) and the target color. We hypothesized that the effect of *prediction-to-target difference* would be modulated by expectations about the distribution of possible colors the target could take. For example, the RT cost is expected to be larger for the target color that is 50-degrees from the mean of the target distribution in the low-variability condition, compared to the high-variability condition. This is because participants will have low uncertainty about the color of the target and therefore expect the target to exactly match; in contrast, RT for the same target color should be faster in the high-variability condition because participants are more uncertain about the target color and may expect any of the possible target colors to appear. This outcome would lead to an interaction in RT between the *prediction-to-target difference* and *variability condition*. Such an interaction would suggest that attentional effects based on relationship between the predicted target and realized target are modulated by expectations about the variance of the target.

We tested this hypothesis by fitting RT data from trials with correct responses to a gamma-distributed hierarchical regression model (Lo & Andrews, 2015), using the lme4 package in R. Trials with incorrect responses or RTs greater than 3 times the interquartile range (IQR) from the median RT were excluded from the analysis (Mean=13 trials total per participant).

The regression model had one regressor for *prediction-to-target difference*, which reflects the slope of the RT increase as the target becomes more dissimilar from the predicted target color. We also included a binary regressor for *variability condition* capturing the overall difference in RT due to uncertainty about the target color, and an interaction term between this regressor and *prediction-to-target difference.* One final regressor was included to capture *target-to-distractor difference*, calculated by the circular difference between target color and the distractor color, which can also influence RT (Duncan & Humphreys, 1989; Wolfe & Horowitz, 2017). The random effect structure included random intercepts for each participant and random by-subject slopes for each of the fixed effects.

All significance testing was done using likelihood-ratio tests between the full model and models with the relevant fixed effect removed, as is appropriate for hierarchical models (Luke, 2017). For example, the significance of the interaction between *prediction-to-target difference* and *variability condition* was tested by fitting a model with and without the fixed-effect of the interaction.

MRI data acquisition

Data was acquired using a Siemens Skyra 3 Tesla scanner. We used gradient-echo-planar imaging (EPI) pulse sequence, with a multi- band acceleration factor of 2, and aligned the slice angle to the anterior-posterior commissure line. We acquired 48 axial slices, 3mm thick with the following parameters: repetition time (TR) = 1500ms, echo time (TE) = 24.6ms, flip angle =75 degrees, field of view (FoV) = 210mm, voxel size = 3x3x3 mm. Slices were acquired in interleaved order. We also acquired a field map to correct for potential deformations with dual echo-time images covering the whole brain, with the following parameters: TR = 500ms, TE1 = 4.92ms, TE2 = 7.38ms, flip angle = 40 degrees, FoV = 192mm, voxel size = 3x3x3 mm. For accurate registration of the EPIs to the standard space, we acquired a T1-weighted structural image using a magnetization-prepared rapid gradient echo sequence (MPRAGE) with the following parameters: TR = 2400ms, TE = 2.98ms, flip angle = 7 degrees, FoV = 256mm, voxel size = 1x1x1 mm.

Pre-Processing

Preprocessing of the data was done in SPM12 (Wellcome Trust Centre for Human Neuroimaging) in MATLAB (2019a Matworks). Data were preprocessed using the default options in SPM. Images were slice-time corrected and realigned to the first volume of each sequence. Realignment was done to correct for motion using a six-parameter rigid body transformation. Inhomogeneities in the field were corrected using the phase of non-EPI gradient echo images at 2 echo times, which were co-registered with structural maps. Images were then spatially normalized by warping subject specific images to the reference brain in the MNI (Montreal Neurological Institute) coordinate system with 2mm isotropic voxels. Finally, for the univariate analysis images were spatially smoothed using a Gaussian kernel with a full width at half maximum of 8mm.

ROI Selection

Based on *a priori* prediction about the role of IFJ, DLPFC and Visual cortex in coding of the attentional template, we defined 3 regions of interest to be used in all neural analyses. The Inferior Frontal Junction ROI was created using IFJa and IFJp (Index 79 and 80, respectively) from the Human Connectome Project dataset (Glasser et al., 2016). The Dorsolateral prefrontal cortex ROI was taken from a probabilistic map of dorsal prefrontal cortex (Sallet et al., 2013). We constructed a visual cortex ROI by using all visual cortical areas from a probabilistic map of retinotopic regions in the brain (Wang et al., 2015). We included all voxels generated from the probabilistic maps without thresholding.

Univariate fMRI Analyses

To model BOLD activity in each voxel we used a GLM with three different regressors; the “cue period” (a boxcar, stimulus presentation time of 250ms plus 750ms after the tone ended), the delay period (a boxcar over seven seconds of the delay) and the search period (a 1 second boxcar beginning at the onset of the search stimuli). We also included two parametric modulators for the search period: The first was the *prediction-to-target difference* and the second was the *target-distractor similarity*. Six motion regressors were included as regressors of no interests in the model to account for translation and rotation in head position during the experiment.

From the first-level analysis, we calculated contrast images of the parameter estimates from all regressors (except the motion regressors) and parametric modulators. We then submitted the contrast for each participant into one sample t-tests in the group-level analyses.

Multivariate fMRI Analyses

The multivariate pattern analysis aimed to identify regions of the brain that coded knowledge of the cued target values in each trial. To test this, we estimated the BOLD activity patterns during the delay period using unsmoothed preprocessed images. The delay period was modeled as a boxcar that had a constant duration lasting 7 seconds from the end of the cue period. No parametric modulators were added. Pattern-based classifiers (linear support-vector machines [SVM]) were trained using LIBSVM (Chang & Lin, 2011) to distinguish between trials where the predicted target color was pink-orange and trials where the predicted target color was blue-green.

We used a searchlight procedure to identify voxels in each ROI that contained representations of the target identity. Each searchlight consisted of a 5x5x5 voxel cube placed around a centroid voxel. Each centroid was required to values in at least 10 of the surrounding voxels to be considered for further processing and were then standardized by z-scoring the beta values within each searchlight. We tested using an iterative, leave-one-out procedure. Trials from all but one block were used to train the SVM, then tested on the left-out block. Accuracy was averaged across iterations. The resulting maps were then spatially smoothed using a gaussian kernel with full width half maximum of 8mm. Group-level analyses were preformed using a one-sample t-test on accuracy maps across subjects.

To calculate trial-by-trial decoding strength of representations, we calculated the distance of each trial to the hyperplane separating the two categories using the equation specified on the LIBSVM webpage (https://www.csie.ntu.edu.tw/?cjlin/libsvm/faq.html). Patterns that are more distant from the hyperplane can be thought of as having more information about a category, and those that are closer to the hyperplane as having less information (Schuck and Niv, 2019). We then signed the distance of each point according to whether the predicted category label was correct (+ for correct, – for incorrect).

Group Level Statistical Inference

Group level testing was done using a one-sample t-test (df=19) on the cumulative functional maps generated by the first level analysis. All first level maps were smoothed prior to being combined and tested at the group level. To correct for multiple comparisons, we used Threshold-Free Cluster Enhancement (TFCE) which uses permutation testing and accounts for both the height and extent of the cluster (Smith and Nichols, 2009). All parameters were set to the default (H=2, E=0.5) and we used 5000 permutations. In all ROI based analyses and whole brain analyses we report effects below a pTFCE < .05 threshold. We first performed group-level inference on independent anatomical ROIs, then performed exploratory whole brain analyses. For ROI analyses, we first extracted voxels from each ROI in each subject's first-level activation map, averaged the maps together, then applied small volume TFCE correction. All other analyses were corrected for multiple comparisons at the whole brain level.

Results

Search behavior is determined by the Predicted Variability of Search Targets

We first assessed the extent to which the variability of each target distribution informed participants search behavior. Specifically, we hypothesized that RT would be slower as the *prediction-to-target difference* increased, but that the slope of this effect would be less steep in the high-variability condition compared to the low-variability condition. This hypothesis was confirmed by the significant interaction between *prediction-to-target difference* and *variability condition* (β =.79ms/deg/condition, model comparison $\chi 2(1)=12.61$, $p<.001$) (Fig 4.2). This means that the slope of the increase in RT per degree difference between the target and the mean of the distribution was higher in the low-variability condition compared to the high-variability condition. While there was a significant main effect of *target-to-distractor difference* on RT (β =.006ms/deg difference, χ2(1)= 11.59, p<.001), there was no interaction with *variability condition* (χ2(1)= 0.001, p=.98). Together these results show that participants are sensitive to the variability of the underlying target distribution and use this knowledge to moderate expectations about the features of an upcoming target object.

We next tested if the accuracy of the responses followed the variability of target distributions during search. Mean accuracy was high across all participants, and we observed only a marginal difference between accuracy in the low-variability (Mean=.929, SD=.052) and high-variability conditions (Mean=.900, SD=.064; β =-.3, χ2(1)=3.56, p=0.059). Importantly, we found no overall effect of *prediction-to-target difference* (χ2(1)=0.329, p=0.57) nor an interaction between *prediction-to-target difference* and condition (χ2(1)=0.47, p=0.49). We did observe a strong main effect of *target-to-distractor difference* (χ2(1)= 13.86, p<.001), that did not depend on *variability condition* (χ2(1)=2.50, p=0.11). These data show that even though the variability of the target affected RT performance, participants were able to correctly identify targets well across the range of values.

Target identity is coded in a prefrontal and visual network during delay period

Our first analysis sought to identify the neural representations of predicted target identity prior to visual search. To test for these representations, we performed a multivariate decoding analysis in which we trained linear pattern classifier distinguish trials in which the target would come from the pink-orange distribution or blue-green distribution using BOLD activity during the delay period. We began by conducting a search-light procedure over each ROI to estimate decoding accuracy at each voxel, independently for the low-variability and high-variability conditions.

This analysis identified a significant cluster of voxels in IFJ ([x,y,z] =[-50, 12,30], t[19] = 3.48, pTFCE < 0.05 ROI corrected; Figure 4.2b), and in a nearby region of left DLPFC ([x,y,z] =[-32, 28,38], t[19] = 4.20, pTFCE < 0.05 ROI corrected). We also found a significant effect in right visual cortex ([x,y,z] =[36, -94,20], t[19] = 5.70, pTFCE < 0.05 ROI corrected). Importantly, none of these regions showed significant differences between variability block (all pTFCE's >.19), neither did a whole brain corrected exploratory analysis (minimum pTFCE =.5). This shows that cue-induced representations of the target identity were not affected by the predictive uncertainty manipulated across *variability conditions.*

To test whether these regions shared information about the target identity during the delay we utilized information connectivity, a multivariate version of functional connectivity. We extracted the trial-by-trial decoding distance from the hyperplane for voxels which significantly decoded the target identity (threshold t(19)=2.54, p<.01, uncorrected). These distances can be taken as measure of the "decoding strength" for target identities on each trial (Schuck & Niv, 2019; Witkowski et al., 2022). Decoding strength between IFJ, and DLPFC was significantly correlated (t(19)= 2.07, p<.05), and similarly DLPFC decoding strength was significantly correlated with that of visual cortex (t(19)=3.19, p<.01). IFJ and visual cortex showed a marginally significant correlation, (t(19)=1.67, p=.06). None of these connections were found to be modulated by the *variability condition* (all t(19)'s < 1.47, all p's > .15) suggesting that the strength of connections in the network was not affected by uncertainty. These results show the

IFJ, DLPFC and visual cortex comprised a network of regions that code the target identity during the delay period which were not dependent on the uncertainty of the target color prediction.

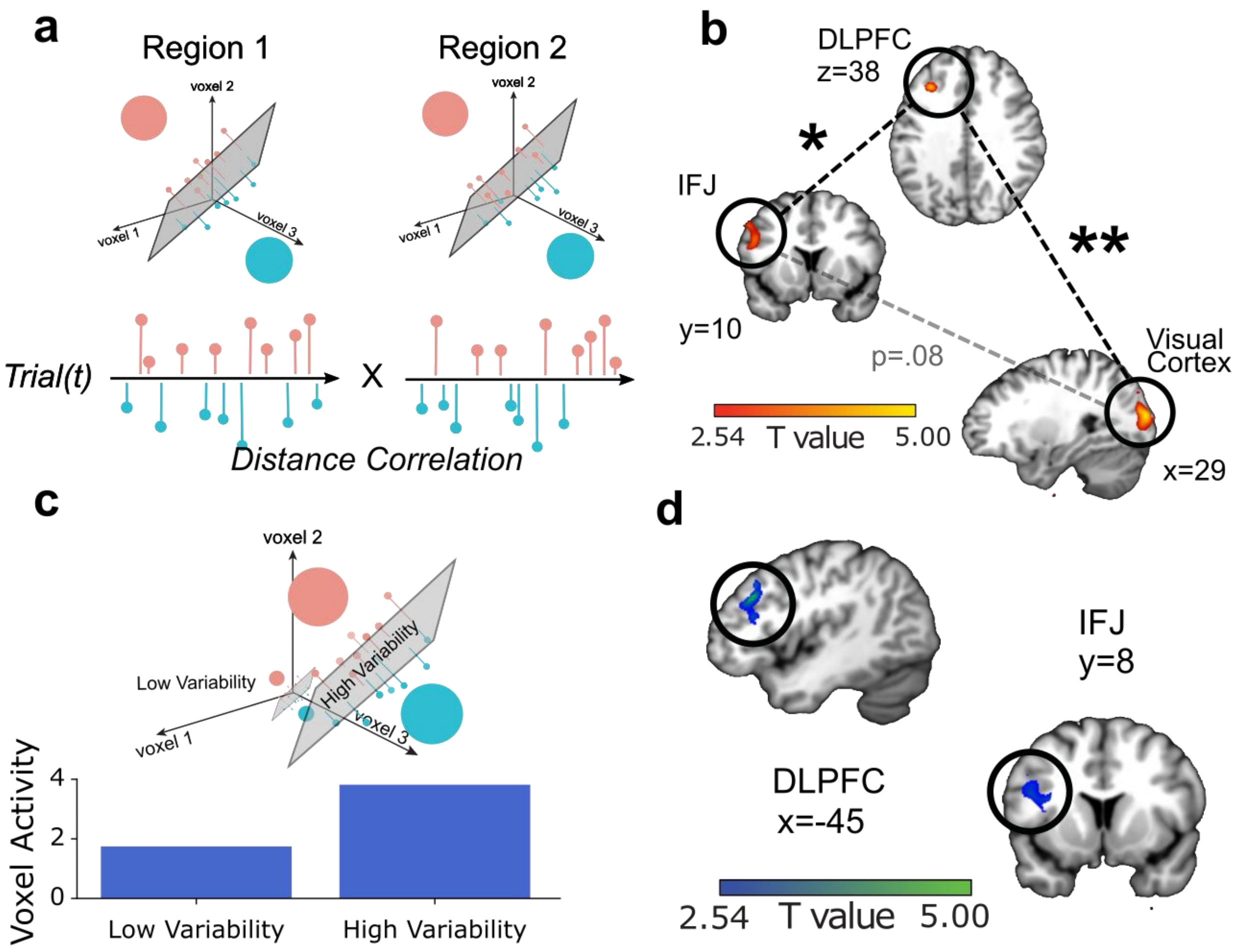

Figure 4.2 Univariate uncertainty codes overlay multivariate coding of target-identity in prefrontal cortex

(a) Analysis scheme for data shown in b. We calculated the trial-by-trial strength of information about target identities in each of two regions of interest. The strength of information about a target identity was taken as the distance between a voxel pattern in multidimensional space and the hyperplane separating categories. We then correlated these distance measures together as a measure of information connectivity. (b) Network of regions involved in coding multivariate representations of the target identity across conditions. Line connecting each area represent the results of information connectivity between regions (** = $p<.01$, * $p<.05$). (c) Hypothetical effect of univariate representation of uncertainty. Note that uncertainty does not change efficiency of decoding but scales only with voxel activity. (d) Results from univariate contrast between conditions showing higher activity in high-variability blocks compared to low-variability blocks. For illustration, maps display regions at a threshold of $t(19) = 2.54$, $p < 0.01$, uncorrected.

Predictive uncertainty coding is supported by univariate signals in prefrontal cortex

Our main hypothesis was that in DLPFC and IFJ would reflect uncertainty of visual features in the target template prior to visual search, in conjunction with target identity. As such, we tested the idea that in addition to multivariate target identity code there is a univariate signal in these regions which indicates variability of the target distribution on each trial. We began by constructing a GLM (see methods) then compared differences in the univariate activation during the delay period between high and low-variability conditions. Our hypothesis was that we would observe higher univariate activation prior to high-variability search trials compared to low-variability search trials, in our prefrontal areas of the target identity network. Our results confirmed this prediction showing higher activity in the left IFJ ([x,y,z] =[-44, 2, 22], $t[19] = 3.38$, pTFCE < 0.05 ROI corrected) and left DLPFC [x,y,z] =[-48, 30, 32], $t[19] = 4.76$, pTFCE < 0.05 ROI corrected; Fig 4.2d) in the high-variability condition compared to low-variability condition. Indeed, this pattern was true even after limiting the voxels included in the analysis to only those that significantly coded the multivariate pattern of target identity (both pTFCE's < 0.05 ROI corrected). However, we found no effect of variability, positive or negative, was found in the visual cortex ROI (all pTFCE's > .3). These results suggests that prefrontal areas specifically, encoded the predictive uncertainty of the target features prior to visual search.

Prediction-to-target and *target-to-distractor similarity have complementary univariate codes at the time of search*

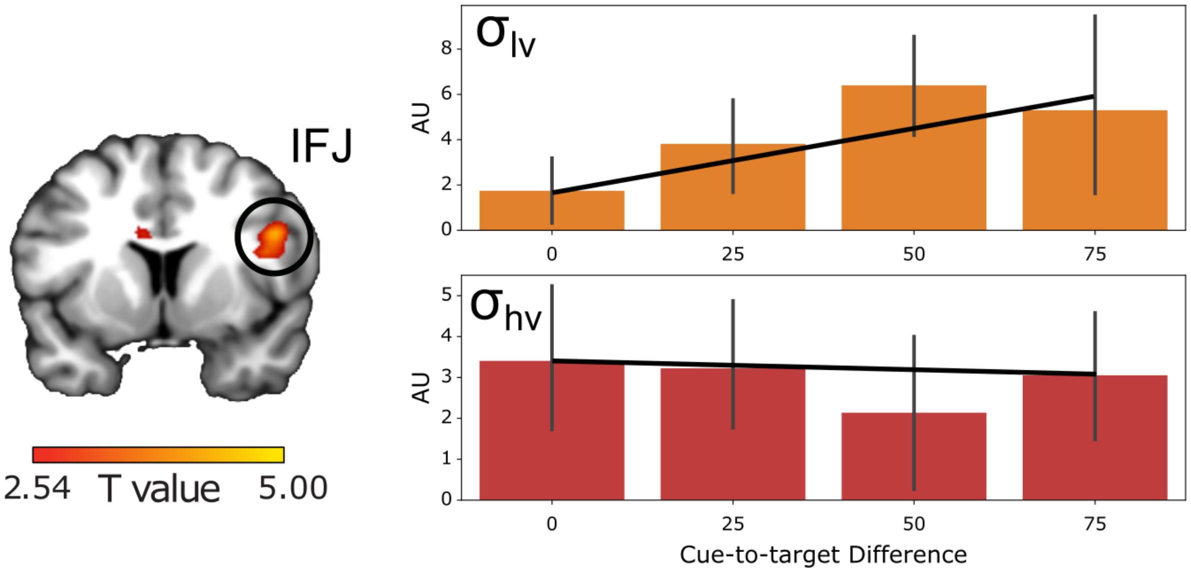

Figure 4.3. Univariate BOLD response to feature-based predictions errors determined by expected variability

Coronal slice of IFJ showing voxels with magnitude differences in response to cue-to-target differences (feature-based prediction errors) on each trial. Right side shows the BOLD response in this region binned by cue-to-target difference. BOLD responses for cue-to-target difference in the low-variability condition are shown on top, with those in the high-variability condition shown on the bottom.

Our final analysis focused on testing whether uncertainty moderated feature-based attentional processing in IFJ and DLPFC during search behavior. The pattern of behavioral data showed that lower *prediction-to-target difference* resulted in longer RTs, but that the slope of this function was shallower in the high-variability condition. This suggested that target colors which were very different from the mean were easier to find and identify when high variability was expected, compared to when low variability was expected. Therefore, we hypothesized that activity in IFJ and DLPFC would be moderated in the same way, indicating that these regions encoded the uncertainty of target features for guiding attention.

We first tested for regions that coded *prediction-to-target difference* in BOLD activity at the time of search, regardless of *variability condition*. However, of these regions, only the error responses in IFJ were negatively modulated by the uncertainty of the predicted stimulus ([x,y,z] =[42, 12, 26], t[19] = 3.13 pTFCE's < 0.05 ROI corrected; Fig 4.3). Neither DLPFC nor visual cortex voxels reached significance within ROIs (both pTFCE >.08 ROI corrected) and no other regions were significant at the whole brain level (all pTFCE>.2). However, all three ROIs showed increasing activation as the target was increasingly dissimilar from the mean color (IFJ [x,y,z] =[42, 12, 26], t[19] = 5.28, DLPFC [x,y,z] =[-32, 10, 28], t[19] = 4.51, and visual cortex [x,y,z] =[20, -92, -14], t[19] = 4.07, pTFCE's < 0.05 ROI corrected). These results implicate IFJ as a key region for coding variability-weighted target similarity at the time of search.

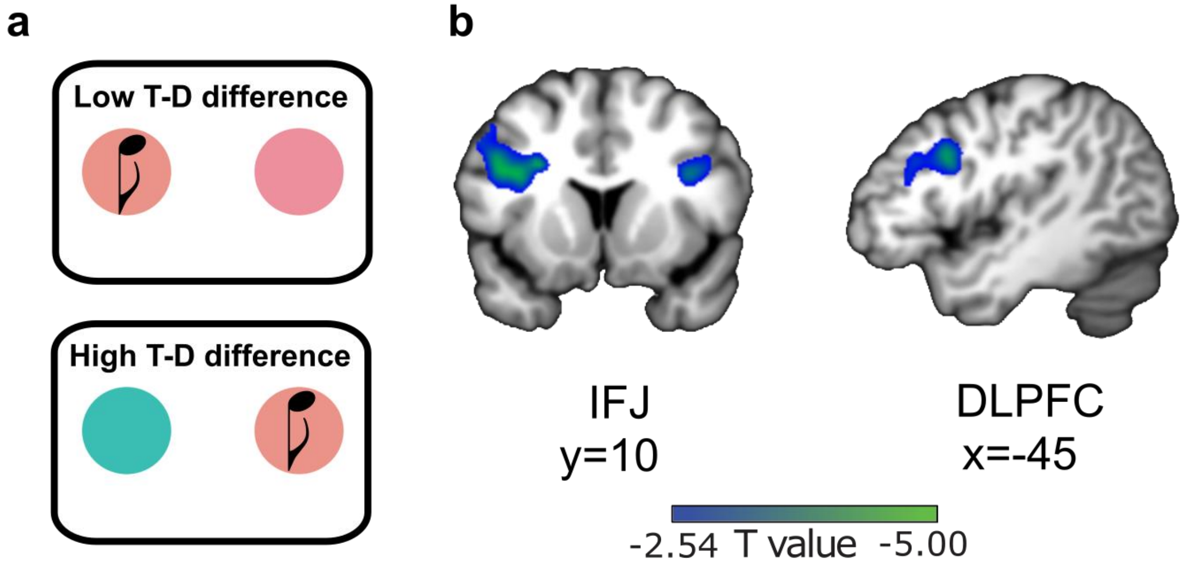

Figure 4.4 Univariate BOLD activity associated with target-to-distractor difference

(a) Example trials showing low target-to-distractor difference (top) and high target-to-distractor differences (bottom). (b) Coronal slice of IFJ and sagittal slice of DLPFC showing voxels in each region with higher activity in response to low target-to-distractor differences on each trial.

Finally, we looked for regions that coded the *target-to-distractor difference* at the time of search. Regions sensitive to this are likely to be involved in the resolution of competition between targets and distractors, showing *greater* activity when their difference is *smaller*. We observed the predicted effect in IFJ ([x,y,z] =[-40, 8, 28], t[19] = 4.30, pTFCE<.05 ROI corrected) and DLPFC ([x,y,z] =[-30, 8, 34], t[19] = 4.20, pTFCE<.05 ROI corrected; Fig 4.4b). This effect did not differ in strength across variability blocks (both pTFCE's<.1 ROI corrected). We found no other associations between brain activity and *target-to-distractor difference* at the whole brain level in the positive (all pTFCE's>.15) or negative directions (all pTFCE's>.35). Together, these results are the first to show that IFJ codes both variability-weighted target similarity, but also target-to-distractor differences that influence the ease of target identification.

Discussion

Understanding how uncertainty is integrated into target templates is essential to understanding how attention operates in the real-world. As we search for objects in dynamic environments, attention systems must consider the natural variability of objects and make predictions with this uncertainty in mind. The current experiments add to a burgeoning literature on how attentional systems use statistical knowledge by looking at the neural mechanisms that underly the representation of uncertainty and how it's integrated into the target template. We find that frontal regions coded both multivariate representations of predicted targets and a univariate representation tracking the uncertainty of the target's actual features. These results add to theories of attention by showing that IFJ and DLPFC are critical regions for integrating the known variability of dynamic search targets into predictions for the target template.

Our multivariate analysis showed that a network of regions spanning IFJ, DLPFC and visual cortex coded the predicted target identity during the delay period prior to search. This result is consistent with previous reports of these regions involvement in generating sensory predictions (Iglesias et al., 2013; Kok et al., 2013, 2017) and top-down influence during attentionally demanding tasks (Baldauf & Desimone, 2014; Giesbrecht et al., 2003; Jackson et al., 2021). The current experiments expand on these findings by showing that the strength of target representations fluctuate in tandem within this network. This suggests that these regions are not only coactive during attentionally demanding tasks but share information about the identity of the target prior to the onset of search.

Our key aim in the study was to test which regions of the brain coded both the identity of the predicted item and the uncertainty of those predictions. We predicted that DLPFC and IFJ would represent the uncertainty of visual features analogous to the way in which these region encode decision uncertainty (Badre et al., 2012; Michael et al., 2015; Summerfield et al., 2011; Tomov et al., 2020). Our results confirmed these predictions by showing that uncertainty was found as a univariate signal which increased when the distribution of possible targets became more variable. Interestingly, we did not observe any significant modulation of the multivariate identity code by uncertainty, standing in contrast to recent reports of joint working memory and uncertainty in this region (H.-H. Li et al., 2021; Van Bergen & Jehee, 2019). One important difference between this work and the results reported here is that our participant was cued to the target with previous associated tones and did not directly view the target stimulus. Thus, construction of the target template in this case may rely on retrieving long-term associations to make predictions, rather than directly maintaining a sensory specific representation. Future work may seek to specify when and how these different coding schemes are utilized. Nonetheless, these results add critical knowledge to theories of attention by showing how abstract, real-world attributes such as predictive uncertainty are integrated into the target template.

One interesting caveat to the symmetry of our finding for IFJ and DLPFC concerns the role of these two regions during search. While both regions responded to competition between targets and distractors, only IFJ coded the variance weighted target similarity, suggesting that this region-maintained information about the predicted target and uncertainty through the search period. DLPFC, however, only coded information about the absolute difference between targets and distractors. This divergence suggests possible differences in the role of IFJ and DLPFC during search. The pattern of results in IFJ suggests it may be more closely tied to global feature-based attention (Meyyappan et al., 2021; Zhang et al., 2018), because it takes into account the distribution of features that must be attended, which serves target identification. This is supported by the observation that IFJ is the only region that showed the same pattern of results as participant’s search response times. However, that DLPFC activity in response to prediction errors did not differ across *variability condition* suggests it may play a more general role in error detection that is independent of the target variability.

Overall, this study adds important knowledge to neural theories of attention by illuminating the neural mechanisms underlying the integration of uncertainty into target templates during visual search. By examining the multivariate and univariate representations of predicted targets and uncertainty, we show that IFJ and DLPFC, are critical to integrating the known variability of target features into predictions for the target template. Furthermore, we revealed that the activity of IFJ during search is moderated by predicted uncertainty. This highlights the role of this region in implementing this statistical knowledge in service of feature-based attention during search. This extends previous studies by showing that prefrontal cortex supports predictive representation of dynamic targets, and flexibly represents the uncertainty of those features to support efficient attentional processing.

References

Badre, D., Doll, B. B., Long, N. M., & Frank, M. J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron*, *73*(3), 595–607. https://doi.org/10.1016/j.neuron.2011.12.025

Bae, G. Y., Olkkonen, M., Allred, S. R., & Flombaum, J. I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *Journal of Experimental Psychology: General*, *144*(4), 744–763. https://doi.org/10.1037/xge0000076

Baldauf, D., & Desimone, R. (2014). Neural Mechanisms of Object-Based Attention. *Science*, *1268*(April), 424–428.

Bedini, M., & Baldauf, D. (2021). Structure, function and connectivity fingerprints of the frontal eye field versus the inferior frontal junction: A comprehensive comparison. *European Journal of Neuroscience*, *54*(4), 5462–5506. https://doi.org/10.1111/ejn.15393

Bichot, N. P., Heard, M. T., DeGennaro, E. M., & Desimone, R. (2015). A Source for Feature-Based Attention in the Prefrontal Cortex. *Neuron*, *88*(4), 832–844. https://doi.org/10.1016/j.neuron.2015.10.001

Bravo, M. J., & Farid, H. (2016). Observers change their target template based on expected context. *Attention, Perception, & Psychophysics*, *78*, 829–837. https://doi.org/10.3758/s13414-015-1051-x

Buschman, T. J., & Miller, E. K. (2007). Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and. *Science*, *315*(March), 1860–1863.

Chang, C. C., & Lin, C. J. (2011). LIBSVM: A Library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3). https://doi.org/10.1145/1961189.1961199

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

D'Esposito, M., Postle, B. R., & Rypma, B. (2000). Prefrontal cortical contributions to working memory: Evidence from event-related fMRI studies. *Experimental Brain Research*, *133*(1), 3–11. https://doi.org/10.1007/s002210000395

Duncan, J., & Humphreys, G. W. (1989). Visual Search and Stimulus Similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10.3758/BF03211797

Ester, E. F., Sprague, T. C., & Serences, J. T. (2015). Parietal and Frontal Cortex Encode Stimulus-Specific Mnemonic Representations during Visual Working Memory. *Neuron*, *87*(4), 893–905. https://doi.org/10.1016/j.neuron.2015.07.013

Geng, J. J., & Witkowski, P. (2019). Template-to-distractor distinctiveness regulates visual search efficiency. *Current Opinion in Psychology*, *29*, 119–125. https://doi.org/10.1016/j.copsyc.2019.01.003

Giesbrecht, B., Woldorff, M. G., Song, A. W., & Mangun, G. R. (2003). Neural mechanisms of top-down control during spatial and feature attention. *NeuroImage*, *19*(3), 496–512. https://doi.org/10.1016/S1053-8119(03)00162-9

Glasser, M. F., Smith, S. M., Marcus, D. S., Andersson, J. L. R., Auerbach, E. J., Behrens, T. E.

J., Coalson, T. S., Harms, M. P., Jenkinson, M., Moeller, S., Robinson, E. C., Sotiropoulos, S. N., Xu, J., Yacoub, E., Ugurbil, K., & Van Essen, D. C. (2016). The Human Connectome Project's neuroimaging approach. *Nature Neuroscience*, *19*(9), 1175–1187. https://doi.org/10.1038/nn.4361

Hansmann-Roth, S., Þorsteinsdóttir, S., Geng, J. J., & Kristjánsson, Á. (2022). Temporal integration of feature probability distributions. *Psychological Research*, *86*(6), 2030–2044. https://doi.org/10.1007/s00426-021-01621-3

Hout, M. C., & Goldinger, S. D. (2015). Target templates: the precision of mental representations affects attentional guidance and decision-making in visual search. *Attention, Perception, and Psychophysics*, *77*(1), 128–149. https://doi.org/10.3758/s13414-014-0764-6

Iglesias, S., Mathys, C., Brodersen, K. H., Kasper, L., Piccirelli, M., denOuden, H. E. M., & Stephan, K. E. (2013). Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning. *Neuron*, *80*(2), 519–530. https://doi.org/10.1016/j.neuron.2013.09.009

Jackson, J. B., Feredoes, E., Rich, A. N., Lindner, M., & Woolgar, A. (2021). Concurrent neuroimaging and neurostimulation reveals a causal role for dlPFC in coding of task-relevant information. *Communications Biology*, *4*(1), 1–16. https://doi.org/10.1038/s42003-021-02109-x

Kok, P., Brouwer, G. J., van Gerven, M. A. J., & de Lange, F. P. (2013). Prior expectations bias sensory representations in visual cortex. *Journal of Neuroscience*, *33*(41), 16275–16284. https://doi.org/10.1523/JNEUROSCI.0742-13.2013

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron*, *75*(2), 265–270. https://doi.org/10.1016/j.neuron.2012.04.034

Kok, P., Mostert, P., & De Lange, F. P. (2017). Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(39), 10473–10478. https://doi.org/10.1073/pnas.1705652114

Li, H.-H., Sprague, T. C., Yoo, A. H., Ma, W. J., & Curtis, C. E. (2021). Joint representation of working memory and uncertainty in human cortex. *Neuron*, 2021.04.05.438511. https://doi.org/10.1016/j.neuron.2021.08.022

Li, S., & Yang, F. (2012). Task-dependent uncertainty modulation of perceptual decisions in the human brain. *European Journal of Neuroscience*, *36*(12), 3732–3739. https://doi.org/10.1111/ejn.12006

Lo, S., & Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171–1187. https://doi.org/10.3389/fpsyg.2015.01171

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*(4), 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes : Evidence from eye movements. *Journal of Vision*, *9*(11). https://doi.org/10.1167/9.11.8.

Meyyappan, S., Rajan, A., Mangun, G. R., & Ding, M. (2021). Role of inferior frontal junction (ifj) in the control of feature versus spatial attention. *Journal of Neuroscience*, *41*(38), 8065–8074. https://doi.org/10.1523/JNEUROSCI.2883-20.2021

Michael, E., De Gardelle, V., Nevado-Holgado, A., & Summerfield, C. (2015). Unreliable evidence: 2 sources of uncertainty during perceptual choice. *Cerebral Cortex*, *25*(4), 937–

947. https://doi.org/10.1093/cercor/bht287

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, *16*(16), 5154–5167. https://doi.org/10.1523/jneurosci.16-16-05154.1996

Noah, S., Meyyappan, S., Ding, M., & Mangun, G. R. (2022). Anticipatory attention is a stable state induced by transient control mechanisms. *Frontiers in Human Neuroscience*, *16*(July), 1–12. https://doi.org/10.3389/fnhum.2022.965689

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences*, *15*(7), 327–334. https://doi.org/10.1016/j.tics.2011.05.004

Press, C., Kok, P., & Yon, D. (2019). The Perceptual Prediction Paradox. *Trends in Cognitive Sciences*, *24*(1), 12–24. http://eprints.bbk.ac.uk/2588/1/2588.pdf

Press, C., Kok, P., & Yon, D. (2020). Learning to Perceive and Perceiving to Learn. *Trends in Cognitive Sciences*, *24*(4), 260–261. https://doi.org/10.1016/j.tics.2020.01.002

Sallet, J., Mars, R. B., Noonan, M. P., Neubert, F.-X., Jbabdi, S., O'Reilly, J. X., Filippini, N., Thomas, A. G., & Rushworth, M. F. (2013). The organization of dorsal frontal cortex in humans and macaques. *Journal of Neuroscience*, *33*(30), 12255–12274. https://doi.org/10.1523/JNEUROSCI.5108-12.2013

Schmidt, J., & Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *Quarterly Journal of Experimental Psychology*, *62*(10), 1904–1914. https://doi.org/10.1080/17470210902853530

Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. *Science*, *364*(6447). https://doi.org/10.1126/science.aaw5181

Summerfield, C., Behrens, T. E., & Koechlin, E. (2011). Perceptual classification in a rapidly-changing environment. *Neuron*, *71*(4), 725–736. https://doi.org/10.1016/j.neuron.2011.06.022.Perceptual

Tomov, M. S., Truong, V. Q., Hundia, R. A., & Gershman, S. J. (2020). Dissociable neural correlates of uncertainty underlie different exploration strategies. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15766-z

Van Bergen, R. S., & Jehee, J. F. M. (2019). Probabilistic Representation in Human Visual Cortex Reflects Uncertainty in Serial Decisions. *The Journal of Neuroscience*, *39*(41), 8164–8176. https://doi.org/10.3389/fnhum.2020.580581

Vickery, T. J., King, L. W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, *5*(1), 81–92. https://doi.org/10.1167/5.1.8

Wang, L., Mruczek, R. E. B., Arcaro, M. J., & Kastner, S. (2015). Probabilistic maps of visual topography in human cortex. *Cerebral Cortex*, *25*(10), 3911–3931. https://doi.org/10.1093/cercor/bhu277

Witkowski, P. P., & Geng, J. J. (2019). Learned feature variance is encoded in the target template and drives visual search. *Visual Cognition*, 1–15. https://doi.org/10.1080/13506285.2019.1645779

Witkowski, P. P., & Geng, J. J. (2022). Attentional Priority Is Determined by Predicted Feature Distributions. *Journal of Experimental Psychology: Human Perception and Performance*, 1–12. https://doi.org/10.1037/xhp0001041

Witkowski, P. P., Park, S. A., & Boorman, E. D. (2022). Neural mechanisms of credit assignment for inferred relationships in a structured world. *Neuron*, 2021.12.22.473879. https://doi.org/10.1016/j.neuron.2022.05.021

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. In *Psychonomic Bulletin and Review* (Vol. 28, Issue 4). Psychonomic Bulletin & Review. https://doi.org/10.3758/s13423-020-01859-9

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 1–8. https://doi.org/10.1038/s41562-017-0058

Woodman, G. F., Carlisle, N. B., & Reinhart, R. M. G. (2013). Where do we store the memory representations that guide attention? *Journal of Vision*, *13*(3), 1–17. https://doi.org/10.1167/13.3.1

Yang, H., & Zelinsky, G. J. (2009). Visual search is guided to categorically-defined targets. *Vision Research*, *49*(16), 2095–2103. https://doi.org/10.1016/j.visres.2009.05.017

Zanto, T. P., Rubens, M. T., Bollinger, J., & Gazzaley, A. (2010). Top-down modulation of visual feature processing: The role of the inferior frontal junction. *NeuroImage*, *53*(2), 736–745. https://doi.org/10.1016/j.neuroimage.2010.06.012

Zhang, X., Mlynaryk, N., Ahmed, S., Japee, S., & Ungerleider, L. G. (2018). The role of inferior frontal junction in controlling the spatially global effect of feature-based attention in human visual areas. *PLoS Biology*, *16*(6), 1–28. https://doi.org/10.1371/journal.pbio.2005399

Chapter 5: General Disscussion

The concept target template is core construct in all models of attention, yet the role of the target template has been tested in a limit set of conditions. Chiefly, previous works have focused on how the identity of a perfectly known target is stored, and the consequences of failures to maintain that representation in memory (Hout and Goldinger, 2015; Malcolm and Henderson, 2009; Schmidt and Zelinsky, 2009). Few, if any, have tested how the target template adapts to uncertainty in target features and makes predictions about how the target *will* look in the future. The current series of experiments addresses how predictive uncertainty shapes representations of target templates, on both cognitive and neural levels. We show that observers readily learn statistical knowledge about targets and use this knowledge to make predictions about the likelihood of various target features. These predictions are then used to set priority for features which minimize uncertainty about the upcoming target, biasing eye-movements and reducing decision times when participants are certain about how the target will look. At the neural level, we show that uncertainty is coded independently from target identity in univariate activity which can allow for flexible recombination of specific features and uncertainty estimates in prefrontal cortex. Together these studies add critical insight to the body of knowledge concerning attentional templates, showing how these critical constructs respond to the variability so prevalent in the real world.

Target templates are shaped by learned target statistics

Across four experiments, we found that attentional priority is tied to statistical knowledge about the distribution of possible target features in the upcoming search trial. These results fit well with recent accounts of the target template, which hold that these representations are not merely veridical copies target objects, but task-adapted representations of “useful” information (Geng and Witkowski, 2019; Yu et al., 2023). We extend these finding by showing that distribution of target features is critical for shaping the contents of the target template, but also elaborate how this information should be used to select these contents. The variability of each feature determines whether this feature may be prioritized or attenuated, minimizing the uncertainty of the predicted target features (Fetsch et al., 2012; Foley et al., 2017). This can allow for rapid localization of targets in dynamic environments and tasks which abound in the real world, by focusing attention on information that is likely to useful for identifying targets (Gottlieb and Balan, 2010; Gottlieb et al., 2013). Thus, these results add new insights into theories of attention by elaborating how target templates adapt to statistical knowledge about uncertain target features.

The dependence of the target template on statistical regularities in the environment ties this construct to a broader literature on learning and suggests that aspects of the target template may develop over time. While theories of perceptual learning have made great strides in understanding how perceptual expectations are learned over time (Friston, 2010; Iglesias et al., 2013; Rao and Ballard, 1999), their ongoing relationship with the target template and attentional processing in general is less well understood. Although it is clear from the work above that the target template is greatly informed by targets feature regularities, its less clear how the target template can be updated when those regularities change. Experiment 1 in Chapter 3 can give us some clues, suggesting that participants quite flexibly update knowledge about the uncertainty of visual features and change priority for those features within the template accordingly. This is corroborated by the results in Chapter 4 which show participant flexibly shifting between anticipated target distributions in each block. This suggests that target templates are not only flexible in response to task-demand but update quite rapidly in response to changes in those demands.

However, future research may account for differences in this updating process for attended versus unattended factors of the environment. For example, more experiments are needed to determine if participants continue to learn about the feature variability of color once orientation is prioritized or simply ignore color until orientation is no longer useful and make an "exploratory shift" (Badre et al., 2012; Gershman, 2019) to test if reliability of this other information source is now lower? Furthermore, how can changes in the distractor environment be rapidly integrated into the template if this information is suppressed during the task (Geng et al., 2019; Won et al., 2021)? Interestingly, DLPFC and IFJ, which have been shown to encode template information in chapter 4, have also been linked to belief updates during learning about the environment and perceptual regularities (Iglesias et al., 2013; Witkowski et al., 2022). This suggesting that regions which encode the template have direct access to information about changes in the environment. Yet, how effectively and how quickly this information can be utilized and for what information will need to be tested.

Remembered versus Predictive Templates

The statistical sensitivity of target templates to task demands has, in part, influenced theories about the nature memory representations, ultimately leading to the idea of "premembering" (Nobre and Stokes, 2019). The core of this idea is that task-demands effect memory representations as early at the time of encoding, such that memories are formatted for the upcoming task rather coding veridical information about previously experienced stimuli (Boettcher et al., 2021; van Ede et al., 2020, 2021). Our behavioral results in Chapters 2 and 3 corroborate this account by showing that the remembered features of target items are more precise for low-variability features compared to high-variability features, suggesting that working memory representations of target features are responsive to task-demands. One caveat to our findings is that while we observed increased precision for low-variability features, our results in Chapter 3 showed that these responses did not track with learned variability. This suggests the memory representations seem to be biased for the prioritized (low variability) feature but did not track its specific variability or relative prioritization.

Rather, when we asked participants to *predict* the likelihood of target features, responses reproduced the relative likelihood of target features tracking with the learned variability. This method of probing expectations of target features in the template was, to our knowledge, unique to our experiment at the time of publication and suggests a tight link between target predictions and target templates. These results indicate predictions about upcoming targets in uncertain situations and memories for target features may rely on different information, although whether these are separate representations entirely or similar representations that can be utilized for different means remains to be determined. The idea of having multiple target template representations for different aspects or stages of visual search is not new to theories of attention (Wolfe, 2021; Wolfe et al., 1989; Yu et al., 2023). However, future work will be needed to confirm whether predictions and memories are truly separate representations, or merely the same representation that is accessed in different ways depending on the task.

Neural encoding of the target template

The experiment in chapter 4 was aimed at understanding how uncertainty was represented within the target template. Our results show that IFJ and DLPFC significantly coded both target identity and uncertainty, but whereas target identity was coded in multivariate patterns, uncertainty was coded as an overall increase in activation when target features were uncertain. This suggests that the uncertainty of the target predictions is carried in prefrontal cortex as independent representations from target identity, which may add flexibility to target template representations. For example, increased uncertainty can be scaled as much as need to reflect the underlying target statistics without changing the readout from the population pattern. These results add critical knowledge to theories of attention by illuminating the neural substrates of the target template and showing how these regions represent knowledge about the uncertainty of target features.

One important question that is left unanswered by this study, however, is how uncertainty is coded in the brain for multifeatured targets. In Badre et al (2016) the authors found that DLPFC coded absolute uncertainty, an effect which we replicated here, but also that rostrolateral prefrontal cortex coded relative uncertainty between decision options. This could imply that different neural regions are required for representing the relative uncertainty of features, versus the total uncertainty of the target. We were not able to test that question here because we only used single feature targets. Nonetheless, out study shows the basic mechanisms of how uncertainty is coded in the target template. Subsequent work may focus on how these mechanisms are extended to multifeatured target templates and what neural regions are required to do so.

Encoding abstract task features within the target template

One interesting question raised by the current experiments is whether the target template is limited to encoding only details about the target identity, or whether the definition should be extended to abstract task-variables. Although historically, the target template has been broadly defined as the information needed to identify a target object and set sensory gain (Duncan and Humphreys, 1989; Geng and Witkowski, 2019), this has been tested in a constrained set of conditions which directly relate to sensory information about the target. For example, when looking for friend who wears a read jacket, you may have a template for “red” which gives red areas of the visual scene higher priority.

Yet, the definition of “information needed to identify the target and set priority” can be much broader than this. Work on visual search has shown that this process can be moderated by the known spatial associations between targets and other elements of the environment (Zhou and Geng, 2023), semantic relationships (Nah and Geng, 2022), or meaningfulness (Henderson and Hayes, 2017). The current series of experiments add to this list by showing that inferred quantities like the uncertainty of target features are encoded target templates and actively utilized to set attentional priority during search. Importantly, uncertainty estimates in these tasks are not specific to any one visual feature. Rather, they are the accumulation of learning about a feature dimension or target properties which can then be generalized to all other features within that dimension. Learning that the color of your friend’s red jacket is unreliable because she has many colors of jacket is not tied to a specific representation of color in the template - it is tied to inferred knowledge about your friend. This was shown throughout the 5 experiments described here, suggesting that abstract knowledge about the reliability of target properties was actively integrated into the target template to set priority and even make target identification decision. Thus, these findings join a growing body of literature suggesting that target templates include abstract information about target properties and not merely tied to sensory representations.

Conclusion

The concept of target templates has been a core construct in all models of attention, but its role has been tested in a limited set of conditions. Our results strongly argue that the statistic of the visual environment, and particularly the uncertainty of target features, are a critical component for constructing target templates in dynamic environments. This current series of experiments has addressed how predictive uncertainty shapes representations of target templates, on both cognitive and neural levels. The results have shown that attentional priority is tied to statistical knowledge about the distribution of possible target features in the upcoming search trial, which knowledge allows for rapid localization of dynamic real-world targets. Furthermore, the results suggest that target templates are not only flexible in response to task-demands but also rapidly update in response to changes in the uncertainty, and thereby task-relevance, of features. Overall, these findings add important insight to theories attentional templates, showing how these critical constructs respond to the variability so prevalent in the real world.