UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
The Influence of Knowledge and Expectations for Color on Episodic Memory

Permalink

Journal

ISSN

Authors
Persaud, Kimele
Hemmer, Pernille

Publication Date
2014

Peer reviewed

The Influence of Knowledge and Expectations for Color on Episodic Memory

Kimele Persaud (Kimele.persaud@rutgers.edu)
Pernille Hemmer (Pernille.hemmer@rutgers.edu)
Department of Psychology, Rutgers University

Abstract

Expectations learned from our environment are known to exert strong influences on episodic memory. Furthermore, people have prior expectations for universal color labels and their associated hue space—a salient property of the environment. In three experiments, we assessed peoples' color naming preferences, and expectation for color. Using a novel experimental paradigm, we then assessed free recall for color. We found that people's color naming preferences were consistent with the universal color terms (Berlin & Kay, 1969), as well as a strong subjective agreement on the hue values associated with these color labels. We further found that free recall for color was biased towards the mean hue value for each preferred color. We modeled this relationship between prior expectation and episodic memory with a rational model under the simple assumption that people combine expectations for color with noisy memory representations. This model provided a strong qualitative fit to the data.

Keywords: Episodic memory; color; prior knowledge and expectations; Bayesian models.

Introduction

"Mere color, unspoiled by meaning, and unallied with definite form, can speak to the soul in a thousand different ways." **— Oscar Wilde**

Color is fundamental to how we identify, define, and organize the world around us. As such, color stands as an essential feature in many facets of society, ranging from the identification of individuals of a racial or ethnic group to the facilitation of simple communication among people. Similarly, colors also serve to represent and unite members of social groups as illustrated by the colors of a country's flag and the colors of sports teams. Color not only holds cultural relevance, but is also an invaluable domain for investigating human cognition.

Color has been employed in efforts to understand cognition ranging from modelling the emergence of language universals as a function of learning biases (Xu, Griffiths, & Dowman, 2010), to measuring the influence of categorization on perception (Webster & Kay, 2012). For example, color naming has been used to understand the perceptual commonalities among different languages and individuals of different cultures (e.g., Davies & Corbett, 1997; Xu, Griffiths, & Dowman, 2010). An important finding from cross cultural studies of the universality of basic color terms is the existence of 11 basic colors: white, black, brown, gray, red, orange, yellow, green, blue, purple, and pink (Berlin & Kay, 1969).

Additional support for the 11 universal color terms comes from studies of perception (Hardin, 2005; Uchikawa & Shinoda, 1996; Webster & Kay, 2012). These terms can be utilized to adequately partition the color space into 11 regions (Uchikawa & Shinoda, 1996), and discretize this space into small sets for categorization and cognitive processing (Webster & Kay, 2012). Taken together, cognitive studies of color naming and color perception suggest that people have clear knowledge and expectations for colors.

Across a broad range of domains, expectations learned from the underlying environment have been shown to influence performance on cognitive tasks, such as perceptual categorization (Huttenlocher, Hedges, and Duncan, 1991; Huttenlocher, Hedges, & Vevea 2000; Jern and Kemp, 2013; Galleguillos and Belongie, 2010), visual perception (Eckstein, Abbey, Pham, & Shimozaki, 2004; Epstein, 2008; Todorovic, 2010), color perception (Mitterer & de Ruiter, 2008), and long term memory (Bartlett, 1932; Hemmer and Steyvers, 2009a). People appear to have strong prior expectations for their natural environment, and use this knowledge optimally. For example, Huttenlocher et al. (1991) showed that having prior knowledge of the underlying stimulus distribution improved average recall. They found that responses regressed toward the mean of the overall stimulus distribution, which enhanced performance.

Furthermore, this influence of prior knowledge has been shown to be hierarchical, such that the structure of the natural environment interacts with recall at multiple levels of abstraction. For instance, recall for the size of objects was shown to be biased towards the overall size distribution, or the distributions associated with specific objects, as a function of familiarity (Hemmer & Steyvers, 2009a; 2009b). Similarly, prior knowledge about the height of people influences recall, not only towards the general height of people, but also at a more fine-grained level based on gender - i.e., knowledge that females on average are shorter than males (Hemmer, Tauber & Steyvers, in revision).

Such behavior is well modeled by a rational model of memory which assumes that noisy data in the mind is optimally combined with prior knowledge about the environment. The question is, how does an observer in a task integrate noisy and incomplete information stored in episodic memory with prior knowledge of the environment? In a real world example, imagine that an individual has witnessed a car accident. Later, when questioned by the police, that individual is asked to recall certain aspects of the event, such as the color of a car seen fleeing the scene. The witness might only have a vague recollection of the events that transpired - the car was greenish. However, they are also likely to have prior knowledge about the possible

colors of cars - green cars tend to be darker green, and are unlikely to be neon green. This knowledge might provide a useful cue when trying to reconstruct the event from memory. In this scenario, and many other real world situations, having prior knowledge and expectations for the regularities of the environment can help fill in vague and uncertain memories and improve average recall.

In this paper, we investigate peoples' prior expectation for color, and the influence of prior knowledge for color on episodic memory. The goals of this investigation are three-fold. First, we seek to determine people's color naming preferences and quantify prior expectation for hue values associated with their preferred color labels. Second, we quantify the contribution of prior knowledge on episodic memory for color. Third, we model this relationship in a simple rational model of memory. We develop a novel experimental approach for assessing free recall for color, where participants generate the hue values associated with color labels using a continuous color wheel, as opposed to recognition of color patches, as has been previously conducted (e.g., Uchikawa & Shinoda, 1996). This paper gives the first characterization of the influence of expectation on free recall for color.

Knowledge and Expectations for Color

In the following two experiments, we determine peoples color naming preferences, and assess their prior expectation for the hue values associated with the preferred color labels.

Experiment 1: Color-Naming Task

The goal of the color naming task was to determine peoples color naming preferences over the hue color space. We predict that color naming preferences will correspond to the universal color terms (Berlin & Kay, 1969). The task required subjects to provide the color label that best represented a given hue value. We will take this as a measure of peoples' preferences for color labels.

Method

Participants Forty-seven Introductory Psychology undergraduate students at Rutgers University participated in this study in exchange for course credit. Participants' ages ranged from 18-23 years of age. All participants provided self-reports of normal color vision. Data from one subject was discarded because no responses were recorded.
Materials The stimuli consisted of 48 colors sampled from the hue color space. Colors varied in hue by 5 units (i.e. hue values of 0, 5, 10, etc) along the full hue range from 0-239, based on the ability to perceptually differentiate two sequential colors in the range. Saturation and luminance were held constant at 100% and 50%, respectively. Stimuli were presented on 23 inch Dell monitors that were all color calibrated using Windows 7 Display Color Calibration. All experiments were written and presented in Matlab.
Procedure A color patch measuring three-by-three inches was presented in the center of the computer screen.

Participants were asked to provide a color label for that specific patch by typing their answer in a response box below the color patch. The patch remained on the screen until the participant was satisfied with their response and clicked continue to view the next patch. Each of the 48 color patches were presented twice in random order, for a total of 96 trials. The experiment was self-paced, and took on average 20 minutes to complete.

Results

Figure 1 shows label frequencies for the 48 color hue values. The top panel shows the 7 most frequent labels (red, orange, yellow, green, blue, purple and pink). Because saturation and luminance were held constant, the presented hue values did not include black, white, brown or gray. The 7 labels comprised 28% of all responses in the experiment. The bottom panel of Figure 1 shows label frequencies for the top 21 labels, comprising 59% of total labels. The cutoff for including the 21 labels was based on a label being given a minimum of 40 times. The results show that participants expressed a large degree of agreement. Furthermore, the 7 preferred color labels coincide with the universal color terms of Berlin & Kay (1969).

Experiment 2: Color Generation Task

The goal of the color generation task was to invert the color naming task, and determine the hue values that people associate with given color labels. In this task, given a specific color label, participants were asked to 'generate' the hue value that best represented that label. We develop a novel experimental approach to allow participants to freely generate color hue responses. We predict a systematic agreement between subjects for the hue values corresponding to the labels centered on the universal color terms. We take this as a measure of peoples' prior expectation for color.
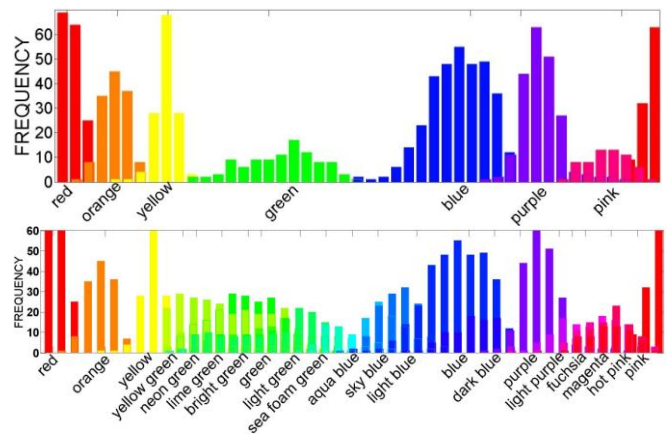

Figure 1. Frequency distributions over color labels in Experiment 1. The top panel illustrates the frequency distributions over the 7 most frequent colors labels. Each bar represents a 5 unit range on the hue scale from 0-239. The colors are presented below the corresponding hue values. The bottom panel illustrates the frequency distributions over the 21 most frequent labels.

Method

Participants Forty-nine undergraduate students at Rutgers University participated for course credit or monetary compensation of $10. All participants provided self-reports of normal color vision. Participants were not involved in Experiment 1.

Materials The stimuli consisted of the 21 most frequent color labels provided in Experiment 1. All stimuli were presented on the same calibrated monitors used in Experiment 1.

Procedure The procedure of Experiment 2 was the inverse of Experiment 1. Participants were presented with a color label and were instructed to generate the color hue corresponding to that label using a color wheel. The label was presented in 24 point Georgia font at the upper right side of the computer screen. To generate a color hue response, participants moved a cursor over a large black circle presented on the left side of the computer screen. The black circle was a mask over a color wheel that varied in hue only. The participant could not see the underlying color wheel - only the black mask[1]. When the participants clicked on the black circle, the corresponding color from that location of the underlying color wheel was shown in a three-by-three inch patch to the right of the wheel and below the color label. The three-by-three inch square was presented in black at the beginning of each trial in order to prevent biased responses. The color wheel, hidden beneath the black mask, was rotated randomly by 45 degrees for each trial so that it was not possible to predict a color's location on the wheel from trial to trial. Participants could click as many times as they wished to generate the color they thought best
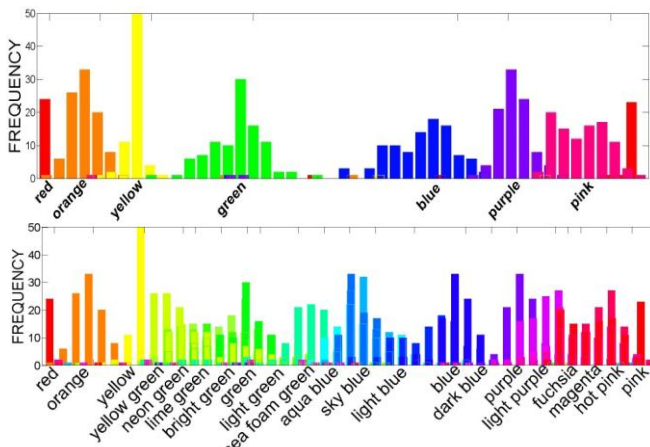

Figure 2 Frequency distributions over hue values from Experiment 2. The top panel shows the frequency at which a particularly hue value was generated on the color wheel for one of the 7 preferred colors. The 7 color labels are presented below their corresponding hue values. The bottom panel shows the hue values generated from the 21 most frequent labels from Experiment 1. Each bar represents a 5 unit range on the hue continuum from 0-239.

---

[1] The color wheel was masked to discourage selection of values only at the edges or directly in the center of each color category. (see Goldstone (1995) for a similar approach).

Table 1. Mean (SD) of hue values and hue ranges for top 7 color labels

	Mean (SD)	Hue Range
Red	1.1 (2.56)	(230–239, 0 – 5)
Orange	20.23(5.59)	(10-30)
Yellow	40.05 (3.04)	(35-50)
Green	79.79 (10.34)	(55-110)
Blue	153.53(12.13)	(115-170)
Purple	189.41 (6.27)	(175-190)
Pink	215.60 (9.57)	(195-225)

corresponded to the given color label. Once participants were satisfied with the color they generated, they pressed the "space bar" to continue to the next trial. Participants generated colors for 21 labels twice each, for a total of 42 trials, presented in random order. The experiment was self-paced, and took on average 30 minutes to complete.

Results

Figure 2 shows frequency distributions over the hue values generated to reflect the given color labels. The color wheel allowed participants to generate colors that differed by 1 unit of hue, resulting in 239 possible hue values. To facilitate comparison between Experiments 1 and 2, responses where binned into the same 48 hue values, as in Experiment 1 (varying by 5 units on the hue range from 0-239, such that all hue values that ranged between 2.5-7.5, where included in the first bin, hue values between 7.5-12.5 fell in the second, and so on). The top panel of Figure 2 shows the hue value frequency distributions for the 7 most frequent labels from Experiment 1 (red, orange, yellow, green, blue, purple and pink). As in Experiment 1, participants expressed a large degree of agreement. Because of the circular nature of the hue space, we fit the frequency distributions with von Mises distributions (a.k.a. the circular normal distribution). Outliers more than 40 hue values from the highest or lowest value in a given colors hue range (see Table 1) were removed before fitting the von Mises distributions, resulting in the removal of 11 responses. The means and standard deviations from the von Mises fits are shown in Table 1. The distributions reflect the notion that a given color label is best represented by a small range of hue values, with some overlap at the edges of the distribution and strongest agreement for the hue value that resulted in the most frequent response of that label in the color naming task. Figure 2, bottom panel shows the hue value frequency distributions for all 21 stimulus labels.

Memory for Color

In the following experiment, we assess free recall for color. We use the hue values from Experiment 1 as the experimental stimuli. Responses were solicited using both the naming task and the color generating task from the two previous experiments. The novelty of this experiment is the

methodology employed to assess free recall of color. We predict that recall will be systematically biased towards the mean of the hue range associated with each preferred color label. For example, for hue values associated with the color label 'red', we predict that darker shades of red (above the mean of the hue range) will be recalled as being lighter, while lighter shades of red (hue values below the mean of the hue range) will be recalled as being darker.

Experiment 3: Color Memory Task

The color memory task was a combination of the two first experimental tasks, in that participants studied a continuous sequence of shapes filled with color selected along the hue range in the same manner as the color patches presented in Experiment 1. At test, participants were then asked to both provide a color label for the color they recalled studying, as well as use the same color wheel as that used in Experiment 2 to generate their reconstruction of the studied color. The goal was to measure the influence of prior expectations for color hue on free recall for color.

Method

Participants Eighteen Introductory Psychology undergraduate students at Rutgers University participated for course credit. All participants provided self-reports of normal color vision. These participants were not involved in Experiments 1 or 2.

Materials The stimuli consisted of 48 shapes uniformly filled with the same 48 hue values used in Experiment 1. The purpose of the shapes was only to cue subjects on test trials to recall the fill-color of the shape. The shapes and colors were paired randomly, and pairings were randomized across subjects. Participants studied each shape and color only once. Stimuli were presented on the same calibrated monitors used in Experiments 1 and 2. See Figure 3 for sample stimuli.

Procedure Participants were shown a continuous study-test sequence of color filled shapes. Shapes were presented one at a time at the center of the computer screen for 2 seconds. They were told to study the color of each shape as they would be asked to recall the color of the shapes. Test trials were randomly interleaved between study trials (see Figure 3 for a sample study test sequence). On a test trial, a shape from a previous study trial, but filled with gray, would appear at the center of the screen. Participants were asked to make three responses: 1) whether or not they recalled studying the shape presented. They responded by clicking either on a yes or a no button presented at the bottom of the
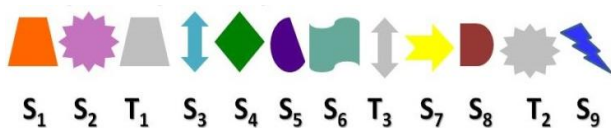

Figure 3. Sample of continuous study/ test sequence. S denotes a study trial with the trial number in subscript. T denotes a test trial for the study trial number in subscript.

screen. 2) the color of the shape at study (this question was posed regardless of their response to the recognition question). Responses were typed into a text box and participants pressed "enter" to continue. 3) to recreate the studied color of the shape using the same color wheel as was used in Experiment 2. Test trials were self-paced.

Results

To measure the influence of prior knowledge, we calculated recall bias as the difference between the recall hue value and the studied value. We restricted the analyzed sample to include only cases in which subjects provided the correct label on the second (color label) question of the test trials (e.g. datum was excluded if the subject recalled blue, when the color studied was red (based on the most frequent label for that hue value in the color naming task), however, responses such as light blue, if the studied color was blue, or yellowish-green if the studied color was green, were acceptable). Hue range for a color category (listed in Table 1) was determined based on the lowest point between two response distributions in the color naming task.

Furthermore, hue responses that deviated by more than 6 standard deviations from the mean of the determined hue range were excluded. In essence this corresponded to someone correctly providing the label 'blue' to a blue hue value, but then reconstructing it as red with the color wheel. This resulted in the removal of 4 data points. Five test trials were also excluded because no response was recorded.

For simplicity and visual clarity, only analysis of the 7 primary labels is presented. Thus, 55% of the data was used in this analysis. The results from this experiment revealed regression to the mean affects as illustrated in Figure 4 top panel. For each of the 7 colors, subjects overestimated values below the mean of each color's hue range and under estimated the values above the mean of each color labels corresponding hue range. A linear regression model was fitted to each subject for each of the 7 preferred colors assuming a single slope and separate intercept for each regression line (see Figure 4 top panel). A one-way analysis of variance revealed a significant main effect of intercepts (F[694]=664, p<.001) across color categories. The negative slope of the lines indicates a regression to the mean effect, such that studied hue values below the mean of that color

Table 2. *Mean slopes and intercepts by color label*

	Slope		Intercept	
	Mean	SD	Mean	SD
Red	-.046	0.13	-3.4137	3.26
Orange	-.046	0.13	10.6451	2.63
Yellow	-.046	0.13	18.2125	3.61
Green	-.046	0.13	37.0389	8.79
Blue	-.046	0.13	64.9861	4.43
Purple	-.046	0.13	88.1715	4.90
Pink	-.046	0.13	92.4914	6.79

Note. N=18

Memory Data and Regression

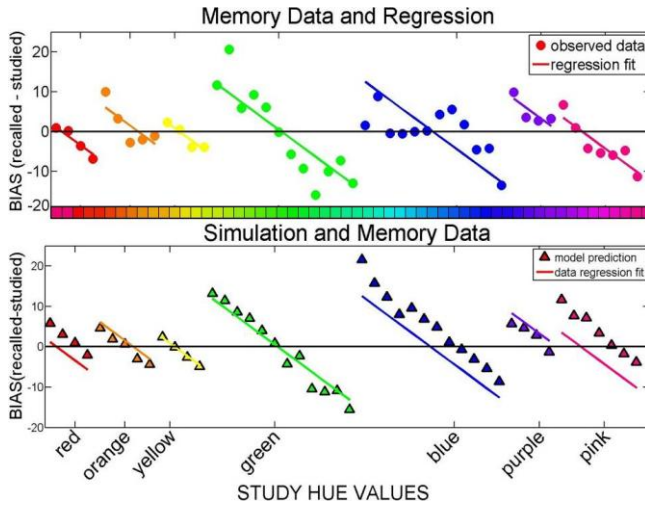**Simulation and Memory Data**

STUDY HUE VALUES

Figure 4. **Top panel**: Recall bias by color category. Positive bias indicates over estimation and negative bias indicates underestimation. The black line indicates no bias. The data points are color coded with the hue for that color range and the corresponding labels are given on the x-axis. The lines give the regression fits for each preferred color label. **Bottom panel**: Model predictions with regression fits from the memory data.

category were overestimated at recall and studied hue values above that color category were underestimated at recall. The different intercepts for each of the color categories indicate regression towards a different mean hue value for each of the color categories. Table 2 shows the slope and intercepts for the 7 color labels.

Modeling

The results show that each of the 7 preferred color labels are associated with pre-experimental prior knowledge of the associated hue range, each exerting an influence on reconstructive memory. That is, hue values less than the mean of the basic color were overestimated and hue values greater than the mean were underestimated.

We propose that this behavior can be modeled with a simple rational model which assumes that prior knowledge for different color categories exert an influence on episodic recall. This rational analysis emphasizes the relationship between behavior and the structure of the environment. For recall, this assumes that the goal of the memory system is to efficiently store and retrieve relevant information, which needs to be combined with prior knowledge and expectations about the environment.

Suppose the observer in our task studies a stimulus feature θ. Based on our experiment, we will assume that the studied features (i.e., hue values) are Gaussian distributed, $\theta \sim N(\mu, \sigma^2)$, with the prior mean μ and variance σ^2 of the features drawn from the environment. When the specific feature θ is studied, we assume that this leads to memory traces y, drawn from a Gaussian distribution centered on the original studied value, and a memory noise process ψ, $y \sim N(\theta, \psi)$. The noise process determines how closely the stored memory trace resembles the original studied stimulus feature. We will further assume the observer has a prior

expectation for the stimulus distribution that mirrors that of the distribution in the environment. At test, the goal of the observer is to recall the studied stimulus feature θ using noisy samples y retrieved from memory and their prior expectation for the distribution of the stimulus. Bayes' rule gives a principled account of how to combine noisy memory representations with prior expectations to calculate the posterior probability,

$$p(\theta|y) \propto p(y|\theta) \ p(\theta) \qquad \text{Eq (1)}$$

The posterior probability $p(\theta|y)$ describes how likely feature values θ are given the noisy memory traces y and prior expectation for the feature $p(\theta)$. Standard Bayesian techniques (Gelman et al., 2003) can be used to compute 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. This rational analysis of recall suggests that the optimal behavior is a trade-off between the strength of the evidence in memory and the likelihood of the event in the natural environment. When our memory representation is strong, recall will closely resemble the studied feature value, but when our prior expectation is strong, and memory content is noisier, recall will more closely reflect the prior expectation.

In this vein, the rational model assumes that the combination of prior expectations and noisy content in memory optimally combine to produce recall of episodic experiences. Furthermore, the model predicts a systematic regression to the mean effect, such that lighter shades (lower hue values) will be recalled to be darker, and darker shades (higher hue values) will be recalled to be lighter.

To implement the model, we specified a prior with mean μ for each color category to be equal to the mean of the von Mises (circular Gaussian) distribution fitted to the frequency distributions in Experiment 2. In other words, we assume these distributions to be representative of peoples prior expectation over hue values for a given color category. In the same way, we set σ^2 for each color category equal to the variances from those same distributions in Experiment 2 (see Table 1). We simulated a memory noise (ψ) that varies for each color category based on the prior standard deviations derived from Experiment 2 (see table 1). We used the model to simulate exactly the same trials that we used in the experiment – including the same sizes for study stimuli.

Figure 4 bottom panel shows the simulated responses from the model. The results show effects of the prior expectation for each preferred color. Lower hue values are estimated to be larger and larger hue values are estimated to be lower, relative to each color category. Overall, the model produces results that are qualitatively similar to the observed data and captures the overall trend in the data. The strength of the current approach is that we make the very simple assumption that peoples prior expectations are drawn directly from the environment. This provides strong support

to the idea that reconstruction from memory is a combination of episodic memory and prior expectations learned from the environment.

Discussion

In this paper we sought to investigate the influence of expectations for color on episodic memory. We measured prior expectation via two tasks: a color naming task which elicited color naming preferences, and a unique task in which participants used a color wheel to generate colors most closely associated with the given color label. The results showed naming preferences that are consistent with the existing literature (Berlin & Kay, 1969), namely red, orange, yellow, green, blue, purple and pink. Subjects also showed a high level of agreement in both Experiments 1 and 2. We then measured the influence of expectation on free recall for color. Results revealed a regression to the mean effect in free recall, such that studied hue values below the mean of that color category were overestimated at recall and studied hue values above that color category were underestimated. This suggests that recall is influenced by expectations for color.

This behavior was modeled with a simple rational model of memory, which assumes that prior knowledge for different color categories exert an influence on episodic recall. In this way, recall is a combination of prior expectations and noisy memory content. The model provides qualitative predictions that are a good fit to the observed data. The model captures the regression to the mean effect for each of the 7 preferred labels. Importantly, the only assumption made in the model was that prior expectations for color were well described by the performance in the color generation task.

Here, we do not provide an analysis of sub-labels (all 21 labels). However, results for hue values within the blue range are interesting in that the pattern of over and underestimation appears to be dispersed. This may be the result of participants separating the hue values in the blue range to account for not just the universal label 'blue', but also high frequency sub-labels (i.e. light blue and sky blue). This suggests that colors might be hierarchically organized, such that blue is the general color label, and sub-labels are based on subjective naming preferences. We believe that this investigation has provided important support for existing understanding of the structures of color categories, as well as a new understanding of relationship between prior expectations and free recall for color.

Acknowledgments

References

Bartlett, F.C. (1932). Remembering: A study in experimental and social psychology, Cambridge, England: Cambridge University Press.

Berlin, B. and P. Kay (1969). Basic Color Terms: Their Universality and Evolution. *Berkeley: University of California Press.*

Davies, I. R. L. & Corbett, G. C. (1997). A cross-cultural study of colour-grouping: Evidence for weak linguistic relativity. *British Journal of Psychology*, 88(3) 493-517.

Eckstein, M. P., Abbey, C. K., Pham, B. T., & Shimozaki, S. S. (2004). Perceptual learning through optimization of attentional weighting: Human versus optimal Bayesian learner. *Journal of Vision,* 4(12), 1006-1019

Epstein, R.A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, 12(10), 388-396.

Galleguillos, C. & Belongie, S. (2010). Context based object categorization: A critical survey. *Computer Vision and Image Understanding,* 114(6), 712-722.

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

Goldstone, R.L. (1995). Effects of categorization on color perception. *Psychological Science, 6*, 298-304.

Hardin, C. (2005). Explaining Basic Color Categories. *Cross-Cultural Research* 39(1), 72-87

Hemmer, P. & Steyvers, M. (2009a). Integrating Episodic Memories and Prior Knowledge at Multiple Levels of Abstraction. Psychonomic Bulletin & Review, 16, 80-87.

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

Hemmer, P., Tauber, S. & Steyvers, M. (in review). Bayesian estimation in rational models.

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

Huttenlocher, J., Hedges, L.V., & Vevea, J.L. (2000). Why Do Categories Affect Stimulus Judgment? *Journal of Experimental Psychology*, (129)220-241

Jern, A. & Kemp, C. (2013). A probabilistic account of exemplar and category generation. *Cognitive Psychology.* 66(1), 85-125.

Mitterer, H. & de Ruiter, J. P. (2008). Recalibrating color categories using world knowledge. *Psychological Science,* (16)629-634.

Todorovic, D. (2010). Context effects in visual perception and their explanations. Review of Psychology, 17, 17-32.

Uchikawa, K. & Shinoda, H. (1996). Influence of basic color categories on color memory discrimination. *Color Research and Application*, 21(6), 430-439

Webster, M.A. & Kay, P. (2012). Color categories and color appearance. Cognition 122, 375-392

Xu, J., Griffiths, T.L., & Dowman, M. (2010). Replicating color term universals through human iterated learning. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society.* Portland, OR: Cognitive Science Society.