MEMORY AND PREDICTION IN A CHANGING ENVIRONMENT

By

TALIA ROBBINS

A thesis submitted to the

Graduate School - New Brunswick

Rutgers, the State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Masters of Science

Graduate Program in Psychology

Written under the direction of

Dr. Pernille Hemmer

And approved by

____

____

____

New Brunswick, New Jersey

January, 2016

ABSTRACT OF THE THESIS

Memory and Prediction in a Changing Environment

by TALIA ROBBINS

Thesis Director:

Dr. Pernille Hemmer

This study aims to understand the effect of changes in the environment on memory and prediction. Previous work has suggested that when recalling events from memory, people use the entire distribution of information they have seen, and not a more recent set. However this work has only tested situations in which the distribution changes slowly, not those in which people encounter larger, more distinct changes in the stimulus distribution. Using an established experimental paradigm in categorical perception, I test memory and prediction in order to quantify performance changes with delineated temporal changes to the distribution of stimuli. I assess whether people use multiple belief distributions for decision making regarding the likely occurrence of future stimuli. I use a Bayesian approach to model how people update their beliefs, and the influence of changing beliefs on memory. This framework assumes that beliefs are a weighted combination of a prior belief and new evidence. Specifically, I will fit three models and discuss the qualitative fit to data. I will conclude with the use of a mixture model—which combines information from an overall distribution and a current distribution—and find that it provides a good explanation of participant behavior in this paradigm.

Keywords: Belief Updating; Decision Making; Bayesian Models; Memory; Prediction.

TABLE OF CONTENTS

1. Introduction

Our beliefs influence the way we navigate the world at all times, and our understanding of the world is predicated on these beliefs being aligned with the statistical regularities of our environment. In this context, beliefs refer to our knowledge and expectations about an environment. Real world contexts are constantly varying, and good decision makers must constantly adjust their beliefs to track environmental change. For instance, as the seasons change, people must adjust their expectations to track changes in the weather. People appear to have clear representations of their environment (a.k.a. beliefs), and use this information optimally in a broad range of cognitive tasks such as reasoning (Oaksford & Chater, 1994), generalization (Tenenbaum & Griffiths, 2001), memory (Anderson 1990), and categorization (Huttenlocher, Hedges, & Vevea, 2000). In this paper we seek to understand both how changes in the environment affect belief representations and in turn how these beliefs influence memory and decision making.

Importantly, if our beliefs about the environment are misaligned with reality, our behaviors will be maladaptive. Beliefs and intuitions are often developed as a result of a person's life experiences; however, such subjective experiences have been found to bias decision processes and be unreliable indicators of judgment accuracy (Kahneman & Klein, 2009). Many investigations in the field of decision making have found that human choice and behavior are not rational. There are two ways in which intuition can go wrong: 1) the environment is not regular and therefore the person cannot develop a good sense of the environmental regularities; or, 2) the person does not have adequate time to learn the statistical regularities of the environment. Yet, even when there are rules for determining the best course of action (that the person is aware of), people tend to use intuition, which leads to suboptimal decision making.

While these findings suggest that beliefs and intuitions are often miscalibrated and lead to bias and errors, a number of studies have also shown that peoples' beliefs are well-calibrated to the statistical regularities of the natural world, and that their predictions are, on average, quite accurate relative to the true statistics of the environment (e.g., Brady & Oliva, 2008; Griffiths & Tenenbaum, 2006; Hemmer & Persaud, 2014; Huttenlocher, Hedges, & Duncan, 1991; Huttenlocher, Hedges, & Vevea, 2000). These beliefs, in turn, influence performance. For example, having prior beliefs and expectations about the stimulus distribution has been shown to improve average recall (Huttenlocher, Hedges, & Duncan, 1991; Huttenlocher, Hedges, & Prohaska, 1992; Huttenlocher, Hedges, & Vevea, 2000). Using categorical perception tasks, Huttenlocher and colleagues also found that people quickly learn the underlying stimulus distribution, and suggested that people can use this knowledge to fill in noisy and incomplete memories for events (Huttenlocher et al., 1991). In addition, they found that responses regressed toward the mean of the overall stimulus distribution, thereby improving average recall. These studies suggest that people use their beliefs to make sense of their environment and appear to use this information optimally in reconstruction from memory.

Additionally, Huttenlocher and colleagues have suggested that the integration of new information influences memory towards an overall accumulation of evidence, and not at a more local timescale. The behavioral evidence for this prediction comes from another series of studies in categorical perception seeking to rule out sequential dependencies as the underlying mechanism in the regression effect. When recalling events from memory, people do not appear to use the most current state of the environment when the frequency distribution of the underlying environment shifts over time. Rather they appear to fill in noisy memory representations with knowledge of the overall stimulus distribution, such that recall is biased towards the running mean of the overall, long-term experience, and not the current distribution (Duffy, Huttenlocher, Hedges & Crawford, 2010). These experiments, however, only illustrated this pattern with subtle changes made over time, and not with more noticeable shifts in the distribution. This is in line with the Kahneman and Klein (2009) argument that the integration of information requires time and environmental consistency. This system is useful for subtle changes in real world settings such as when the seasons change, but may be less adaptive for more clearly delineated changes in the stimulus distribution.

Two important questions that remain unexplored are: 1) what is the effect of clearly delineated changes in the environment on subjective beliefs, as often experienced in real world settings; and 2) what is the influence of such changing beliefs on memory and decision making?

1.2 A Bayesian Approach to Understanding Beliefs

The type of behavior observed in the Huttenlocher studies is well modeled by a Bayesian approach, which assumes that noisy data in the mind is optimally combined with prior knowledge about the environment. Assuming that people do Bayesian inferences in their head, Bayes' rule gives a principled account of how people should update their beliefs in light of new evidence:

$$p(B|E) \propto p(E|B)p(B) \qquad \text{Eq (1)}$$

The posterior probability $p(B|E)$ gives the probability of the belief B given the observed evidence E. This posterior probability is based on a combination of $p(B)$, which is the prior probability (or strength) of the belief, and $p(E|B)$, which is the likelihood of observing the current evidence given the belief. After observing the evidence E, $p(B|E)$ becomes the new belief $p(B)$ in the next iterative time step.

This approach is useful because it characterizes the computational problem people face when trying to make sense of the world given evidence with varying degrees of uncertainty. Bayes' rule predicts a tradeoff between prior beliefs and observed evidence, such that when our prior belief is strong and we encounter weak evidence, our new belief will closely reflect the prior. Conversely, in situations with a weak prior belief and strong evidence, the new belief will closely reflect the evidence. In situations where both the evidence and prior belief are strong, the new belief will lie somewhere in between. In a standard belief-updating model, the prior distribution for a given belief should be combined with the data from the environment to create a posterior distribution, which should become the new belief.

1.3 A Model for the Influence of Beliefs on Memory

A Bayesian approach can also be used to characterize how our beliefs influence our memory when we are trying to learn a new environment. A simple Bayesian model of memory (see Figure 1) predicts that memory is a combination of our expectations about an environment (μ, τ) and the noisy samples retrieved from memory (y).

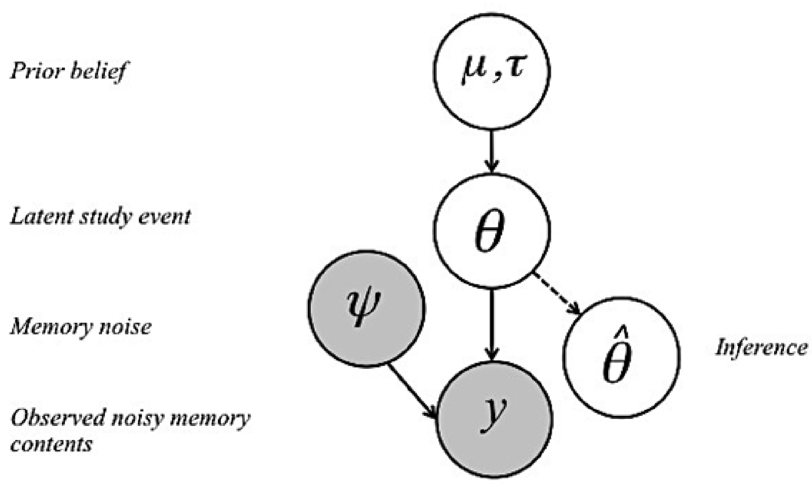

Figure 1. Graphical representation of a simple Bayesian model of memory. The model is a representation of inference in the head of the observer in a given task. Shaded nodes are observed variables for the observer. The inference task is to recall the now latent study event (unshaded node), which is a combination of the prior probability and noisy memory samples.

For example, when you cannot find your keys, you will be faced with recalling features θ of where you left your keys. Each time you encounter new features, θ, of the location of your keys this is assumed to lead to noisy representations y in episodic memory, where the memory representation is centered on the original location of the keys and is stored with some noise ψ. The resulting inference problem when trying to recall the location of your keys is $p(\theta| y, \psi, \mu, \tau)$, and where you remember them to be will be based on the posterior probability $p(\hat{\theta}| y, \psi, \mu, \tau)$, which describes how likely the location of the keys, θ, is given the noisy memory contents, y, and prior beliefs about where you usually leave your keys.

1.4 Experimental Objectives

This paper seeks to investigate some of the most fundamental components of human behavior: how we update our beliefs given changes in our environment, and how dynamically changing beliefs influence memory and prediction. We begin by experimentally quantifying the influence of changing beliefs on memory and prediction. We then use a Bayesian (a.k.a. rational) approach to model how changing beliefs affect recall, and ask why using beliefs in memory and prediction is an optimal strategy.

The aim of the experiment described in the next section is two-fold. The first aim is to find evidence for the influence of beliefs on episodic memory as a function of changes in the underlying environment. This aim seeks to replicate the regression to the mean effect observed in previous research, and to answer two questions: 1) how do people update their beliefs in an environment with delineated temporal changes to the distribution of stimuli; and 2) is there evidence that the regression effect changes as a function of the changing belief? The second aim is to find evidence for the influence of beliefs in prediction about the likely occurrence of future stimuli. This aim seeks to determine if people use the same belief for short-term and long-term prediction. That is, what constitutes the current belief that prediction is centered on, and do people use multiple beliefs distributions—one for the long-term running mean and one for the most recent stimulus environment?

1.5 Experiment

We used an established experimental paradigm from categorical perception (Huttenlocher et al., 1991) to assess memory for spatial location. In this task, participants were asked to recall the spatial location of a dot in a circle. While this paradigm relies on estimation of relatively artificial stimuli, for which participants are unlikely to have strong pre-experimental beliefs, the strength of the paradigm is that it has shown that beliefs can be created during the course of the experiment by training participants on the underlying stimulus distribution. That is, participants appear to update their beliefs about the overall distribution of dots and their recall judgments are biased by this belief about the dot environment. This paradigm lends itself to extensions that can be used to test changes in the stimulus distribution and assess prediction. As an extension of the Huttenlocher et al. (1991) paradigm, in this experiment participants will also make predictions for future dot locations. Furthermore, the number of presentations at a given dot location will be manipulated in order to simulate a changing stimulus environment where optimal performance can be achieved by using the belief about the current cluster of dots when reconstructing the study location from memory.

2. Method

2.1 Participants

Participants were recruited from Rutgers University, New Brunswick. There were eight participants in this study and they were compensated with $10 for their participation, which lasted approximately 30 minutes.

2.2 Procedure

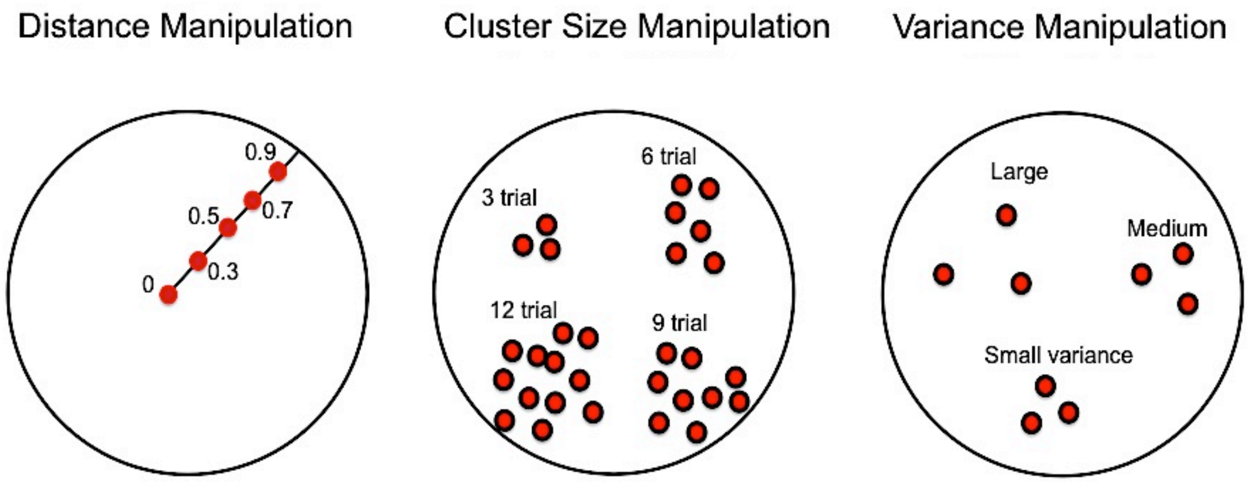

Figure 2. The three manipulations of dot locations presented to participants. The left circle represents the distance manipulation (either 0 .3 .5 .7 or .9 from the circle center to edge based on a 1 unit circle). The middle circle represents the number of trials presented to participants at a particular distance (either 3 6 9 or 12). The right circle depicts the variance of the cluster of dots (small, medium, or large).

In this experiment, participants were asked to study the location of a dot presented in a circle (see Figure 2) and reconstruct that location from memory. On the first 20 trials a sequence of dots were presented near the center of the circle in order to familiarize participants with the task. These trials were excluded from the analysis. Three manipulations on dot location and distribution were used to assess updating in this task: location (consisting of a distance measure across 24 angles around the circle), number of trials, and variance (henceforth referred to as clusters). The distances and angle measures were informed by Huttenlocher et al. (1991). Four different distances measuring out from the center of the circle to the circumference were chosen (0, .3, .5, .7, and .9), and represented in each quadrant. Dots were presented in clusters (3, 6, 9, or 12 presentations at a mean location), sampled from a multinomial normal distribution with a mean of a given radius and one of three variances (0.01, 0.04, and 0.06 in a unit circle) chosen respectively to represent weak, average, and strong spatial coherence. Each of the relative angles had a different distance, variance, and number of trials. The trial order was randomized, starting with 20 dots presented for training.

Each dot was viewed for one second followed by a combined visual mask and distractor task designed to remove the dot from participants' visual field and introduce uncertainty in the memory process. This mask consisted of a grid of black and white squares; after this mask was removed, an "X" appeared on the screen and participants were asked to report the color of the square (black or white) previously in that location. Data from the distractor task was recorded but not analyzed. After the completion of the distractor task, participants were asked to reconstruct the location of the dot from memory by clicking on the recall location in the circle.

After every three trials, participants were asked to make a prediction about a future dot location. Prediction trials alternated between prediction for the next trial and prediction for five trials from now. Each cluster was followed by a prediction for the expected dot location 10 trials from the current trial. This resulted in a total of 280 trials per participant: 80 prediction trials and 200 recall trials.

3. Results

3.1 Recall

Figure 3, bottom panel shows the error distribution (recalled minus studied dot location) averaged across all trials and participants. Participant responses were approximately normally distributed, with a mean centered on the true stimulus location (an error of 0) and a standard deviation of .1. Error distributions are commonly used in the literature to visualize response bias; however, aggregating across all data masks any possible underlying regression patterns. Therefore, we further visualized recall bias as a function of study radius for all trials (see top

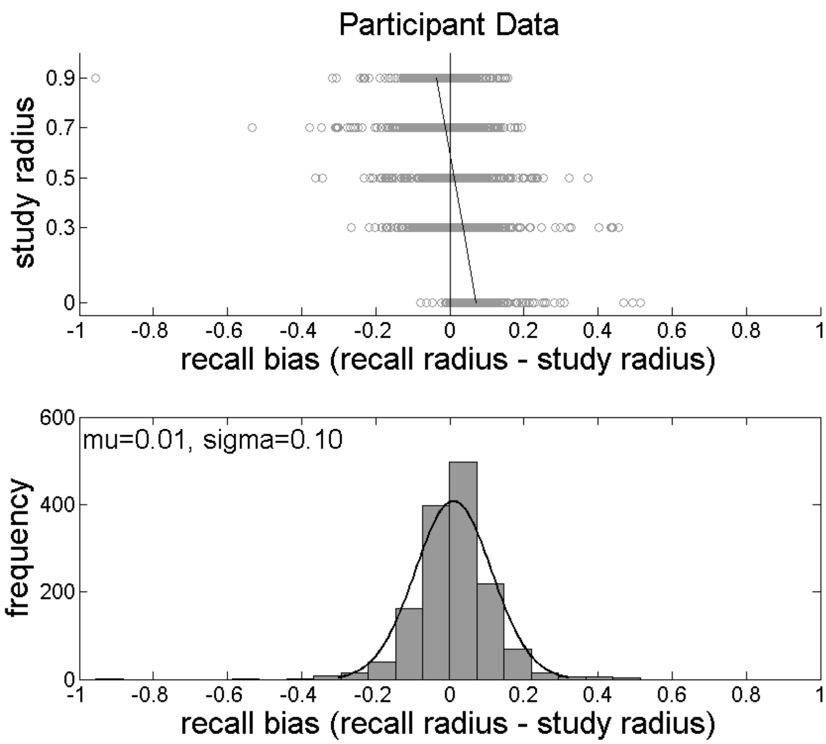

Figure 3. The top panel shows all the participant recall trials plotted with study radius as a function of recall bias. The bottom panel shows the error distribution (recalled minus studied dot location) for all trials.

panel Figure 3). This visualization clearly displays a regression pattern that is masked by this error distribution. We found that participant responses regressed towards the mean radius such that dots studied at smaller distances from the circle center were overestimated while dots studied at the larger distances from the circle center were underestimated. This replicates the findings of Huttenlocher et al. (1991).

We also sought to evaluate recall bias as a function of cluster size. Figure 4 shows the mean recall bias as a function of the five radius locations, separated by time step in a cluster (defined as trials within the current stimulus cluster) at trials 1-3, 4-6 and 7-12. This suggests that bias is reduced as the strength of evidence in a cluster increases; that is, after only one trial in a given cluster there was a greater regression to the overall mean, whereas after nine trials in a cluster there was significantly less regression to the overall mean location and accuracy was

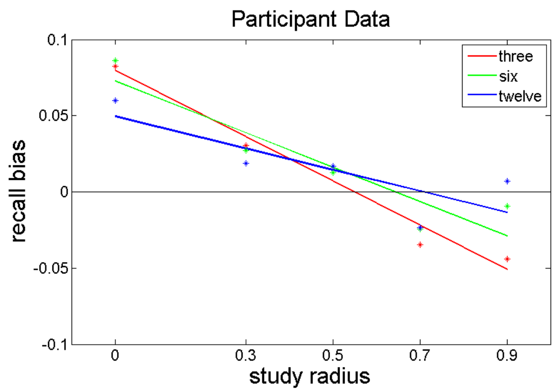

Figure 4. Shows recall bias as a function of study radius and trial number within a dot cluster (either trial 1-3, 4-6, or 7-12).

closer to the true mean of the current cluster. This suggests that participants were learning the underlying distribution of the dot cluster.

In order to further evaluate whether participants improved over the course of the cluster, we conducted an F-test for differences in variance between groups. There was a significant difference between recall bias for trials 1-3 and 4-6 ($p<.05$), recall bias for trials 4-6 and 7-12 ($p<.01$) and recall bias for trials 1-3 and 7-12 ($p<.01$). Predictions for locations ten trials in the future are the most interesting because they give some indication that subjects might hold multiple beliefs about the stimulus environment—one for the current cluster and one for the overall dot distribution.

3.2 Prediction

Figure 5 shows error distributions for the prediction trials. Here, prediction bias refers to the difference between the predicted future dot location and the mean of the current cluster. The top panel shows predictions for one trial in the future, where prediction bias is approximately normally distributed around the mean of the current cluster (an error of 0) and a standard deviation of .21. This indicates that participants believe that in one trial from now, the dot will be

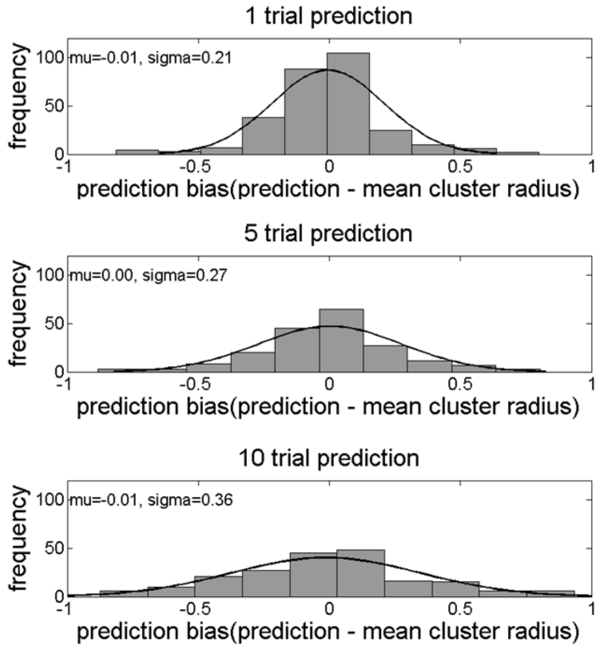

Figure 5. Shows the error distributions for prediction bias (predicted location minus the mean of the current cluster) for the three prediction types. The top panel is 1 trial in the future, the center panel is 5 trials in the future, and the bottom panel is 10 trials in the future.

in the same general location that it is now. The center panel shows predictions for 5 trials in the future, where participants begin to make predictions further from the current cluster mean. They are still centered on the mean of the current cluster, however the standard deviation increases to .27. For 10 trials in the future (bottom panel), participants are making predictions much further from the mean of the current cluster, showing that they do not believe the dot is likely to be in the same location it is now in 10 trials. They are still centered on the mean of the current cluster, but the standard deviation has now increased to .36. To evaluate whether the difference in variance between prediction-type was significant, we conducted an F-test for differences in variance between groups. There was a significant difference between predictions for 1 and 5 trials from now (p<.01), predictions for 5 and 10 trials from now (p<.01) and predictions for 1 and 10 trials from now (p <.01). Predictions for locations ten trials in the future are the most

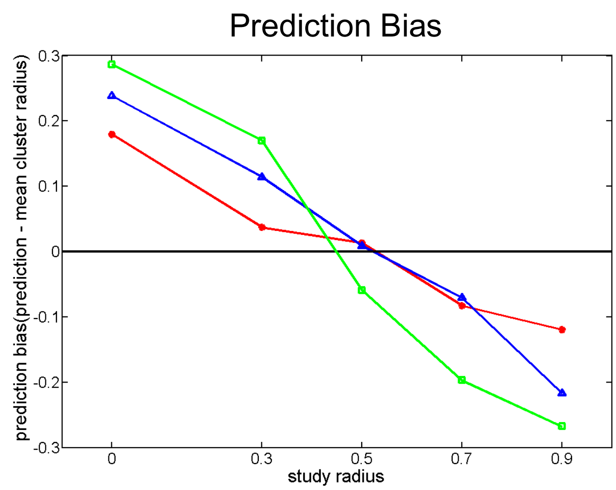

Figure 6. Shows prediction bias relative to the mean of the current cluster as a function of study radius and prediction type.

interesting because they give some indication that subjects might hold multiple beliefs about the stimulus environment—one for the current cluster and one for the overall dot distribution.

Further, Figure 6 shows prediction bias relative to the mean of the current cluster as a function of study radius and prediction type. This shows the same pattern as the error distributions whereby participants make predictions further from the current cluster for predictions further in to the future. Additionally, Figure 6 shows the same pattern observed in the recall portion of the task where dots near the edge of the circle lead to underestimated predictions, and dots near the center of the circle lead to overestimated predictions. This may indicate that predictions are biased toward .5, which is the overall mean of all trials seen. Taken together, this suggests that predictions for the next dot location were drawn from the belief about the current state of the environment as quantified by the mean location of the current cluster. For predictions of dot locations further in the future, participants no longer appeared to use their belief about the current dot environment; rather, they appeared to be using a different belief based on the overall stimulus distribution and biased to the circle center.

4. Discussion

The results suggest that participants may be using more than one prior distribution when recalling dot locations from memory, and making predictions for future dot locations. For predictions in the near future, participants may believe the environment is likely to be the same as it is right now, but for long-term predictions they may not believe that things will be the same and instead base predictions on long-term accumulated evidence. It appears that in making predictions for the distant future, participants rely primarily on the overall distribution of dots (the long running mean), but when they make predictions about the immediate future they rely on the local distribution of dots (the current cluster).

Results in the recall task indicate that as participants advance through trials within a given cluster of dots, they assign progressively more weight to the evidence from that local cluster when recalling the dot locations from memory. This evidence motivates our use of a Bayesian framework, which integrates multiple prior distributions, to model incremental belief updating. It was found that bias decreased as time steps in a cluster increased, and the inclusion of multiple prior distributions was further supported by results from the prediction trials. Next we will implement four models. The first assumes recall is based on a prior accumulated over all trials the participant has seen, henceforth called the global model. The second assumes a local prior accumulated only over the current cluster the participant is observing, henceforth called the local model. The last two models will assume recall is a combination of these two distributions. The third model assumes a static mixture of these two distributions, henceforth called the mixture model. The fourth assumes that the mixture of these two distributions is dependent on the number of trials a participant has seen within a cluster, henceforth called the mixture by trial model.

5. Modeling

Reframing the memory model from section 1.3, an observer in our experimental task is faced with recalling features θ of a stimulus presented at study (i.e. locations of dots). Based on our experimental design we will assume that these features are Gaussian distributed, $\theta \sim N(\mu, \tau)$, where μ and τ are the prior mean and precision of the dot locations. When a specific dot location θ is studied, we assume this leads to noisy representations y in episodic memory, where $y \sim N(\theta, \psi)$. That is, the memory representation is centered on the original studied dot location and is stored with some noise ψ, where ψ expresses the resemblance of the stored representation to the studied location. The goal of the observer on a test trial is to recall the studied dot location θ as best as possible using noisy samples y retrieved from memory. Extending Equation 1 from Section 1.2 to the memory task, the inference problem for the observer is $p(\theta| y, \psi, \mu, \tau)$. The posterior probability $p(\hat{\theta}| y, \psi, \mu, \tau)$ describes how likely dot locations θ are given the noisy memory contents y and prior beliefs about the dot locations. We assume that the observer has a prior belief that corresponds to the observed stimulus distribution in the experiment (i.e. the environmental statistics).

Standard Bayesian techniques (Gelman, Carlin, Stern & Rubin, 2003) can be used to calculate the mean of the posterior distribution:

$$\hat{\theta} = w\mu + (1 - w)\bar{y} \qquad \text{Eq (2)}$$

where $w=(1/\sigma_0^2)/[(1/\sigma_0^2)+(n/\sigma_m^2)]$ and n is the number of samples taken from episodic memory.

In this way, recall can be modeled as a weighted linear combination of beliefs and memory content, where the strength of the prior belief and episodic memory trades off as described in section 1.2. Here, the Bayesian model is applied to the experiment without directly estimating any parameters. Instead, it is assumed that the observer has a belief that corresponds to a mixture of the local and overall environmental statistics. The mean of the overall distribution μ_o was set to 0.5 for all radius locations. The precision for the overall distribution τ_o was set to the exact precision in the total set of observed data. The means for the local clusters μ_c were set to the true radius locations, and the precision τ_c was set to the mean precision for clusters at each radius. Finally, the memory noise ψ was set to 0.25 to reflect the true variance in all the trials presented to participants, and remains constant in all the models discussed below. The goal of this analysis, and all the model analyses presented here, is to compare the predictions of the Bayesian model and the empirical data at a qualitative level.

The first model we will explore is the global model, and assumes participants are using one long-running mean over all the trials they have seen, similar to the assumptions made by Duffy et al. (2010). This is the model introduced in section 1.3 (Figure 1). Importantly the prior is assumed to be based on the cumulative evidence from all trials. Therefore, μ is set to .5 and τ

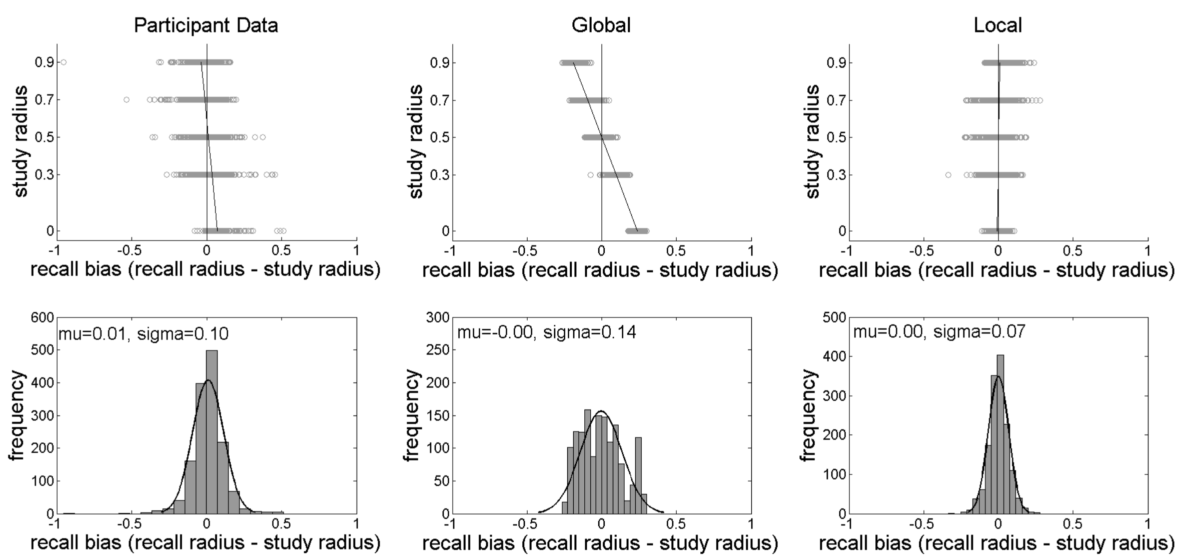

Figure 7. The left panel shows the data from participants, with recall bias as a function of study radius. The center panel shows simulated data from the global model, which assumes a prior based on all trials. The right panel shows simulated data from the local model, which assumes a prior based on the current cluster being presented.

is set to the true variance (.28) of all the trials (Figure 7 shows the performance from the global model in the center panel relative to the true data in the left panel). It is clear that this model does not provide a good qualitative fit to the data. This model fails to capture two important trends in the data: 1) the regression pattern observed in this model is too steep; and 2) the variance is too large.

The second model we explore is a local model, which assumes participants are using a mean for only the current cluster of dots. Therefore, the mean of the prior, μ, is set to the true mean of each cluster and τ is set to the true variance of each cluster. This model also does not provide a good qualitative fit to the data (see Figure 7, right panel). Specifically, it fails to capture the regression pattern observed in the data. This is because the model regresses to each individual cluster, which means performance can never be biased to the overall mean of .5, obscuring any influence of the overall mean.

Bearing in mind the poor qualitative fit of the first two models, the experimental results suggest that observers hold multiple beliefs about the environment: one for the local cluster, and one for the overall stimulus distribution. This can be modeled with a mixture model (as seen in Figure 8) where the mean and precision (μ, τ) of beliefs are a combination of overall (μ_o) and cluster specific (μ_c) beliefs, such that $\mu = z\mu_c + (1 - z)\mu_o$, and $\tau = z\tau_c + (1 - z)\tau_o$, where ($\mu_c$, τ_c) represents the belief associated with cluster i and (μ_o, τ_o) represents the overall belief about the stimulus distribution (Hemmer & Steyvers, 2009). The variable z weights the contribution of the cluster belief relative to the overall belief. First, we hand fit the weighting of the cluster specific contribution such that z is set to a fixed value of .8, meaning that more relative weight is given to the local cluster. This value was chosen because it provides the best qualitative fit in this implementation of the model.

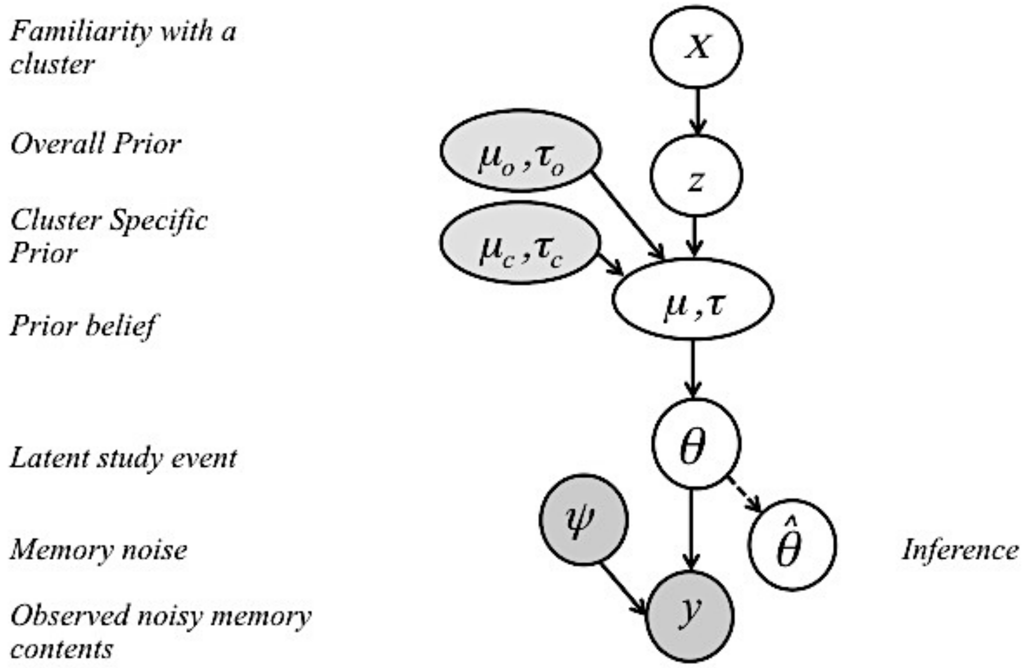

Figure 8. Graphical representation of the mixture model used to describe the recall data. This model is identical to the graphical model of memory in Figure 1, expanded to include a mixture component on top.

While this model does seem to provide a good qualitative fit to the data upon first inspection – for example, it both captures the variance and the regression pattern (see Figure 9, top center panel) – it fails to capture the pattern by which participant recall improves for later time steps in a cluster (see Figure 9, bottom center panel). Therefore, we implement the full model where we assume both a mixture and a progressive mixing component, which is informed by the time step within a cluster. This pattern would suggest that the mixing component should change relative to the time step within a cluster. With this in mind, the weighting is determined by $z \sim Bernouli(X_i)$, where χ_i is a constant that represents the familiarity of a cluster. In this way, familiar clusters lead to a belief that is more dependent on the cluster than the overall distribution. This implements the intuitive notion that for unfamiliar clusters it is unlikely that

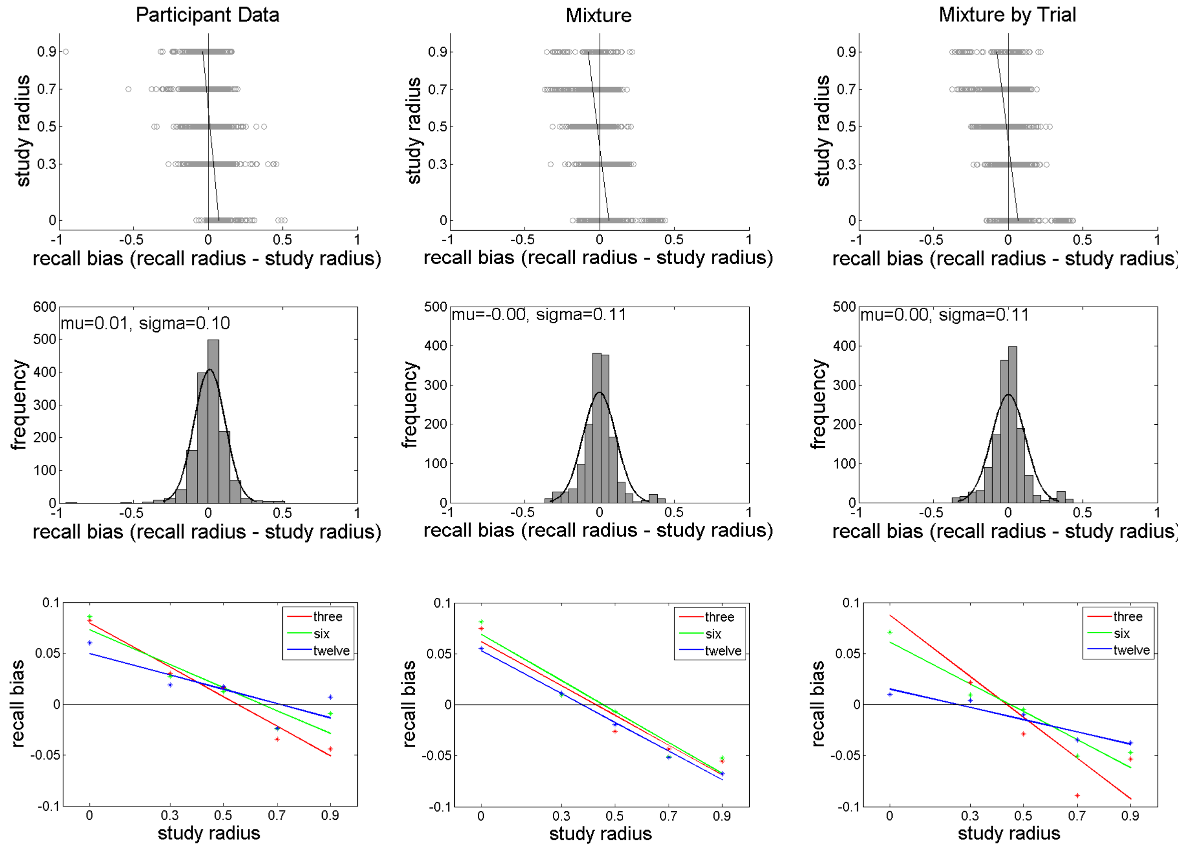

Figure 9. The left panel shows the data from participants, with recall bias as a function of study radius. The center panel shows simulated data from the mixture model, which assumes that the prior distribution is a combination of the entire distribution of dots the participant has seen and the most recent set of stimuli.

the cluster belief is reliable, and an individual's inference instead reverts to a higher-level belief based on the overall stimulus distribution.

On the first time step, the observer is assumed to have a belief about the mean stimulus location that is biased towards the overall distribution, but by the ninth time step is assumed to have a belief that is identical to the local cluster's environmental distribution (time steps 1-3, 4-6, 6-12 are simulated here). Therefore, χ (i.e., familiarity with the local cluster) was set to 0.7, .8, and .95 for each of the three successive time step groupings modeled here to simulate an increasing level of familiarity with the cluster environment.

As seen in Figure 9, the mixture model produces results that are qualitatively consistent with the responses given by human observers. This model is able to capture both the regression pattern found in the data (see Figure 9, top left panel), and the improvement in recall for later time steps in a cluster (see Figure 9, bottom left panel). It appears as though participants incrementally change their beliefs with an increasing number of time steps within a cluster. Results show that memory estimation errors can be explained by the use of beliefs about the environment, since smaller radius distances from the circle center were later recalled to be further away and larger radius distances were later recalled to be closer. Furthermore, beliefs change as a function of increasing familiarity with the underlying local environment.

The observer in our task was also asked to make predictions about future stimulus locations. The posterior predictive distribution of future dot locations $p(\theta_{future}| \theta)$ is determined by averaging the predictive probability across all possible values of beliefs weighted by the strength of the belief. The mean of the posterior predictive distribution can then be shown to be equal to the prior mean, with the variance drawn from both the variance of the observed stimulus and the uncertainty in the current belief. Therefore, prediction for a future stimulus is centered on the mean of the current belief. Yet, as demonstrated in the experimental results, this only holds for short-term predictions; for long-term predictions, people appear to use a mixture of the current belief about the cluster and the overall belief, similar to that of recall. It is now trivial to extend the Bayesian model to assume long-term predictions to be a mixture of belief, but that will be outside of the scope of the current paper.

6. General Discussion

The results indicate that people hold multiple beliefs that simultaneously affect decision making. When experiencing immediate, noticeable, shifts in a stimulus distribution, people use multiple beliefs about the stimulus environment (one about their current environment and one about the overall environment). This provides a more comprehensive understanding of how people integrate changing information from the environment. When combined with previous research from Duffy et al. (2010), these two lines of work together suggest that people use not only the entire distribution of stimuli in a stable, slowly changing environment but can also quickly adapt to changing environments when recalling events from memory.

We found that bias in recall decreased as time steps in a cluster increased, and that a mixture model provided a good explanation of this pattern. This is consistent with previous research suggesting that increased familiarity with a statistically regular environment is necessary for high accuracy (Kahneman & Klein, 2009). This also provides support for the initial hypothesis that people sample from multiple prior distributions when making future predictions.

Beliefs are one of the most influential components of human behavior. The results above indicate that people both apply beliefs learned over a lifetime and are sensitive to sudden environmental changes. This suggests that people possess the ability both to track slow changes, and to adapt quickly to more sudden changes. While this experimental paradigm was successful in simulating a real world belief-updating scenario, future investigations should be expanded into more realistic environments in order to further our understanding of how people make choices in other contexts—for example, how people remember and make predictions about their health. It is possible that people might hold a global distribution for their entire lifetime of health, and a local distribution for the illness they are currently experiencing. If this is the case, a mixture model might help elucidate how people understand their illness, and how they make predictions about future health states. With this in mind, the work presented here has important theoretical implications in a number of cognitive domains, including learning, memory, and decision making, as well as in applied settings such as health decisions.

Acknowledgements

This material has been presented and published in the proceedings of the Annual conference of the Cognitive Science society. The concept of this study was developed by T. Robbins and advisor, P. Hemmer. Stimulus creation, testing, and data collection were performed by undergraduate Y. Tang under the supervision of T. Robbins. Data analysis and interpretation were performed by T. Robbins and P. Hemmer. The thesis was drafted by T. Robbins and edited by advisor P. Hemmer.

This work was supported by a Rutgers University Research Council Grant Award.

References

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Erlbaum.

Brady T.F. & Oliva, A. (2008). Statistical learning using real-world scenes: Extracting categorical regularities without conscious intent. *Psychological Science, 19(7),* 678-685

Duffy, S., Huttonlocher, J., Hedges, L.V., & Crawford, L.E. (2010). Category effects on stimulus estimation: Shifting and skewed frequency distributions. *Psychonomic Bulletin & Review, 17*, 224-230.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2003). Bayesian data analysis. Boca Raton, FL: Chapman & Hall.

Griffiths, T.L. & Tenenbaum, J.B. (2006). Optimal predictions in everyday cognition. *Psychological Science*. 17(9), 767-773.

Hemmer, P. & Persaud, K. (2014). Interactions between Categorical Knowledge and Episodic Memory across Domains. *Frontiers in Psychology,* 5:584.

Hemmer, P., & Steyvers, M. (2009). A Bayesian Account of Reconstructive Memory. *Topics in Cognitive Science, 1,* 189-202.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98,* 352-376.

Huttenlocher, J., Hedges, L. V., & Prohaska, V. (1992). Hierarchical organization in ordered domains: Estimating the dates of events. *Psychological Review, 95,* 471-484.

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Categories and particulars: Prototype effects in estimating spatial location. *Journal of Experimental Psychology: General, 129,* 220-241.

Kahneman, D., & Klein, G. (2009). Conditions for Intuitive Expertise: A failure to disagree. *American Psychologist, 64,* 515-526.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*, 608-631.

Tenenbaum, J. B. and Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences, 24*, 629-641.