University of Alberta

Distortions of Memory for Stimuli

by

Eric Verbeek ©

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Psychology

Edmonton, Alberta
Fall 2006

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-22395-6
Our file Notre référence
ISBN: 978-0-494-22395-6

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Abstract

This thesis investigates two factors that can lead to memory distortion. In Chapter 2, the effects of test stimulus range on generalization gradients in humans were assessed for discriminations between faces that varied in brightness, faces that varied in orientation in the picture plane, and between morphed faces. A bias in responding was found to depend on the difficulty of discrimination in the training task, as well as on the magnitude of bias in the test range. Sufficiently biased testing ranges create shifts in response distributions (generalization gradients), and this may be amplified by using relatively difficult discriminations during training. In Chapter 3, the stimulus displays for a dot location memory task altered the pattern of response bias. In blanks circles, biases have previously been shown to be towards spatial category prototypes. With sectioning radial lines, responses were biased towards the lines, indicating a landmark based strategy, rather than resulting in biases towards alternate category prototypes.

Acknowledgements

I would like to thank my supervisor, Marcia Spetch, for helping me through the difficult steps towards the culmination of this thesis.

The paper that makes up Chapter 2 was accepted for publication in Learning & Behavior, published by Psychonomic Society Publications, on February 3, 2006. That research was supported by an NSERC research grant to Marcia Spetch, and a Queen Elizabeth II Fellowship from the Australian Research Council to Colin Clifford. We are grateful to Joel Pearson for help in generating the face stimuli, and to Alisha Brown, Scott McAnsh and Carlo Dimailig for their help in data collection and experiment preparation.

The paper that makes up Chapter 3 was supported by an NSERC research grant to Marcia Spetch. We appreciate Alisha Brown, Kirsten Williams, Jennifer Harris, and Bowie Lui for their help in data collection. We would also like to offer thanks to Jared Bialowas for inspiring this study, and for help in the programming required for this project, and to Peter Hurd for helping us settle on a more elegant approach to the data analysis.

I dedicate this thesis to my wife, Alisha Brown, because she is such a darling, and puts up with my work late at night. The patience will be returned.

Contents

Tables

Figures

Chapter 1: Introduction

Introduction

Memory is not perfect. People might forget the location of their car keys. Many of us have certainly had the experience of walking into a room with great conviction, only to forget why we had gone to this location. We might also misspell a word that we have spelled correctly many times in the past, or forget someone's name just after learning it, or worse still, call the new acquaintance by the wrong name altogether. Perhaps, as the author had done as a child, one might awake at 10 A.M., realize that he's late for school, and after rushing to get ready and get to school, find the school doors to be locked. It was a Saturday.

These examples all involve a memory problem that involves simple tasks that have been performed correctly in the past. Some of them involve an utter failure of memory, leading to only a feeling of confidence that one does not remember, while others show a more graceful, but more dangerous failure, where one misremembers (Reisberg, 1997, pp.217-218). This thesis is concerned with memory failure of the latter type, in experimental settings. The memory tasks result in memory distortions, with systematic, quantitative changes in memory resulting from the experimental manipulations. These changes in memory and response patterns are elicited by changes in the conditions in which people's responses were collected.

In particular, this thesis investigates the role of stimulus properties as influences on memory. In the first set of studies (Ch. 2), the role of the test stimulus range is investigated. After participants learned to discriminate a target face from another that looked fairly similar to it, their task was to discriminate the target face from an entire series of faces that varied in their similarity to the target face. The results indicate a distortion in memory for the target face, but only under particular experimental parameters. If the test range is not centered relative to the training stimulus values, the response distribution resulting from generalization testing can be biased in the same direction as the test range (Thomas, 1993). This is called a range effect, and theoretically stems from a shift in the adaptation level. The adaptation level is theorized to be an average stimulus value; participants experience stimuli, and over time develop an adaptation level which serves as a dynamic reference point by which they encode other stimuli. Often, the range of stimuli experienced in different parts of an experiment will change, thus leading to a different adaptation level in different phases. If the adaptation level changes between training (when they encode stimuli) and testing (when they decode what they have learned in relation to the adaptation level), then there can be a response shift exactly because this reference point has moved. In Chapter 2, we investigate this effect in face recognition, and demonstrate that the extent of test range bias and the training difficulty are both determinants of the range effects. In particular, there seems to be an optimal training difficulty that facilitates range effects. For faces, this appears to be a harder, rather than easier training discrimination task, which runs counter to some previous findings with other kinds of stimuli (Thomas, Svinicki & Vogt, 1973; Thomas, Mood, Morrison & Wiertelak, 1991).

In the second set of studies (Ch. 3), the effects of the stimulus display field on spatial memory was tested. By drawing different radial lines within a circular field (like slices of a pie), participant's memory for the position of a recently seen dot was distorted towards the lines, as a function of the dot's actual position. Previous research in this paradigm has demonstrated that participants normally respond with biases towards implicit category prototypes (Huttenlocher, Hedges and Duncan, 1991). People implicitly divide the circle into four quadrants, which are theoretically spatial categories. They then bias their responses towards the center of the quadrant that the dot appeared in, towards the category prototype. Huttenlocher at al. theorized that very distinct category boundaries were helpful, and one might then expect that the radial lines that we used would allow participants to produce new categories. In our experiments, participants biased their responses towards radial lines rather than towards the centers of categories defined by the lines. When the lines were removed from only the response circle, participants did not show a strong tendency to follow either the categorical or landmark response strategy. When participants practiced responding using the line, and the lines were subsequently removed from the stimulus and response circles, participants returned to the categorization strategy. It thus appears that when visual aides are available, people make use of an alternate response strategy for remembering dot locations, which results in a different pattern of response bias.

The context in which we are expected to recall sensory stimuli has a subtle effect on our recall accuracy. These two series of experiments systematically investigate two ways in which memory for a stimulus can be distorted, and will give us a better understanding of cognitive mechanisms of memory, by highlighting reliable quirks of remembering.

References

Huttenlocher, J., Hedges, L.V. and Duncan, S. (1991). Categories and Particulars: Prototype Effects in Estimating Spatial Location. Psychological Review, 98(3), 352-376.

Reisberg, D. (1997). Cognition, Exploring the Science of the Mind. New York, NY: W. W. Norton & Company, Inc.

Thomas, D.R. (1993). A model for adaptation-level effects on stimulus generalization. Psychological Review, 100(4), 658-673.

Thomas, D.R., Mood, K., Morrison, S. and Wiertelak, E. (1991). Peak shift revisited: A test of alternative interpretations. Journal of Experimental Psychology: Animal Behavior Processes, 17(2), 130-140.

Thomas, D.R., Svinicki, M.D., and Vogt, J. (1973). Adaptation level as a factor in human discrimination learning and stimulus generalization. Journal of Experimental Psychology, 97(2), 210-219.

Chapter 2: Determinants of Range Effects in Face Recognition

Determinants of Range Effects in Face Recognition

Eric L. Verbeek
University of Alberta
Marcia L. Spetch
University of Alberta
Ken Cheng
Macquarie University
Colin W.G. Clifford
University of Sydney

Accepted for publication by Psychonomic Society Publishing in Learning & Behavior,
February 6, 2006.

Determinants of Range Effects in Face Recognition

Responding to a stimulus as though it were another previously experienced stimulus is referred to as generalization, and is a major feature of behavior. The likelihood of making a false positive identification or a generalized response to a stimulus is a function of its similarity to the target stimulus, and is has been theoretically described as either a Gaussian or an exponential distribution (Ghirlanda & Enquist; 2003; Shepard, 1987). This distribution is normally centered on the target stimulus, but particular manipulations may reliably shift the distribution. For example, a participant may be trained to discriminate between the target stimulus (S+) and a foil (S–) that differs along a particular stimulus dimension (brightness, vertical orientation, volume). The participant is reinforced for responding to the S+, and not reinforced for responding to the S–. When good discrimination develops, a larger range of stimuli along the manipulated dimension is presented. A common finding is that the distribution of responding is systematically shifted away from the S+. For example, many studies have found that the peak of responding is to a stimulus beyond the S+, in the opposite direction of the S–; this is referred to as the peak shift effect (Hanson, 1959; Spence, 1937). This is interesting because the maximum of responding occurs to a stimulus that is different from the S+. The closer together the S+ and S– are along the stimulus dimension, the greater the peak shift tends to be during testing (Thomas, 1962). Sometimes, the peak of responding does not shift away from the S+, but participants nevertheless respond more to stimuli on the S+ side of the distribution than to stimuli an equivalent distance away from the S+ on the S– side. This is referred to as area shift (Nallan, McCoy, Pace, & Welch, 1979). Peak or area shifts occur with intra-dimensional testing wherein test stimuli vary along the dimension in which the discrimination training stimuli differed; these effects do not typically occur with inter-dimensional testing, which varies the training stimuli in one dimension and the test stimuli in another (Ghirlanda & Enquist, 2003).

Two theories have been proposed that can account for peak shift: Thomas' (1993) adaptation level hypothesis, which is often used to explain peak shift in humans, and Spence's (1937) conditioning account, which is often used to explain the effect in non-human animals. Spence's account claims that an excitatory gradient forms around the S+ and an inhibitory gradient forms around the S–. His theory (which was proposed before peak shift had been demonstrated) predicted that if the S+ and S– were sufficiently close, the combination of the two gradients would result in a net excitatory gradient that is shifted away from the S+ in the direction opposite to the S– (Spence, 1937). Thomas offered a very different account of peak shift based on Helson's (1947) notion of adaptation level. Thomas's account holds that responding is based upon the relationship between the positive stimulus and an adaptation level that develops during training (Thomas, 1993). This adaptation level is thought to reflect the subjective average of the stimulus values experienced during training. When testing proceeds, the adaptation level changes due to exposure to a new set of stimuli, but the learned relation between the adaptation level and the positive stimulus is still used to determine responses. If the new adaptation level differs from that in training, the peak of responding will shift in the same direction, and to about the same degree that the adaptation level itself has changed. This hypothesis thus predicts that an asymmetrical change in test range relative to training range will be a critical determinant of the resulting response gradient. Several experiments with humans, using stimulus dimensions such as weight, line orientation, brightness and color have confirmed that manipulations of test stimulus range can shift the peak of responding in a manner consistent with the adaptation level hypothesis (MacKinnon, 1972; Thomas, Mood, Morrison & Wiertelak, 1991; Thomas & Jones, 1962). Particularly powerful evidence for the adaptation-level hypothesis comes from demonstrations that, with the appropriate test range manipulation, the response distribution can shift toward the S– side of the distribution, a result that is inconsistent with Spence's theory.

Lewis and Johnson (1999) demonstrated a peak shift effect in humans' face recognition following training in which human participants discriminated between two morphed faces. The S+ (target) and S– (non-target) faces were selected from a series of blended faces. Blended faces vary in numerous ways including changes to local features and global facial configuration, and hence morphed faces can be said to differ along a complex dimension. During testing, the participants were presented with additional stimuli that fell between the S– and S+ and on the S+ side of the distribution. The peak of responding was shifted away from the S+ in the direction opposite to the S–. The range of stimuli presented during testing was not manipulated, so that the contribution of adaptation level to this peak shift could not be assessed. These results spurred Spetch, Cheng, and Clifford (2004) to investigate whether range effects could be found with human face stimuli. In their first experiment, participants were trained to discriminate between a pair of faces and to press a "Yes" button for the specified target face (S+), and a "No" button for the other face (S–). The face stimuli were derived from a series of photographs, produced through morphing, that ranged from a photo of a unique face to an averaged photo. Following discrimination training, participants were given either blocked tests or probe tests in which they were presented with a range of stimuli biased toward either the S+ or the S– end of the series. For either method of testing, their results showed an area shift (more responding on the S+ side than on the S– side), but no range effects were found. They conducted additional experiments with modified designs but still failed to demonstrate range effects on face discrimination. Moreover, they demonstrated that the procedure was not at fault for the failure to see a range effect; with a procedure used by Thomas et al. (1991), Spetch et al. found range effects with line tilt stimuli but not with face stimuli.

Spetch et al. (2004) suggested that "the more complex and multidimensional the stimuli, the more likely area shift rather than range effects will be found" (p. 239). In Experiment 1 we tested whether face stimuli, as complex (multi-elemental) stimuli, are resistant to range effects even when they vary in only a single dimension. Experiment 1a replicated Spetch et al.'s (2004) failure to find adaptation effects in discrimination of average to unique morphed faces, whereas Experiments 1b and 1c, respectively, assessed whether range effects would occur if the face stimuli varied along the brightness dimension (an example of manipulation along an intensity dimension) or in their orientation (an example of an arrangement dimension; Ghirlanda & Enquist, 2003). In Experiments 2 and 3 we assessed the role of instruction, discrimination difficulty, and width of the training and testing ranges on the occurrence of range effects with face stimuli. Because previous studies have found the classic peak-shift effect with faces, our investigations focused exclusively on range effects. Our studies were designed to determine whether there are any conditions under which range effects could be found with complex stimuli.

Experiment 1a

In this experiment, we used a series of morphed faces that ranged from an average face to a unique face. We tested for adaptation effects after varying the range of faces experienced during testing. The test range for one group was biased in the region beyond the positive training stimulus, and the other group was biased toward the negative stimulus end. Six faces that were common to both ranges allowed us to assess whether judgements of these faces were influenced by the range experienced. Specifically, a shift in responding toward the range experienced in testing would provide evidence for an adaptation level effect.

Methods

Participants

The participants in all of the experiments were University of Alberta undergraduate students enrolled in introductory psychology courses. They received course credit for their participation. Forty participants were included in Experiment 1a (20 female and 20 male).

Design

Participants were randomly assigned in blocks to four conditions (N = 10 per condition). Half of the participants in each condition viewed faces from the female stimulus set and the rest viewed faces from the male stimulus set. The experimental factors consisted of: position of the positive stimulus within the stimulus series (toward the unique end or the average end of the series); and test stimuli range (biased toward the side of the positive stimulus or the negative stimulus). The 10 test stimuli presented to each participant formed a within-subject factor. For statistical analyses, only the stimuli common to both range conditions (the middle six stimuli) were included. Thus, the design used for statistical purposes was a mixed design, with range, positive stimulus, and face gender as between-subject factors (each with two levels) and test stimulus as a within-subject factor (with six levels).

Stimuli

The stimuli were drawn from those used in Spetch et al. (2004). Briefly, these stimuli had been created by photographing 20 people's faces under controlled conditions and then averaging them into a single photograph using techniques specified in the original article (p. 225). One original photograph and the average photograph were then combined through weighted averaging to create 41 stimuli between them, resulting in a morphed series. Weighting began at 0% average and 100% unique, and progressed in increments of 2.5%.

For the present study, one set of each gender was used. Of the 41 stimuli in each set, every second stimulus, beginning at the 8th stimulus and ending at the 34th, was selected. This resulted in fourteen stimuli per set, each 5% apart, that ranged from 82.5% unique to 17.5% unique. The pictures were grayscale bitmaps with on-screen dimensions of 4.3 cm by 5.0 cm (distance between upper jaw tips, by top of forehead to tip of chin). The monitors were 17" CRT monitors with a 5:4 aspect ratio. Display resolution was 600 x 480 pixels, and stimuli were displayed on a black background. Seated participants viewed the screen from a distance of 30 to 60 cm.

Absolute Stimulus Steps

Experiment 1c (degrees from vertical picture plane orientation)

Lower Range	12	17	22	27	32	**37**	42	47	**52**	57				
Upper Range					32	**37**	42	47	**52**	57	62	67	72	77

Experiment 1b (relative brightness difference in CorelDraw brightness units)

Lower Range	-14	-12	-10	-8	-6	**-4**	-2	0	**2**	4				
Upper Range					-6	**-4**	-2	0	**2**	4	6	8	10	12

Experiment 2b (relative brightness difference in CorelDraw brightness units)

Lower Range	-6	-5	-4	-3	-2	**-1**	0	1	**2**	3				
Upper Range					-2	**-1**	0	1	**2**	3	4	5	6	7

Experiment 1a (% morph between unique and average face)

Lower Range	17.5	22.5	27.5	32.5	37.5	**42.5**	47.5	52.5	**57.5**	62.5				
Upper Range					37.5	**42.5**	47.5	52.5	**57.5**	62.5	67.5	72.5	77.5	82.5

Experiment 2a (% morph between unique and average face)

Lower Range	2.5	10.0	17.5	25.0	32.5	**40.0**	47.5	55.0	**62.5**	70.0				
Upper Range					32.5	**40.0**	47.5	55.0	**62.5**	70.0	77.5	85.0	92.5	100

Experiment 3, Testing Width Compressed, Training Width Narrow (% morph between two unique faces)

Lower Range	23	26	29	32	36	**41**	.	50	55	**59**	64	68			
Upper Range				32	36	**41**	45	50	.	**59**	64	68	72	76	80

Experiment 3, Testing Width Compressed, Training Width Wide (% morph between two unique faces)

Lower Range	23	26	29	32	**36**	41	.	50	55	59	**64**	68			
Upper Range				32	**36**	41	45	50	.	59	**64**	68	72	76	80

Experiment 3, Testing Width Extended, Training Width Narrow (% morph between two unique faces)

Lower Range	0	8	16	24	32	36	**41**	50	**59**	64	68				
Upper Range					32	36	**41**	50	**59**	64	68	76	84	92	100

Experiment 3, Testing Width Extended, Training Width Wide (% morph between two unique faces)

Lower Range	0	8	16	24	32	**36**	41	50	59	**64**	68				
Upper Range					32	**36**	41	50	59	**64**	68	76	84	92	100

Table 2-1. Diagram showing the design used in all three experiments. Bolded values acted as training values (S+ and S- counterbalanced). Those highlighted in grey were in the common testing range, and were used for analysis. "Lower Range" and "Upper Range" correspond to the range conditions (positive and negative range) depending on the value of the positive stimulus for a given subject.

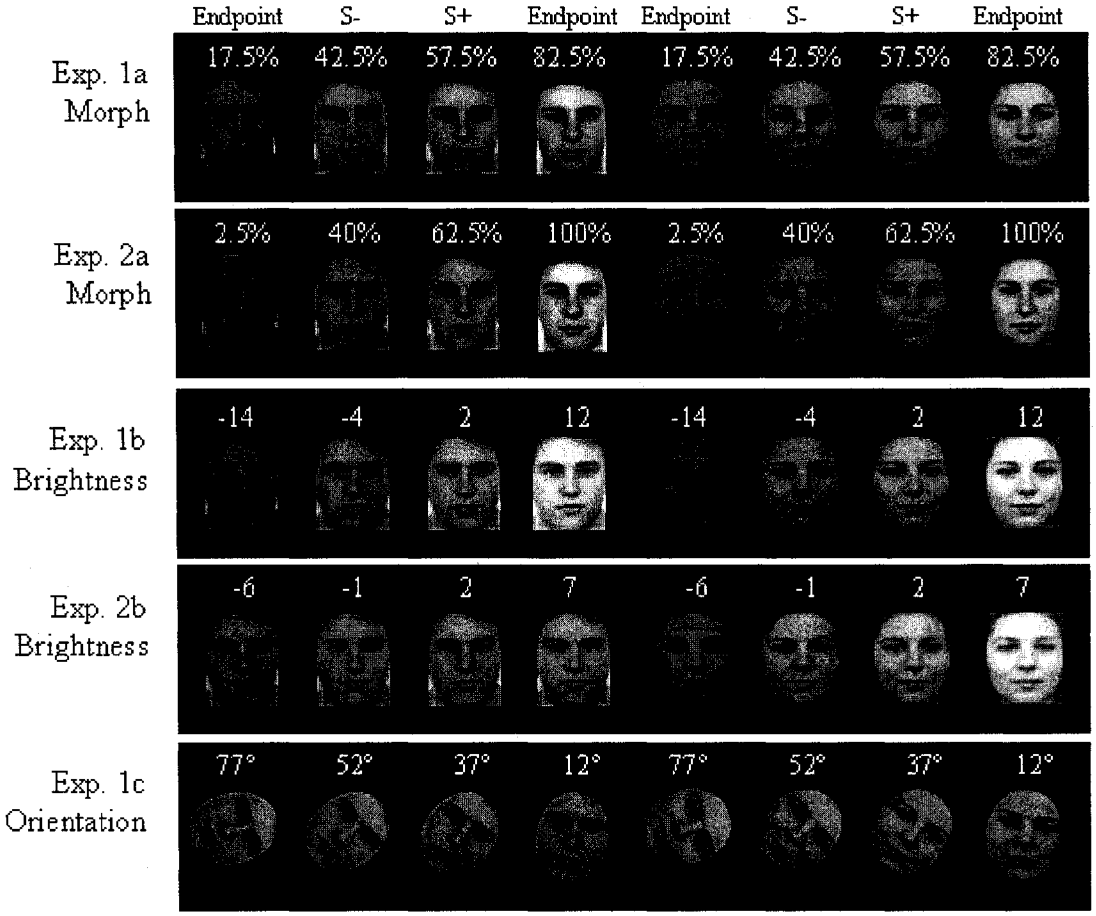


Figure 2-1. Sample face stimuli used in Experiments 1 and 2. Rows correspond to experiments (morph is 'a', brightness is 'b', and orientation is 'c'), and stimulus identity is designated by column, as labelled. Only one endpoint would be seen per subject, as described in the experimental designs. Stimulus values appear above each face. Positive and negative stimulus values were counterbalanced across subjects.

Procedure

The stimuli were presented on 17-inch CRT monitors, and auditory feedback during training was presented through headphones. Participants were told they would learn to discriminate between a pair of similar-looking photographs of faces. On-screen instructions informed participants that the face would appear for only 1.5 seconds, that the first face would be the target face, and that they should press the space bar whenever they see the target face and refrain from doing so for all others.

When training commenced, participants received auditory feedback after each trial indicating whether responses were correct or wrong. The face stimuli were presented for a maximum of 1.5 seconds or until the space bar was pressed. The inter-stimulus interval was 3 sec, during which time feedback was given, and the screen was completely black.

During training, participants learned to discriminate between the 42.5% and 57.5% morphed faces (6th and 9th stimuli of the 14 in testing; see Table 2-1). Depending on the participant's condition, one of these served as the target (positive) stimulus and the other served as the negative stimulus.

The training phase included at least two blocks of eight stimulus presentations, the first of which had the same order of positive (+) and negative (–) stimuli across all participants: +, –, +, –, +, +, –, –. The second block presented four + and four – stimuli in a random order. Testing commenced if participants made six or more correct responses within the second block. If fewer than six responses were correct, additional 8-trial blocks of randomly ordered trials were presented until six or more correct trials occurred within a block.

At the onset of testing, on-screen instructions informed participants that they were to continue the same task of recognizing the target face, except that additional face stimuli would be presented, and there would no longer be auditory feedback on performance. During testing, each participant was presented with 10 stimuli, which were each presented 8 times in random order. For all participants, these included both training stimuli (6th and 9th positions), two stimuli between the training stimuli (7th and 8th positions), and one outside of each training stimulus (5th and 10th positions). Participants in the positive range condition received four additional stimuli from beyond the positive end of their training range, whereas participants in the negative range condition received four stimuli from beyond the negative training stimulus. Common test stimuli thus ranged between 37.5% and 62.5%, and the entire range from which test stimuli were selected was between 17.5% and 82.5% morphed. See Table 2-1 for further details, and Figure 2-1 for examples.

Analysis

Responses to test stimuli were recoded based on their serial relation to the positive stimulus used in training. The dependent measure was the proportion of positive responses (trials on which the space bar was pressed) to each test stimulus. To assess whether the distribution of responses shifted as a result of our range manipulation, a weighted peak statistic was calculated for each subject using only the six stimuli common to both range groups (essentially a weighted mean for frequencies of interval data; see Thomas 1962, and Hays 1994, p.173). The peak was calculated by multiplying the morph percentage of each test stimulus within the common range by the proportion of responses given to that stimulus, and dividing by the sum of all six proportions.

A univariate ANOVA was performed on this weighted peak statistic, with range, positive stimulus, and face gender as factors, each with two levels. An adaptation effect would be expected to shift the response gradient in the direction of the padded range. Therefore, a significant range effect in which the peak is smaller for the negative range group than for the positive range group would provide evidence of an adaptation effect.

Predicted peaks were found by first getting the weighted average of the test stimuli, to find the final adaptation level, and then adding the distance between S+ and the training AL. This method is useful for determining ordinal predictions to compare conditions but is not intended to provide accurate quantitative predictions, because we do not have subjective scales for our stimuli (see Thomas 1993). In addition, the predicted values are very likely to overestimate range effects since the adaptation levels will likely not have

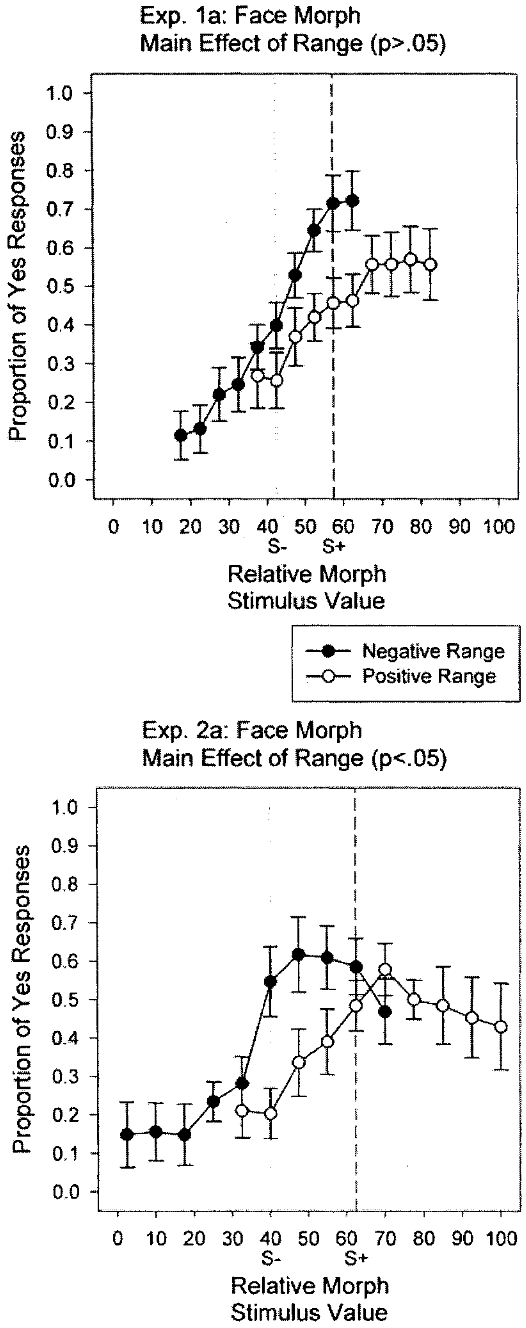


Figure 2-2. Proportion of "Yes" responses to test stimuli in morphed face experiments, broken down by Stimulus Range condition (M ± SE). Face Gender and Positive Stimulus conditions have each been collapsed across. Only stimuli common to both ranges were used in analysis.

fully shifted by the end of testing, and the predicted values refer to the location of peak responding rather than the weighted peak statistic used in the analysis.

Results and Discussion

The ANOVA on the weighted mean position scores (positive range = 53.3% and negative range = 51.8%) did not reveal a significant range effect ($F_{(1,32)}=0.452$, $p>.05$). Predicted values were 67.5% and 47.5% respectively. No other effects in the model were significant. Thus, this experiment failed to reveal evidence of an adaptation level effect, and therefore replicates the results found by Spetch et al. (2004). In Figure 2-2, one can see that the six common values are virtually parallel for the two range groups, and the greater overall responding in the negative group is not indicative of a range effect.

Experiment 1b

The lack of range effect with morphed faces is consistent with previous results (Spetch et al., 2004) and leads to a consideration of what might block the effect. The morphed dimension is complex, in that the stimuli change in configuration, at both an elemental level and in their entirety. In the present experiment, we varied the brightness of the face images to produce a stimulus series that instead varies along a simple intensity dimension (Ghirlanda & Enquist, 2003). If the absence of an adaptation effect was due to the complexity of the changes in morphed faces, then adaptation effects might occur if the faces are varied along one simple dimension, such as brightness. The effect should fail to occur, however, if faces themselves are somehow resistant to adaptation effects, perhaps due to stimulus complexity, as hypothesized by Spetch, Cheng and Clifford (2004, p.239).

Participants

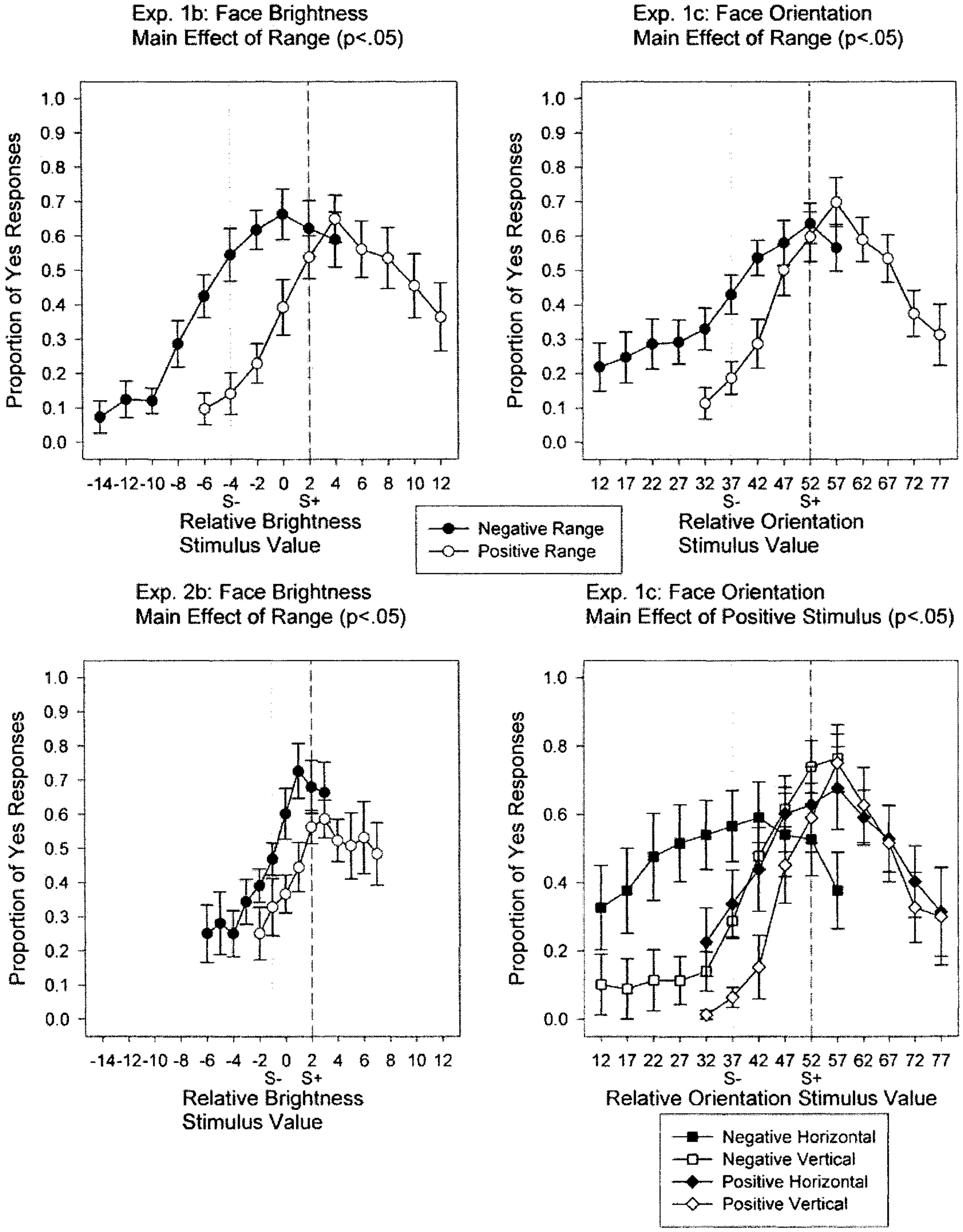
Forty undergraduate University of Alberta students were included in Experiment 2 (23 female, 17 male).

Design

The design of this experiment was identical to that of Experiment 1a except that the stimulus series consisted of faces that varied along the brightness dimension.

Stimuli

From the forty-one photographs available from the original morph set, one was selected from each gender (the 18th face stimulus of the originals). This original was edited with CorelDraw to create fourteen faces that ranged in steps of 2 from +12 to –14 brightness units away from the original. The face at +12 units was the lightest face, and the –14 face was the darkest. The positive and negative training values were –4 and 2; the common testing range was from –6 to 4.



Exp. 1b: Face Brightness
Main Effect of Range (p<.05)

Exp. 1c: Face Orientation
Main Effect of Range (p<.05)

Exp. 2b: Face Brightness
Main Effect of Range (p<.05)

Exp. 1c: Face Orientation
Main Effect of Positive Stimulus (p<.05)

Figure 2-3. Proportion of "Yes" responses to test stimuli in brightness (1b and 2b) and orientation (1c) experiments, broken down by Stimulus Range condition (M ± SE). Face Gender and Positive Stimulus conditions have each been collapsed across in the first three graphs. The fourth plot collapses across Stimulus Range, showing gradients for Positive Stimulus conditions. Only stimuli common to both ranges were used in analysis.

Procedure

The procedure was the same as in Experiment 1, except that the instructions informed participants that the pictures would vary in brightness, and that they should pay attention to brightness in order to respond correctly.

Results and Discussion

As predicted by the adaptation-level hypothesis, participants in the positive range condition had a higher mean response position score (1.45) than did participants in the negative range condition (–1.0; $F_{(1,32)}=19.269$, $p<0.05$). Predicted values were 6 and -2 respectively. No other effects were significant. In Figure 2-3, one can see a distinct shift in the two gradients, which otherwise appear to have an identical overall shape. Each group thus demonstrated a shift in responding toward the side that was padded with more test stimuli, as expected if the adaptation level shifted during testing and responding continued to be based upon the relation between the training adaptation level and the target stimulus. This result is very interesting because, despite an identical procedure, Experiment 1a did not find this result. This implicates the complex changes occurring in morphed faces in the lack of a range effect.

Experiment 1c

In this experiment, faces were varied in their orientation. Orientation is an example of an arrangement dimension (Ghirlanda & Enquist, 2003) but, like brightness, it is a much simpler dimension than the morphed faces. Because the faces differed only in this one dimension, we expected that the results would be similar to those found in Experiment 1b for variations in brightness.

Participants

Forty undergraduate University of Alberta students were included in Experiment 1c (21 female, 19 male).

Design

This design was the same as in the previous pair of experiments except that the stimulus series consisted of faces that varied along the orientation dimension.

Stimuli

The faces were the same as those used in Experiments 1a and 1b, but they were cropped to fit within an ellipse and then a stimulus series was created by manipulating the orientation of the faces in the picture plane using CorelDraw. The series ranged from 12 degrees (i.e., close to vertical) to 77 degrees (i.e., close to horizontal) in five-degree increments. This places the positive and negative training stimuli at 37 degrees and 52 degrees, centered on 44.5 degrees, and the common test range was between 32 and 57 degrees.

Procedure

This procedure was identical to the previous experiments except that participants were told that the photographs would vary in orientation and that they were to identify the target stimulus based upon this property. Participants were also instructed to not touch the screen in any way. This prevented participants from using a mark on the screen or their finger as a cue. An experimenter was in the room at all times and no transgression of this rule was observed.

Results and Discussion

The mean response position was higher in the positive range condition (50.7 degrees) than the negative range condition (45.3 degrees) which indicates an adaptation effect ($F_{(1,32)}=16.112$, $p<0.05$). Predicted values were 62 degrees and 42 degrees respectively. The plot in Figure 2-3 shows the two range gradients, and it is clear that the response gradients differed depending on which side was padded with more stimuli, as predicted by the adaptation level hypothesis.

There was also a significant effect of which stimulus was positive during training (more vertical or horizontal) on the weighted peak location (vertical condition at 50.0 degrees, horizontal condition at 45.9 degrees; $F_{(1,32)} = 9.29$, $p<.05$). Recall that for analysis and plots, stimulus values were recoded relative to the positive stimulus value during testing, to place expected peaks for each group in the same region. This effect of positive orientation can be seen in Figure 2-3, which shows the gradients of four groups, composed by crossing the range and positive stimulus conditions. Discrimination between stimuli was better (i.e., a higher peak and a steeper slope away from this peak) in the groups for which the positive stimulus was closer to vertical than in the group for which the positive stimulus was closer to horizontal. Relative to the vertical positive groups, the curves for the horizontal positive groups (particularly the negative range subset), are flatter and shifted toward the negative side. One plausible explanation for this result is that humans might be more sensitive to changes in the vertical orientation of faces because of greater experience in viewing vertically oriented faces than horizontally orientated faces. Indeed, Collishaw and Hole (2001) found that when faces were blurred to disrupt their featural information and then were subsequently tilted away from vertical, recognition became poorer. This suggests that the configural processing is more difficult for faces that are tilted away from vertical, which might cause participants in the horizontal-positive condition to be less accurate in discriminating the target orientation from the other orientations, if configural cues are used to perceive orientation. Extensive prior experience with vertically oriented faces might set up a strong pre-experimental adaptation level of vertically oriented faces, which is not completely overcome by a brief experimental experience. A near vertical S+ works in concert with the pre-experimental adaptation level, whereas a near horizontal S+ works against the pre-experimental adaptation level.

Experiment 2

The results of Experiments 1b and 1c clearly indicate that faces per se are not resistant to range effects. Robust range effects were found when people discriminated between faces that varied along the single dimensions of brightness and orientation. At the same time, we replicated the findings of Spetch et al. (2004) in failing to find a range effect when people discriminated among faces that varied along the complex dimension created by morphing faces. However, in addition to the complexity of the dimensions, there were two other potentially important differences between the morph manipulation and the manipulations of brightness and orientation. First, the instructions differed because participants were alerted to the dimension varied (different-looking faces in Experiment 1a, photographs varying in brightness or orientation in Experiments 1b and 1c, respectively). Second, the morph discrimination may have been more difficult than the brightness and orientation discriminations. Indeed, the average number of training trials required to meet the criterion to move to testing was higher in Exp 1a (50.0) than in Exp 1b (37.1) or Exp 1c (23.0). Therefore, Experiment 2 replicated the conditions of Experiments 1a and 1b, but with two changes. First, we used the identical instructions for both the morph and the brightness discrimination. Second, we made the morph discrimination easier by selecting values during both training and testing that were spaced further apart, and the brightness discrimination harder by spacing stimuli closer together.

Experiment 2a

This experiment is a replication of Experiment 1a, with training stimulus discriminability increased, and with greater distance between adjacent test stimuli, and thus a greater distance between the test series endpoints.

Methods

Participants

Thirty-two undergraduate University of Alberta students were included in Experiment 2a (22 female, 10 male).

Design

This design was the same as in the previous experiments. The stimulus series consisted of morphed faces selected from the same large stimulus set as in Experiment 1a.

Stimuli

The total range of the 14 stimuli was between 2.5% and 100% morphed from unique to average, with 7.5% steps between stimuli. Training stimuli were 40% and 62.5%, and the common test range was 32.5% to 70%. The comparison between Experiment 1a and this experiment can be seen in Table 2-1.

Procedure

Procedures were the same as in Experiment 1a, including the instructions, which specified that participants would need to discriminate between faces that looked very similar.

Results and Discussion

The stimulus changes did indeed allow for a range effect; the mean response position was 51.3% for the negative range condition and 58.9% for the positive range condition ($F_{(1,32)}=10.23$, $p<0.05$). Predicted values were 47.5% and 77.5% respectively. Results are plotted in Figure 2-2, wherein one can see a distinctly different center for each response gradient. The effect contrasts with Experiment 1a, in which there was a greater overall rate of responding in the negative group, but no obvious difference in gradient location. The average number of training trials to criterion was 42.75, as compared to 50.2 for Experiment 1a.

Experiment 2b

This experiment is a replication of Experiment 1b, with training stimulus discriminability decreased, and with a test stimulus range that had less distance between adjacent stimuli. Instructions were also made identical to those in Experiment 1a and 2a.

Methods

Participants

Thirty-two undergraduate University of Alberta students were included in Experiment 2b (23 female, 9 male).

Design

This design was the same as in the previous experiments. The stimulus series consisted of brightness modified faces produced in the same manner as described in Experiment 1b.

Stimuli

The 14 stimuli ranged between –6 and 7 brightness units, in steps of 1 unit. Training stimuli were –1 and 2, and thus the common test range was from –2 to 3. Comparisons between Experiment 1b and 2b may be made by using Table 2-1.

Procedure

The only change in procedures was the neutralization of the instructions, which now told participants that they need to discriminate between faces that look quite similar, rather than between faces that differ in their brightness (as in Experiment 1b).

Results and Discussion

Despite increasing training difficulty (mean 48.0 trials to criterion compared to 31.2 in Experiment 1b) and decreasing test range, a range effect still developed; the negative range peak (0.63) and positive range peak (1.27) were significantly different ($F_{(1,32)}=5.39$, $p<.05$). Predicted values were 0 and 4 respectively. Results are plotted in Figure 2-3. Thus, range effects with faces that vary in the brightness dimension were robust across variations in instruction and in the extent of the training and test range.

Experiment 3

The contrast between the results of Experiment 1a and Experiment 2a suggests that either the training discriminability or the extent of the test range may be important factors in the occurrence of range effects with morphed faces. Previous studies suggest that both of these factors can influence the magnitude of range or peak shift effects with other stimulus dimensions (Baron, 1973; Thomas, Mood, Morrison, Wiertelak, 1991). Experiment 3 was designed to determine the impact of both training discriminability and the magnitude of range manipulation on range effects with morphed faces.

Methods

Participants

Ninety-six undergraduate University of Alberta students participated in Experiment 3; data for 85 were kept (53 female, 32 male), and 11 subjects failed to pass training (unrecorded gender). Five were omitted from the statistical analysis because they failed to respond to the common stimulus range, thus making it impossible to calculate a peak within that range. These subjects did respond to stimuli outside that range, and are thus included in Figure 2-5.

Design

Building from the previous designs, this experiment includes two new between-subject variables: training width, and testing width. Training width was the distance between the S+ and S–, and was either wide (36% and 64%) or narrow (41% and 59%). Testing width corresponded to the total test stimulus range, and was either compressed (23% to 80%) or extended (0% to 100%). The test stimulus intervals were uneven within a series. This allowed us to produce particular differences between the training and testing adaptation level: the difference was approximately 7% of morph for the compressed range, and 14% of morph for the extended range. Table 2-1 details the exact setup of stimulus values within conditions. The complete set of factors includes testing width (compressed or extended), training width (narrow or wide), range (positive or negative), positive stimulus (above or below 50% morph stimulus), and face gender (male or female).

Stimuli

New face stimuli were used for this experiment, both to generalize to a new set of faces, and to increase the difference between the two ends of the series. Whereas the previous stimuli were morphed between a single unique face and an averaged face, the faces for Experiment 3 were morphed between two very different looking unique faces. The original faces were acquired from the Max-Planck Institute for Biological Cybernetics (Department Bülthoff URL: http://www.kyb.mpg.de/bu/index.html; Face Database URL: http://faces.kyb.tuebingen.mpg.de/). These faces were produced from 3D scans of people's heads; scans were then strategically morphed together to eliminate individual identity from the face database. We selected particular face pairs from the 200 available front-facing faces, and manipulated them with the same morphing procedure used by Spetch, Cheng and Clifford (2004). One female and one male set were used. New 17"monitors were used for this study. The monitor aspect ratio was 4:3; the resolution was set to 1024 x 768 pixels, and the faces were approximately 4 cm by 6 cm (distance between upper jaw tips, by crown top to chin tip). The faces were presented in color rather than grayscale as in the previous studies. Sample stimuli appear in Figure 2-4.

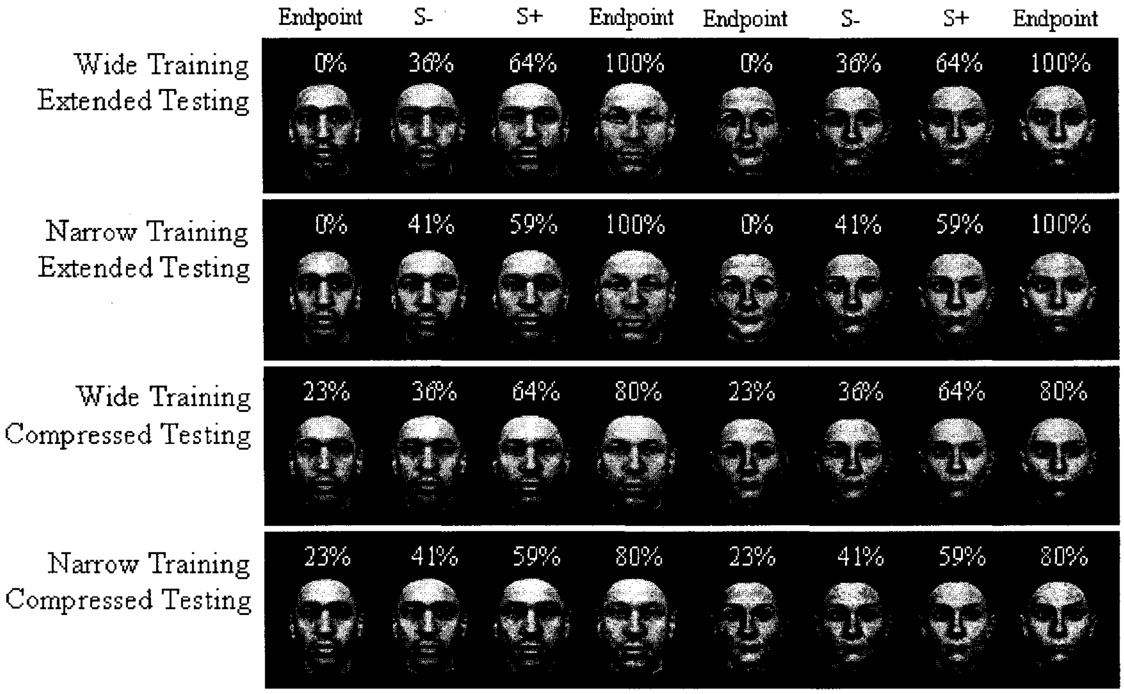


Figure 2-4. Sample face stimuli used in Experiment 3. Rows correspond to conditions as labelled, and stimulus identity is designated by column labels. Only one endpoint would be seen per subject, as described in the experimental designs. Stimulus values appear above each face. Positive and negative stimuli were counterbalanced across subjects.

Procedure

The procedure was a slightly modified version of the previous experiments. Instructions were the same as in Experiment 1a. The training criterion was changed to 20 out of 24 trials correct, over three blocks of training trials. Eleven participants who did not meet this criterion within 30 min of training were discontinued prior to testing. For the 85 participants that met criterion, seven blocks of the eleven randomly-ordered test stimuli were presented. Each subject's response proportions were thus made up of 7 responses at each stimulus value. All other procedural details were the same as in previous experiments.

Results and Discussion

The data were analysed with an ANOVA with five between-subjects factors: training width (narrow or wide), test width (compressed or extended), range (positive or negative range), positive stimulus value, and face gender. This analysis revealed four significant effects: a main effect for range ($F_{(1,48)}=20.79$, $p<0.05$); a main effect for test width ($F_{(1,48)}=11.00$, $p<0.05$); a training width by range interaction ($F_{(1,48)}=6.22$, $p<0.05$); and a test width by range interaction ($F_{(1,48)}=10.63$, $p<0.05$). Analyzing the simple effects on the training width by range interaction reveals that there were only range effects for the narrow training width condition (narrow: $F_{(1,48)}=22.66$, $p<0.05$, and wide: $F_{(1,48)}=2.36$, $p>0.05$). Simple effects analysis of the range by testing width interaction revealed that range effects occurred only for extended conditions (extended: $F_{(1,48)}=29.15$, $p<0.05$, and compressed: $F_{(1,48)}=0.89$, $p>0.05$). Plots of the results for the four combinations of training and test width are presented in Figure 2-5.

Additional planned ANOVAs on each of the four groups produced by crossing the training width by testing width revealed a significant range effect in both narrow training groups. In the extended test width condition the negative range peak was 44.8%, and the positive peak was 62.1% ($F_{(1,8)}=16.34$, $p<0.05$, $\eta_P^2 = 0.67$). Predicted values were 45.2% and 72.8% respectively. The five subjects who were excluded from the analysis because they did not respond to any of the common test stimuli were all from the narrow extended condition. All of these subjects responded to the non-common stimuli on the padded end of the test stimulus range and thus demonstrated extreme response shift; this means that their exclusion would lead to an underestimation of the response shift (see Figure 2-6). It is unlikely that these outliers exemplify a powerful adaptation effect. Instead, they may have forgotten the target face far more than other participants, or perhaps they used a different response strategy. In the narrow training and compressed testing width the negative range peak was 57.9%, and the positive peak was 61.6%, ($F_{(1,13)}=5.99$, $p<0.05$, $\eta_P^2 = 0.32$). Predicted values were 52.9% and 65.6% respectively. The range effect just missed significance in the wide training and extended testing condition (the negative peak was 51.7%, and the positive peak was 58.5%, $F_{(1,15)}=3.05$, $p>0.05$, $\eta_P^2 = 0.17$, predicted to be 50.2% and 77.8%); and it did not approach significance in the wide training and compressed test range condition (a negative range peak of 58.2% and a positive peak of 57.8%, $F_{(1,12)}=0.004$, $p>0.05$, $\eta_P^2 = 0.00$, predicted to be 57.9% and 70.6%).

These results suggest that the occurrence of a range effect with morphed faces depends both on the discriminability of the training stimuli and on the extent of the range manipulation. Larger range effects occur with a large change in adaptation level, as produced by an extensive range manipulation, and with a more difficult discrimination.

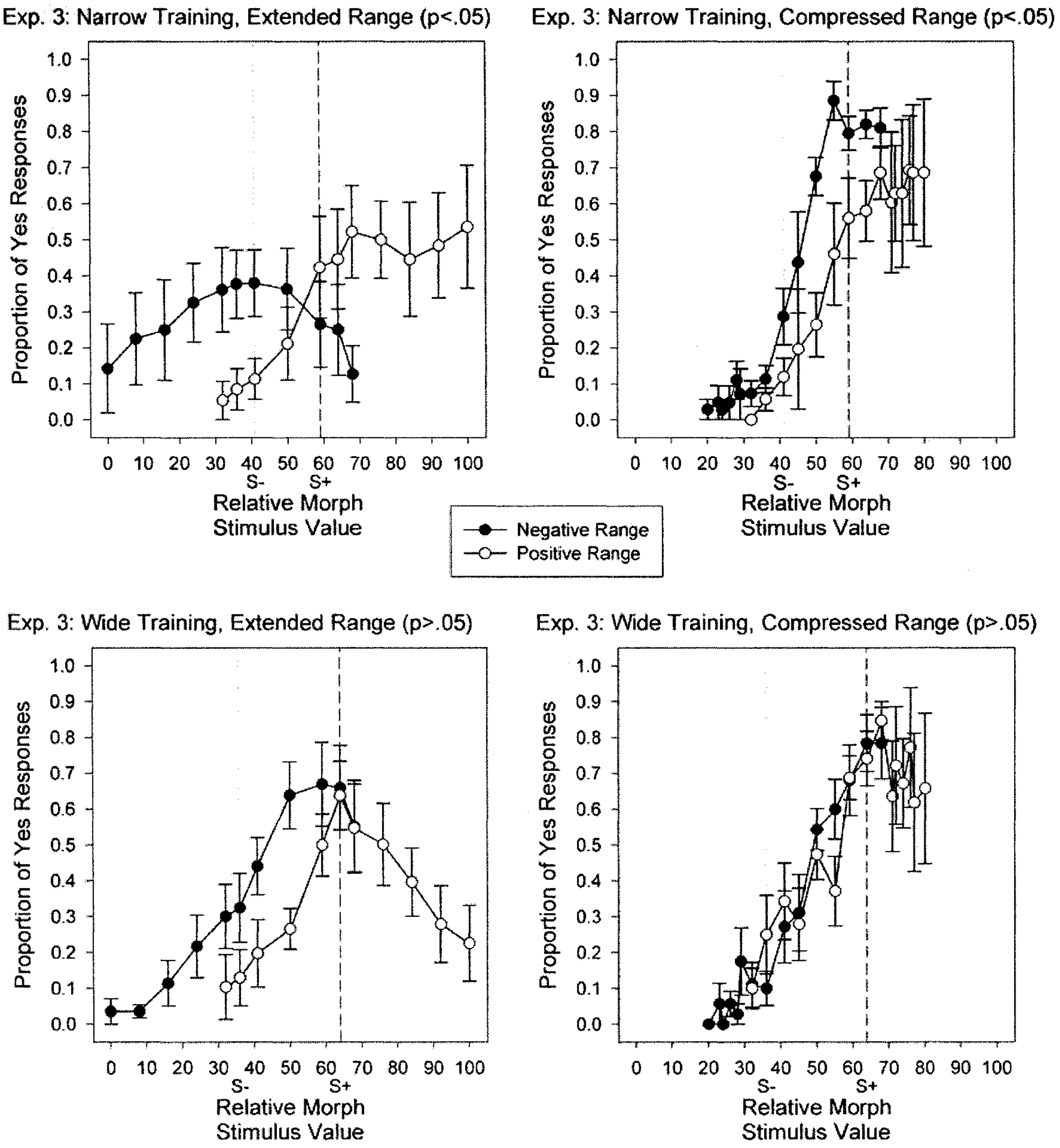


Figure 2-5. Proportion of "Yes" responses to test stimuli in Experiment 3 for crossings of Training Width and Range Width conditions, broken down by Stimulus Range condition (M ± SE). Face Gender and Positive Stimulus conditions have each been collapsed across. Only stimuli common to both ranges were used in analysis.

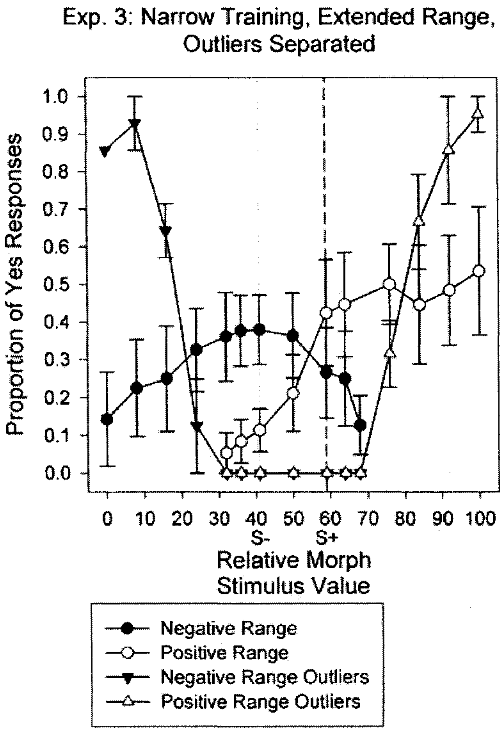


Figure 2-6. Proportion of "Yes" responses to test stimuli in Experiment 3 for the Narrow Training Width and Extended Range Width condition, broken down by Stimulus Range condition, and by outlier status (M ± SE). Face Gender and Positive Stimulus conditions have each been collapsed across. This shows the difference between subjects used in the analysis and those that were necessary to leave out.

General Discussion

These three experiments demonstrate that range effects can be found with face stimuli, but that only certain conditions allow it to emerge. First, range effects appear readily when faces are varied along the simple dimensions of brightness and orientation. Second, range effects appear with morphed faces only when the training conditions are relatively difficult, and only when the range manipulation produces a sufficiently great change in adaptation level.

Although we did not find a range effect in the morphed faces of Experiment 1a, increasing the range manipulation, and slightly increasing the discriminability of the training stimuli resulted in a range effect in Experiment 2a. Conversely, when faces varied in brightness, a range effect occurred in Experiment 1b and remained even when we weakened the range manipulation and decreased discriminability in Experiment 2b. In Experiment 3 we varied both discriminability and extent of range manipulation for morphed faces, and found that both factors were important. Specifically, a difficult training discrimination and a wide test range manipulation each contributed to the occurrence of a range effect. Indeed, when these conditions were both present (extended range, narrow discrimination), a very sizable range effect occurred, and when they were both absent (compressed range, wide discrimination) no range effect occurred. Thus, range effects appear to readily occur when faces are varied along a simple dimension, but when they are varied along the complex morphed dimension, training difficulty and extent of the range manipulation are important factors.

Direct comparisons between Experiment 3 and the morphed face conditions of the first two experiments are problematic because of differences in the faces used, and the way in which the morphed range was created (i.e., morphing between a unique and average face versus morphing between two unique faces). However, in the context of Experiment 3, it is interesting that no range effects were found in Experiment 1a which had a relatively difficult discrimination, and a weak range manipulation, whereas range effects were found in Experiment 2a, which had an easier discrimination and a much stronger range manipulation. It seems likely that the increased range manipulation was responsible for the occurrence of range effects in Experiment 2a, but with morphed faces there may also be an optimal level of discriminability for producing range effects and this level is possibly dependent on the strength of the range manipulation. By contrast, the occurrence of range effects on the brightness dimension does not appear to require a difficult discrimination: participants in Experiment 1b learned the brightness discrimination readily yet showed strong range effects.

The impact of extent of range manipulation is sensible, according to the adaptation level hypothesis: if encoding and responding are relative to the adaptation level, and the adaptation level changes, then response rate to a given stimulus will change following the shift in adaptation level. The greater the change in adaptation level, the greater the change in the location of peak responding.

The relationship between training difficulty and range effects is less clear. Previous predictions and empirical results by Thomas and colleagues suggest that range effects should increase with greater discriminability between training stimuli (Thomas, Svinicki & Vogt, 1973; Thomas, Mood, Morrison & Wiertelak, 1991). Although the range effects seen with the brightness dimension in Experiments 1b and 2b seem consistent with this prediction, the results of Experiment 3 are not, because range effects for morphed faces were stronger with the more difficult discrimination. Similarly, Baron (1973) found that humans showed a larger response shift during generalization testing for discriminations with a narrow difference (100Hz) than discriminations with a wider difference (200Hz) in tone frequency. Clearly more research is needed to determine why increases in discriminability of the training stimuli sometimes increase and sometimes decrease range effects.

Overall, our results clearly show that range effects can occur with complex face stimuli. When face stimuli are varied along the complex dimension produced by morphing between individuals, the range effect is fragile and is sensitive to the both the discriminability of the training stimuli and the extent of the range manipulation. Nevertheless, with a difficult discrimination and an extreme range manipulation, we found a large range effect with morphed faces, thus lending further support for Thomas's (1993) contention that adaptation level is an important determinant of stimulus generalization in humans. It is very likely that Spetch, Cheng and Clifford et al. (2004) did not find range effects in morphed faces because of smaller range manipulations; common stimulus regions were larger in some cases, and ranges differed in width rather than direction, which has less power than the bidirectional approach used in the present study.

Although our results clearly show that range effects can occur when complex stimuli are varied along a complex dimension, the present results, together with those of Spetch et al. suggest that such effects may be more sensitive to procedural variations compared to range effects that occur with simple stimuli. Moreover, the possibility exists that some participants may have attempted to simplify the complex stimuli into a simple dimension for discrimination for example, their head width or nose length. In future research it would be interesting to include probe tests in which a single stimulus feature is varied to determine the strategy used and then to compare generalization and range effects under conditions in which this strategy can be used and one in which it is prevented.

There is one provocative set of results demonstrating what appears to be a natural occurrence of range effects in face perception (Webster, Kaping, Mizokami, & Duhamel, 2004). Webster et al. found that after adapting to an exemplar from the "Asian face" category, the categorical boundary between Asian and Caucasian moved further into the Asian category: a face that was previously seen as being midway between Caucasian and Asian was seen as clearly Caucasian after adaptation to an Asian face. Although their experiment dealt with effects that are better termed adaptation after-effects (due to the very short time frame involved; see O'Toole, Vetter, & Blanz, 2001, and Webster & MacLin, 1999), some further correlational results appear to match our findings of range effects in faces. Asian students who had been in North America for at least one year set different boundaries from their newly arrived exchange student counterparts, and they showed a negative correlation between how long they had been in North America and how Asian their categorical boundary was. In other words, students who had been in North America for a long time were more likely to judge a face in the middle of spectrum as being Asian than were those who had been in North America for a shorter period of time. There was also a positive correlation between that boundary and the amount of time they reported spending with people of the same ethnicity. Thus, greater exposure to Caucasians caused the Asian student's category boundary to become less Asian, and more Caucasian. Although Thomas' (1993) adaptation level hypothesis focuses on explaining peak shift in generalization testing, it might also have implications for the role of adaptation level in determining categorical boundaries for stimuli. If categorical boundaries are encoded relative to adaptation level, these boundaries should move in the same direction as the adaptation level moves, and experimental results that parallel Webster et al.'s correlational findings might be possible.

References

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The "oblique" effect in man and animals. Psychological Bulletin, 78(4), 266-278.

Baron, A. (1973). Postdiscrimination gradients of human subjects on a tone continuum. Journal of Experimental Psychology, 101(2), 337-342.

Collishaw, S.M., & Hole, G.J. (2001). Is there a linear or nonlinear relationship between rotation and configural processing of faces? Perception, 31, 287-296.

Ghirlanda, S., & Enquist, M. (2003). One century of generalization. Animal Behaviour, 66, 15-36.

Hanson, H.M. (1959). Effects of discrimination training on stimulus generalization. Journal of Experimental Psychology, 58, 321–334.

Hays, William L. (1994). Statistics. Belmount: Wadsworth.

Helson, H. (1947). Adaptation-Level as Frame of Reference for Prediction of Psychophysical Data. American Journal of Psychology, 60(1), 1-29.

Leopold, D.A., O'Toole, A.J., Vetter, T., and Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. Nature Neuroscience, 4, 89-94.

Lewis, M.B., and Johnston, ,R.A. (1999). Are Caricatures Special? Evidence of peak shift in face recognition. European Journal of Cognitive Psychology, 11(1), 105-117.

MacKinnon, M. (1972). Adaptation-level theory, anchor theory, and the peak-shift phenomenon. Journal of Motor Behavior, 4(1), 1-12.

Nallan, G.B., McCoy, D.F., Pace, GM., & Welch, R. (1979). Generalization gradients following differential intradimensional autoshaping. Perceptual and Motor Skills, 48(2), 671-677.

Spence, K.W. (1937). The differential response of animals to stimuli differing within a single dimension. Psychological review, 44, 430-444.

Spetch, M.L., Cheng, K., & Clifford, C.W.G. (2004). Peak shift but not range effects in recognition of faces. Learning and Motivation, 35, 221-241.

Shepard, R.N. (1987). Toward a universal law of generalization for psychological science. Science, 237, 1317-1323.

Thomas, D.R. (1962). The effects of drive and discrimination training on stimulus generalization. Journal of Experimental Psychology, 64, 24-28.

Thomas, D.R. (1993). A model for adaptation-level effects on stimulus generalization. Psychological Review, 100(4), 658-673.

Thomas, D.R., and Jones, C.G. (1962). Stimulus generalization as a function of the frame of reference. Journal of Experimental Psychology, 64(1), 77-80.

Thomas, D.R., Mood, K., Morrison, S. and Wiertelak, E. (1991). Peak shift revisited: A test of alternative interpretations. Journal of Experimental Psychology: Animal Behavior Processes, 17(2), 130-140.

Thomas, D.R., Svinicki, M.D., and Vogt, J. (1973). Adaptation level as a factor in human discrimination learning and stimulus generalization. Journal of Experimental Psychology, 97(2), 210-219.

Webster, M.A., Kaping, D., Mizokami, Y., Duhamel, P. (2004). Adaptation to natural facial categories. Nature, 428, 557-561.

Webster, M.A. & MacLin, O.H. (1999). Figural aftereffects in the perception of faces. Psychonomic Bulletin & Review, 6(4), 647-653.

Chapter 3: Categories, Landmarks, and Spatial Distortions

Categories, Landmarks and Spatial Distortion

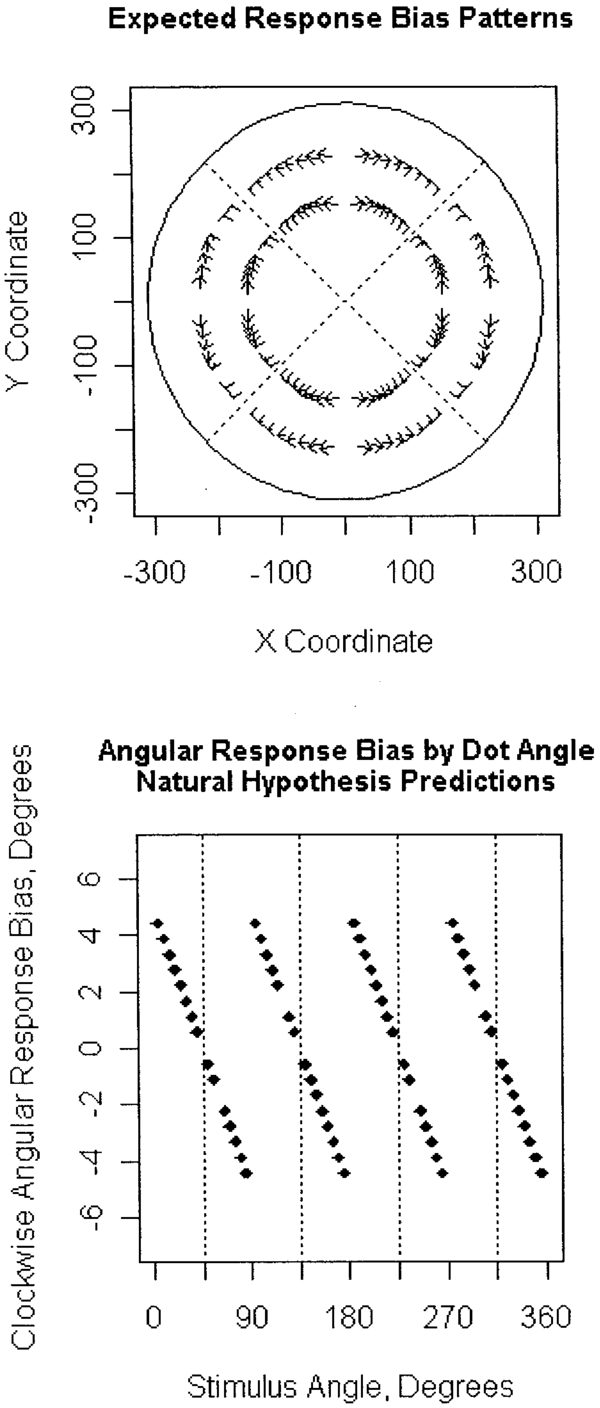
Eric Verbeek
Marcia Spetch
University of Alberta

Categories, Landmarks and Spatial Distortion

Much of human memory is influenced by categorical knowledge. In particular, memory for a category member can be distorted towards the prototype of its category. Evidence of this has been demonstrated in a variety of domains: memory of an artificially colored object will shift towards the usual color of the object (Belli, 1988); when attention is compromised, judgments of people bias towards any stereotypical categories they might fit into (Neuberg and Fiske, 1987); and spatial memory of a dot's position is affected by spatial category membership (Huttenlocher, Hedges and Duncan, 1991).

Huttenlocher et al. (1991) developed a categorical adjustment model for spatial memory distortions. They tested it with experiments wherein a dot was presented in a circle, and participants responded in a blank circle by indicating where they thought the dot was originally. Huttenlocher et al. theorized that fine-grained information about a position would be lost as one forgets, but the spatial category in which the dot was positioned would be remembered. As forgetting occurs, the degraded, fine-grained-memory of the position would be averaged with the prototypical value of the category. Furthermore, they theorized that memory of the value would be truncated to be more consistent with the category, rounded, and placed on more coarse-grained measurement scales. All of these factors result in a memory for the position that is biased toward the center of the category (i.e., toward the prototype). They also claim that this results in better response accuracy because the bias in responding allows for a large reduction in the variability in responding. In Huttenlocher et al.'s experiments, the dot's position was remembered as being angularly closer to the oblique axes of the circle, towards the 45 degree angles of the four Cartesian quadrants. From this they inferred that the categorical breakdown of the circle is into four quadrants, and they demonstrated that polar coordinates seem to capture participant's responding most appropriately. They also found that the angular and radial biases in dot reproduction are independent of one another, and that the dot angular biases roughly followed a linear function of within-category angle, with high bias near the boundaries, and low bias near the prototypes. Also, the angular bias toward the prototype was least when the dots were extremely close to the category boundaries (the cardinal axes, horizontal and vertical), but greatest when just a small distance from the boundaries. Some of these general patterns are displayed in Figure 3-1.

Huttenlocher, Hedges, Corrigan and Crawford (2004) tried to influence category formation in the dot and circle task. Each of their experiments involved the use of a distribution of dots that was not well captured by the natural quadrant scheme. Dots were clustered fairly heavily toward the cardinal axes, and no dots appeared near the oblique axes. Theoretically, it would be optimal to form four categories that corresponded to the natural quadrants, but rotated by 45 degrees. There would thus be the top, bottom, left and right categories, as opposed to the top-right, bottom-right, bottom-left, and top-left categories that people naturally use. Their strongest manipulation was to show participants the distribution from which the sample dots were drawn, and then get them to categorize each stimulus dot into the top, bottom, left or right category prior to making their dot position estimate. Surprisingly, even this rather direct manipulation did not affect the response biases, and the evidence indicated that participants still used their natural categorization scheme.



Expected Response Bias Patterns

Angular Response Bias by Dot Angle
Natural Hypothesis Predictions

Figure 3-1. These plots show an idealized pattern of bias for remembered dot locations for a blank circle. The dotted lines represent the category prototypes. The first panel shows a spatial representation, and the second panel demonstrates those biases in a scatterplot.

In the present study, we attempt to modify participant categorical schemes using what we believe to be an even more direct approach. Rather than implying categories with dot distributions and a categorization task prior to dot estimation, we drew sectioning radial lines in the circle stimulus display, and sometimes in the response display. These appear to be obvious categorizations of the circle, and took on four forms as described below in the experiments. This gives subjects the opportunity to encode the dot position in the context of explicit boundaries, and we may determine if they use these new categories or if they continue to use the natural Cartesian quadrant categories. It is also possible that people will not use the lines as category boundaries, but rather as landmarks. If this is the case, the response bias will be towards the lines, rather than towards the prototypes of the categories that may be formed by the lines.

Experiment 1

This experiment was designed to determine whether visible divisions in a circular space could induce alternate category use in participants performing a dot location memory task. Either 3 or 4 radial lines in one of two configurations produced four possible visible breakdowns of space for the circle. We will be testing three hypotheses in this experiment: the Boundary hypothesis, the Landmark hypothesis, and the Natural hypothesis. The Boundary hypothesis is that the sectioning lines will be treated as category boundaries, and that the dot location estimates will be biased towards the centroid regions of those categories. The Landmark hypothesis is that biases will be towards the sectioning lines, using them as landmarks rather than categorical boundaries. The Natural hypothesis is that the sectioning lines will have no effect on responding, and that response biases will be towards the centroid regions of the four Cartesian quadrants of the circle, just as in previous studies with completely unsectioned circles.

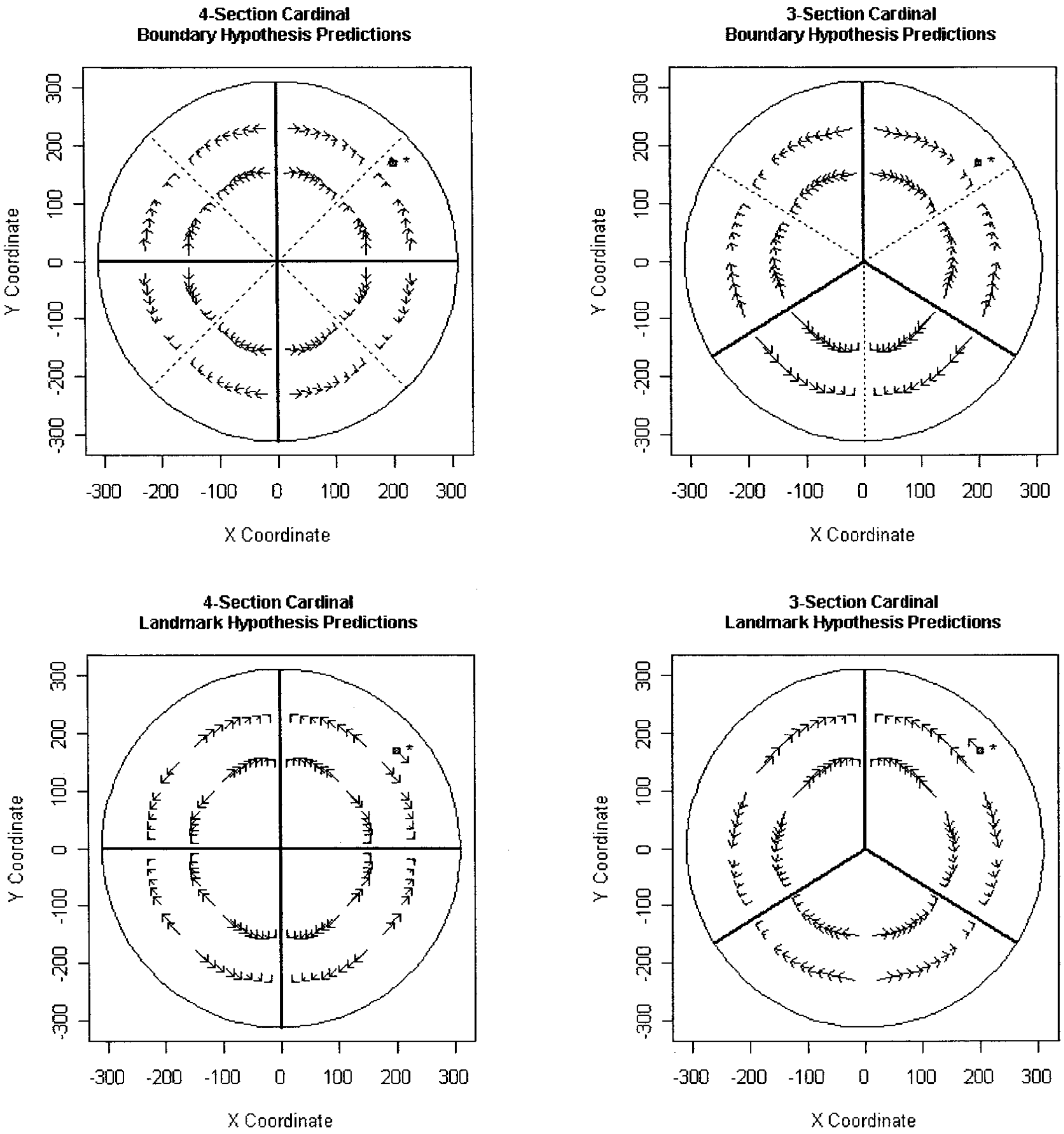
Methods

Participants

Participants were University of Alberta undergraduates enrolled in first year psychology courses. They received course credit for participation. Data from 112 students were used in the analysis.

Design

Participants were randomly assigned to four category conditions, which determined how their circle would be sectioned: 4-section cardinal orientation, 4-section oblique orientation, 3-section cardinal, and 3-section oblique (Figure 3-2). The experiment consisted of two phases with 60 trials each. Each trial used one dot location. Both phases had stimulus circles sectioned according to the condition, but one of the two phases had a sectioned response circle, while the other phase left the response circle blank. Phase order was counterbalanced across participants.



4-Section Cardinal Boundary Hypothesis Predictions

3-Section Cardinal Boundary Hypothesis Predictions

4-Section Cardinal Landmark Hypothesis Predictions

3-Section Cardinal Landmark Hypothesis Predictions

Figure 3-2. All four panels represent idealized patterns of bias under different situations. The top two panels correspond to the Boundary hypothesis, wherein participants treat lines as category boundaries. The resulting biases are away from the visible category boundaries (dark lines), and towards the inferred category prototypes (dashed lines). The bottom two panels correspond to the Landmark hypothesis, that subjects will use the lines as reference points and bias their responses towards them. The oblique conditions would be simple rotations of the cardinal patterns and thus are excluded for brevity.

Stimuli

Stimuli were presented on either 17" CRT or 17" LCD monitors; resolution was set to 1280 x 1024 pixels. Stimuli were yellow and displayed on a grey background. Seated participants viewed the screen from a distance of 30 to 60 cm. The circle and its sections were drawn with a 1 pixel (0.25 mm) thick line, and had a radius of 305 pixels (76.25

mm). Dots were squares with 6 pixel (1.5mm) sides. Dot locations had one of two set radius values, approximately either 38 pixels or 57 pixels (9.5 mm or 14.25 mm). This provided a wide distribution of dot locations while keeping the absolute number of dots to a minimum. The dot angular locations were kept at least 5 degrees away from any of the possible section lines (at 0, 45, 60, 90, 120, 135, 180, 225, 240, 270, 300, and 315 from vertical), but otherwise occurred every 5 degrees. Combining radius and angles, there were a total of 60 dot locations, each dot presented once during each of the two experimental phases described above. Circle section diagrams are available in Figure 3-2.

Procedure

Stimuli were presented on two computer stations concurrently so that two participants could be tested at the same time. The experimenter, who remained in the testing room, gave participants simple verbal instructions, directed them to read the instructions on the screen, then reiterated those instructions and asked if they had questions. The verbal and written instructions included the following information. A yellow circle will appear on screen, and soon after a dot will appear in the circle. A short time after the dot appears, the screen will clear for a moment, and then another yellow circle will appear somewhere else on the screen. In this new circle, the participants should click where they remembered the dot to be, relative to the circle, and not relative to anything else, such as the monitor.

No mention of categories or of sections was made. Thus, participants were not provided with either a cover story or a true explanation of the sections.

Halfway through the session, the response section condition was reversed for each participant and a message screen informed them of the change. Participants who had sectioned response circles in phase 1 were informed that the lines would no longer be present in the response circles, whereas participants who had blank response circles in phase 1 were instructed that the circles would now have dividing lines. In both cases, participants were instructed to continue performing as they had before.

The circle was placed on screen for 1 second, at which point the dot was placed within it. After 1.5 seconds, the circle and dot were removed, and the response circle was immediately drawn at a random location at least 102 pixels away from the top or bottom edge of the screen, and 128 pixels away from either side. The response circle was presented until a response within the circle was made, although responses within the first second of the display were blocked. Following a response, the response dot was presented for 1/2 second. The inter-trial interval then began, presenting only a grey screen, for 6 seconds. The experiment took approximately 25 minutes, including instructions and debriefing.

Analysis

Our four sectioning-line conditions divided the same circular space in different ways as seen in Figure 3-2. The sectioning-line conditions are named according to the number of sections and the orientation of the lines (e.g., 4-section cardinal has lines following the cardinal, as opposed to oblique, axis).

Circle categorization is typically demonstrated by linear regression of angular response bias on stimulus dot angle, separately for each of the four Cartesian quadrant categories (Huttenlocher et al., 2004). The angular bias (response angle minus stimulus angle) is calculated such that positive values are clockwise biases, and negative values are counter-clockwise biases. Significant negative slopes in each category indicate a bias toward the central portion of each category, or, in other words, towards the category prototype.

Since the regressor is the angle of the dot within the category, and not relative to the 360 degrees within the circle, category-relative angles are required as regressors. The three hypotheses discussed above lead to different categorization schemes for the different conditions and three possible regressors: a Natural regressor, a Boundary regressor, and a Landmark regressor. Looking at Figure 3-1 and 3-2, we see a sample dot within each circle. This dot is at 50 degrees relative to the entire circle, with 0 degrees at the top of the circle. As you can see in the figure, this dot will be in a different category for each section condition, and thus will have a different category-relative angle per condition. Under the Natural hypothesis (which assumes that the divisions for each condition will not matter), this dot will be in the top-right category, and will be given a Natural regressor value of 50 degrees in all conditions. Under the Boundary hypothesis, the Boundary regressor value depends on the section condition: it is 50 degrees in the 4-section cardinal condition, 5 degrees in 4-section oblique, 50 degrees in 3-section cardinal, and 55 degrees in 3-section oblique. Under the Landmark hypothesis, this dot will have Landmark regressor values of 5 degrees in 4-section cardinal, 50 degrees in 4-section oblique, 55 degrees in 3-section cardinal, and 50 degrees in 3-section oblique.

If a given hypothesis is correct, then the within-category stimulus dot angle will be a good predictor of the angular bias, resulting in a negative slope, as exemplified in the regression plot in Figure 3-1. Using a regressor from an incorrect hypothesis will result in inferior predictive power. We can thus distinguish between the three hypotheses by computing regressions using each of the three regressors under each of the four category conditions. Support for a given hypothesis will be provided if the R-square values, or model fits, are greater for that regressor than for the others regressors under each condition.

We regressed angular bias on each of these three regressors, separately for each subject. The resulting R-square values (model fits, or explained variances) were used as the data for inferential analysis, with the understanding that R-square indicates the extent to which a subject's response angle was influenced by the category boundaries that a given hypothesis specifies. For example, if the Natural categories were used by a subject exclusively, or even most of the time, then the Natural regressor would result in a strong negative slope, showing a positive bias (clockwise toward prototype) in the first part of the category, and a negative bias (counter-clockwise toward prototype) when the stimulus dot angle was beyond the prototype. If that participant used the lines as category boundaries instead, then the Boundary regressor would lead to a stronger negative slope (and thus R-square), which would support the Boundary hypothesis.

Note that the predictions of the three hypotheses overlap in each of the 4-section conditions such that two of the regressors are identical. In the cardinal 4-section condition, the regressor for the Boundary hypothesis is the same as for the Natural hypothesis, because the induced categories are the same as the natural ones. In the oblique 4-section condition, the Natural and Landmark regressors are the same. Both of the 3-section conditions have distinct regressors for the Natural, Boundary, and Landmark hypotheses. Thus, we have two regressors to consider for the 4-section conditions, and three regressors for the 3-section conditions.

This also leads to the possibility that one of these overlapping, or “fully correlated” regressors may explain more variance than the other regressions. For example, if the Landmark hypothesis is true, then the Natural (and thus also Landmark) regressor for the 4-section oblique condition will explain any variance attributable to natural categorization effects, and any possible additional variance that the actual visible sections could impart. This prediction will only hold if natural categorization is influencing responses either partially or alternately with any induced categorization, or if different participants use different strategies. This gives us one way to determine whether induced categories replace, or instead merely work alongside the natural categorization scheme.

Recall that participants went through two counterbalanced phases of the experiment. In one phase, both the stimulus and response circles had visible sections. In the other phase, the stimulus circle had divisions while the response circle was left blank. In line with our other predictions, we expect that the induced category effect will be larger when both circles are sectioned than when the response section is blank. Although this might seem like a trivial prediction, it will offer some indication as to whether the categorization context is important solely for encoding the dot locations, or also for decoding during responding.

Results and Discussion

In order to see which regressor predicts best under each of the category conditions, we conducted separate within-subjects ANOVAs for each of the four section conditions. This set of ANOVAs use only data from sectioned response circle phases. Each used R-Square as the dependent variable, and regressor as the independent variable. The 4-section conditions had only two levels of regressor, whereas the 3-section conditions had three levels for the reasons discussed above. The 4-section cardinal condition had the Natural/Boundary regressor and the Landmark regressor, the 4-section oblique had the Natural/Landmark and Boundary regressors, and both 3-section cardinal and oblique conditions had each of the Natural, Boundary and Landmark regressors.

Both ANOVAs for the 4-section conditions were significant ($F_{(1,27)} = 12.58$ (cardinal) and 66.51 (oblique) , $p<0.05$). As seen in Figure 3-3 the Landmark regressor was greater than the Natural/Boundary regressor for the 4-section cardinal condition, and the Natural/Landmark regressor was greater than the Boundary regressor for the 4-section oblique condition. In fact, the Natural/Landmark regressor in the 4-section oblique condition accounts for much more variance than any other regressor presumably because it captures all variation due to both the Natural hypothesis and the Landmark hypothesis.

The 3-section conditions also had significant main effects of regressor, after correcting the degrees of freedom for sphericity, using the Geisser-Greenhouse correction ($F_{(1.336,54)} = 30.04$ (cardinal) and $F_{(1.424,54)} = 40.23$ (oblique), $p<0.05$). Post-hoc t-tests demonstrate that the Prototype regressor was greater than the Boundary regressor for both ($t_{(27)} = 7.21$ (cardinal) and 6.64 (oblique), $p<0.0083$ Bonferroni corrected). The Prototype regressor was also superior than the Natural regressor for both ($t_{(27)} = 6.60$ (cardinal) and 5.50 (oblique), $p<0.0083$ Bonferroni corrected). These results are apparent in Figure 3-3.

The last set of analyses investigates the role of response circle sections. Recall that the stimulus circle always had visible sections drawn in it, and that each subject went through a phase with sections drawn within the response circle, and with the response circle empty. We expected that with blank response circles, the Natural categorization scheme might replace induced Landmark effects identified in our initial analyses. We excluded the 4-section oblique condition from this analysis because the Landmark and Natural regressors are indistinguishable in this condition. We performed separate one-way ANOVAs for the remaining three conditions, including only data from the experimental phase with an unsectioned response circle. The results showed no difference between the

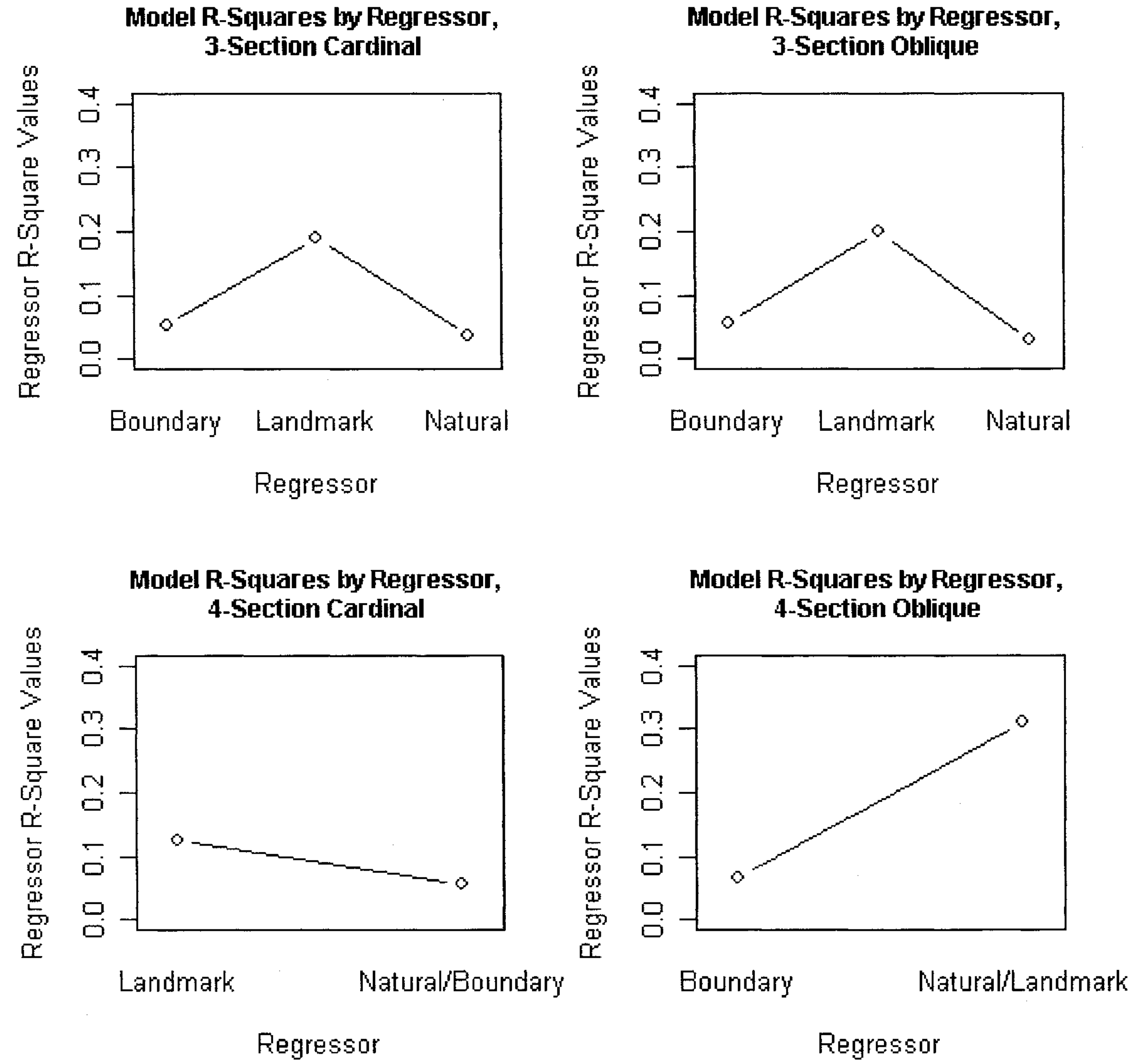
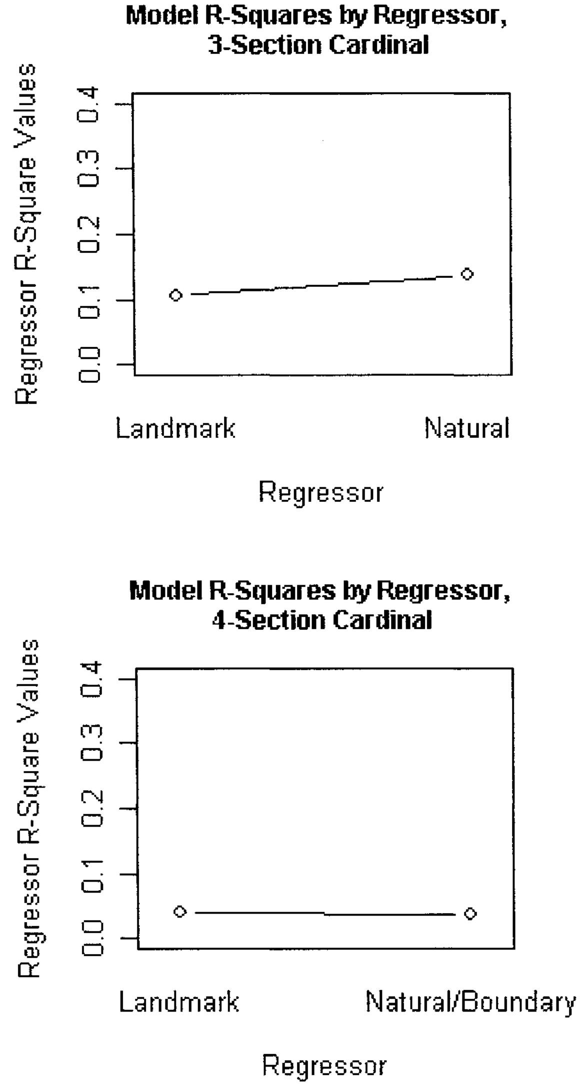


Figure 3-3. These plots show the mean R-squares or model fit values in Experiment 1, for each regressor applicable, separately for each sectioning line condition.

Landmark and Natural regressors under any of the three conditions (largest $F_{(1,27)} = 2.01$, $p<0.05$). Relative to when participants responded in sectioned circles, when they responded in blank circles, the model fits for the Landmark regressor are drastically reduced, while the model fits of the Natural regressor increased. See Figure 3-4 to see a plot of these results, which are readily comparable with Figure 3-3. Over all three conditions, the mean R-squares for the unsectioned data were 0.092 for Landmark regressor, and 0.094 for Prototype regressor.



Model R-Squares by Regressor, 3-Section Cardinal

Model R-Squares by Regressor, 3-Section Oblique

Model R-Squares by Regressor, 4-Section Cardinal

Figure 3-4. These plots show the mean R-squares or model fit values for the Landmark and Natural regressors in Experiment 1, for three of the sectioning line conditions. The fourth condition, 4-section oblique, is excluded because the Natural and Landmark regressors are functionally identical for that condition.

These results demonstrate successful category induction, and support the Landmark hypothesis. Regressors based on the hypothesis that people use sectioning lines as landmarks or category centers fit the data better than those based on the Natural categorization scheme, or on a hypothesis that people treat the sections as category boundaries. Furthermore, the Landmark/Natural regressor in the 4-section oblique condition explained the greatest variance, presumably because of the complete congruence between the landmark-based induction and the natural categorization scheme. This is indirect evidence for a mixture of strategies within or across participants, despite the support for the Landmark hypothesis. Lastly, it seems that participants use the Natural categorization scheme more and the Landmark scheme less when the sections are removed from the response circle. In this case, neither categorization scheme appears to dominate the other, since the mean R-squares are virtually identical. It is possible that participants were using a mixture of response strategies, either on a per-response level or with the two categorization schemes influencing responding conjointly, or that different participants used different strategies.

Experiment 2

This experiment was designed to further examine whether exposure to section lines would have any enduring effects on categorization of a circular space. Specifically, would participants that are exposed to sectioned stimulus and response circles in the first phase form a long-term categorical framework in which to remember dot locations, or are they only influenced by the sight of the sections as they remember and respond?

Methods

The methods were those used in Experiment 1, with some exceptions. First, all participants used 17" LCD monitors. Second, for all participants, both the stimulus circle and the response circle were sectioned (according to the participant's category condition) during the first experimental phase, whereas both the stimulus circle and the response circle were unsectioned during the second phase. Participants were University of Alberta undergraduate students enrolled in first year psychology courses. Sixty-four participated in the study, and received course credit.

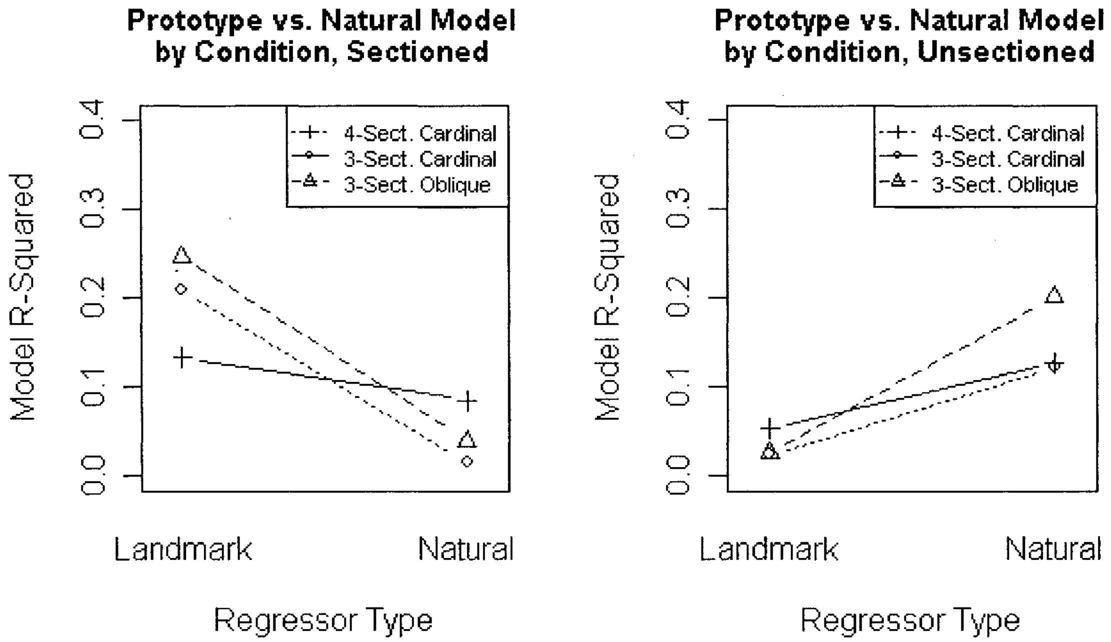
Results and Discussion

To check for differences in responding, we first performed a mixed model ANOVA with regressor type (Landmark and Natural regressors only) and section presence as within-subject factors, and condition as a between-subject factor. The four-section oblique condition was excluded from this analysis because of the complete correspondence of its Natural and Landmark regressors. There was a significant interaction between regressor type and section presence ($F_{(1,45)}=100.24$, $p<0.05$), and a significant three-way interaction ($F_{(2,45)}=8.25$, $p<0.05$).

Subsequent two-way analyses revealed that the interaction of regressor type and response sections was significant for each of the three category conditions (smallest $F_{(1,15)}=9.89$, $p<0.05$). A final set of analyses examined the effect of regressor type for each of the six combinations of category condition and section presence. (see Figure 3-5). Except for the 4-section cardinal sectioned condition ($F_{(1,15)}=1.48$, $p>0.05$), all show a significant main effect of regressor (smallest $F_{(1,15)}=21.47$, $p<0.05$). In the sectioned conditions, the Landmark regressor had a higher R-square value, whereas in the unsectioned conditions, the Natural regressor resulted in greater R-square values (smallest $F_{(1,15)}=15.12$, $p<0.05$). For the sectioned conditions (including the non-significant 4-section cardinal condition), the mean R-square was 0.215 with the Landmark regressor, and 0.103 with the Natural regressor. The corresponding means under the unsectioned conditions were 0.076 and 0.163.

In a separate analysis, we examined the effect of section presence in the four-section oblique category. The t-test was significant ($t_{(31)}=2.31$, $p<0.025$), with a sectioned mean R-square of 0.269 and an unsectioned mean of 0.199.

We reasoned that by looking only at the Natural regressor and the unsectioned phase, and comparing the 4-section oblique condition to the remaining three category conditions, we have one last way of checking for any potential carry-over of the sectioned phase. If there is a significant difference, it would be because the categorization scheme during the sectioned phase wasn't immediately eradicated when the sections were removed, thus raising the R-square of the Natural regressor in the 4-section oblique case, and lowering it in all others. The Natural regressor mean was not significantly higher for the 4-section oblique condition (0.199) than the average of the other three category conditions (0.151, $t_{(26.73)}=1.66$, $p>0.05$), thus offering no evidence of influence of the sectioned phase of the experiment on the unsectioned phase.



Prototype vs. Natural Model by Condition, Sectioned

Prototype vs. Natural Model by Condition, Unsectioned

Figure 3-5. These plots show the mean R-squares or model fit values for the Landmark and Natural regressors in Experiment 2, for three of the line section conditions. The 4-section cardinal condition is excluded because the Natural and Landmark regressors are functionally identical for that condition. Sectioned and Unsectioned circle conditions are shown separately. In Experiment 2, the Unsectioned condition followed the Sectioned condition, within subjects, and removed sectioning lines from both the stimulus and response circles.

These results are evidence that the induced categories have little longevity when they are removed from sight, and that people have a tendency of returning to their natural categorization scheme. Although other methods may succeed in changing a person's circle categorization scheme, using explicit visual indicators alone will likely not have a lasting effect.

General Discussion

Our results clearly show that the strategies people use to remember locations in a circular space can be influenced by making visible radial divisions. Somewhat unexpectedly, participants used the lines as landmarks, rather than using the natural categorization scheme, or using the lines as new category boundaries. We originally expected that the section lines would serve as category boundaries, because salient category boundaries should allow for more accurate categorization and hence better responding overall (Huttenlocher et al., 1991). Although participants did not appear to use the section lines as category boundaries, they did benefit from the presence of the lines. Looking at the overall accuracy of responses, we see that there was a substantial benefit to using the lines rather than using the natural categorization scheme. In Experiment 1, the average absolute angular bias was 3.02° when the response circle was sectioned and 6.28° when it was left blank. In Experiment 2 angular bias was 3.24° and 4.86° for sectioned and unsectioned conditions, respectively. Thus it appears to be beneficial to change from a categorical strategy to a landmark strategy. The advantage of using landmarks over the natural categorization scheme may be due to the greater salience or stability of an external reference point compared to an internal inferred prototype. Specifically, it might be easier to remember the distance of a dot from a visible landmark line than from the invisible inferred center point of a quadrant. As we will discuss later, there is existing evidence that landmark-based responding results in response biases similar to those observed in our experiments.

Our results suggest that participants use a landmark strategy when section lines are present, and a category strategy when section lines are absent. However, it is possible that the same strategy was used in both cases, but the presence of the section lines altered how the strategy was applied. It is possible that participants might have used the sectioning lines not as simple landmarks around which they base their responses, but rather as category prototypes, and then subsequently inferred the category boundaries. This would presumably be a reversal of the process used in the natural categorization scheme, but would nevertheless be a categorization strategy. Although possible, this approach seems counterintuitive because it contradicts tenets of the categorical approach (Huttenlocher, et al., 1991). For example, it would result in categories with poorly defined boundaries under the 4-section cardinal condition, and in both of the 3-section conditions, due to the oblique effect (Appelle, 1972). The resulting inferred category boundaries would be difficult for people to use, and this could result in a larger number of miscategorized dots.

Alternatively, the landmark response strategy could be the basis for the natural categorical strategy. Perhaps categories are formed so as to provide multiple, easily inferred points of reference. The prototype would thus be an invisible (but imagined) landmark, its position inferred from the category boundaries selected. This idea has face values, since the increase in bias towards the point of reference (landmark or prototype) becomes greater with distance between the reference and the location to be recalled, as we will discuss later.

It is important that the theoretical relations between landmarks and categorical prototypes be explored to see which relation between categories and landmarks is supported, and whether it can lead to unique hypotheses that we may test. If so, we may be able to determine whether the categorical strategies are meta-strategies using the landmark strategy mechanisms, or whether they are distinct strategies. Also, if both are distinct but not exclusive strategies, we can ask how landmarks and spatial categories interact.

Several studies, using somewhat different procedures have also found biases in memory produced by visible landmarks. The direction of the bias, however, is not always the same. At least one study reported an opposite bias to that found in our study (Schmidt, Werner and Diedrechsen, 2003). They presented either one or two small circular landmarks in a rectangular field and found that dot position memories were distorted away from the landmarks. Furthermore, they inferred the existence of a "virtual landmark", based on repulsive distortions away from the midpoint of the two landmarks.

More consistent with our results, Sheth and Shimojo (2001) found a bias in remembered location toward a visible marker, an effect they referred to as a compression of distance in memory. They found that dots presented in a rectangular field, positioned on the horizontal meridian of the screen, were remembered as closer to the fixation point in the center of the screen, in both a mouse-click task, and a sequential judgment task. In one of their experiments, they required participants to gaze at the central fixation point, then at one on the far right or left side of the screen during the mouse-response phase, and still found bias towards the central fixation point. As with Huttenlocher, et al. (2004), and our own results, the magnitude of these distortions were positively correlated to the distance of the stimulus dot from the point of bias fixation point. Furthermore, they also performed an experiment wherein participants viewed a central fixation point, as well as a vertical line near the far right of the screen. Dots that were presented on the far left had a greater extent of bias towards the fixation point, and those presented on the right, between the fixation point and the line, had a smaller bias toward the center of the display. Lastly, when the fixation point was removed, and participants were merely told not to gaze at the line on the right hand side, a strong bias toward the line appeared for dots presented on the right half of the screen, and a much smaller bias appeared for dots presented on the left half of the screen. Thus, their data show a reliable tendency to remember dots as being closer to salient features, a result that is congruent with our results.

Bryant and Subbiah (1994) also found a bias towards landmarks and toward what they called subjective landmarks. Dot stimuli were presented in a square field with 3 distance markers on the left and bottom sides of the square. They found that dots presented on the imaginary intersection of the markers had better production accuracy, whereas dots placed elsewhere were biased towards the nearest intersection point. In another experiment, they found an attraction bias towards single '+' marks placed on the intersection nearest to the stimulus dot. Notably, in their fourth experiment, they found that verbal instructions regarding mnemonic tactics could eliminate the bias. Some participants were told to imagine lines projecting from the tick marks on the edges of the square, and to find the intersection closest to the dot. This resulted in the same intersection bias found before. Other participants were told to imagine the projected lines, but to identify the four intersections nearest to the dot, and to imagine those intersections as being four corners of a box surrounding the dot. In this case, there was no bias towards the intersection point. The authors suspect that participants were biasing their responses towards the center of the imagined square, rather than towards the nearest intersection, although their test of this hypothesis was not significant. It would warrant another experiment with greater power, or perhaps more powerful instructions, to see if this new bias may be induced.

Some of Hubbard's evidence speaks for the possibility of using larger objects or surfaces as landmarks (1998). This set of experiments dealt with targets moving vertically alongside a filled black area. The moving target vanished at an unpredictable time, and subjects used a mouse to click where they believe they last saw it. Hubbard found that the targets were biased towards the black area, whether it was on the right or left side of the target, and were not biased towards either if the target was moving between two such black regions. Similarly, with no black area presented at all, no horizontal bias was detected. Although this result appears to coincide with our results, it would be worth replicating Hubbard's experiments, with a condition that uses a line rather than a large object edge.

Further evidence of discrete landmark attraction bias was found by Hubbard and Ruppel (2000). In their first experiment, a small target square (20 to 60 pixels wide) was presented, at one of three distances from a larger landmark square (120 pixels wide). The target was presented only in line with the cardinal axes of the landmark, above, below or to either side. After 1 second, either both the landmark and the target were removed from the screen, or just the target was removed while the landmark remained. Subjects were then able to respond by clicking where they thought the target was positioned. There was an overall downward bias (already documented as representational gravity), but there was also a tendency to bias target estimates towards the landmark. Hubbard and Ruppel found that the representational gravity effect and landmark attraction effect combined. When the target was on either side of the landmark, there was both a strong downward and landmark attractive bias. When the target was above the landmark, there was a very large downward bias; when it was below the landmark, there was only a slight downward bias, or a slight upward bias (depending on the size of the target square, with larger squares having the stronger downward tendency). Participants also showed a greater bias toward the landmark when the target was further from it than when it was closer. Finally, on trials where both the landmark and target were removed from sight prior to responding, all of the landmark attraction effects were increased. This seems incongruent with our results; we would have expected the overall display to be spontaneously split into some set of categories, resulting in a new pattern of response biases. Also, in our Experiment 1, you will recall that when the sectioning lines were removed following stimulus display, the Landmark regressor lost much of its explanatory strength, while the Natural regressor gained strength, suggesting a mixture of response strategies. Their results predict that we would have found an even greater model fit for the Landmark regressor when we removed the lines. As we hypothesized before, this might be due to a switch in the frame of reference in our experiments, and perhaps participants were not predisposed to any other particular categorization scheme for the display that Hubbard and Ruppel used.

This notion of switching frames of reference is supported by Tversky and Schiano (1989). In their fifth experiment, participants were shown a straight line at an angle (ranging between 20 and 70 degrees), which was shown in an L-frame with tick mark (as would frame a graph). The figures were labeled either to appear as graphs, or to appear as maps, and subjects were given blank L-frames in which they were to draw the line they had seen. There was a distinct tendency for people to bias the line towards the 45 degree angle for the graph condition, while no such bias was found in the map condition. The extent of bias was also positively correlated with the angular distance between the true line and the 45 degree angle, reflecting other experimenter's results discussed thus far. This sort of instructional frame of reference effect is also seen in the results of Bryant and Subbiah (1994) which we discussed above; biases towards subjective landmarks at inferred intersections were eliminated when participants were asked to imagine a matrix of squares rather than focusing on the intersections.

The evidence suggests that higher-level, conceptual frameworks might be able to overshadow lower-level spatial memory strategies. The results of our present experiments suggest that the opposite could be true as well, that a simpler landmark-based response strategy can become more predominant than a hierarchical categorical memory strategy. Again, this might be due to the superiority of explicit landmarks in a given context. We think it is important to determine the relations between categories and landmarks. It does not seem terribly likely that all categorical judgment effects can be explained by a landmark account, although it seems possible that the categorical strategy is used to produce implicit landmarks. If it turns out that we cannot explain all of spatial memory biases the results with a simple landmark hypothesis, then we can take lead from Hubbard (1998); just as he investigated the relations among referential gravity, friction, and landmark attraction effects, we can ask how categorization, landmarks, and other frames of reference interact with one another. It might be the case that explicit landmarks such as drawn lines, circles, and other figures, and implicit, virtual, subjective, or emergent landmarks (such as category prototypes, frames of reference, inferred intersections) are both making use of the same landmark based response mechanisms. If so, it is the selection of landmarks, or frame of reference that differs under various conditions. Personal experience, education, experimental instructions, and stimulus affordances could each have a strong influence on frames of reference, and all of these factors might be considered in discovering how people remember locations in space.

References

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The "oblique" effect in man and animals. Psychological Bulletin, 78(4), 266-278.

Belli, R.F. (1988). Color blend retrievals: Compromise memories or deliberate compromise responses? Memory & Cognition, 16(4), 314-316.

Bryant, D. and Subbiah, I. (1994). Subjective landmarks for perception and memory for spatial location. Canadian Journal of Experimental Psychology, 48(1), 119-139.

Hubbard, T.L. (1998). Some effects of representational friction, target size, and memory averaging on memory for vertically moving targets. Canadian Journal of Experimental Psychology, 52(1), 44-49.

Hubbard, T.L. and Ruppel, S.E. (2000). Spatial memory averaging, the landmark attraction effect, and representational gravity. Psychological Research, 64, 41-55.

Huttenlocher, J., Hedges, L., Corrigan, B., and Crawford, E. (2004). Spatial categories and the estimation of location. Cognition, 93, 75-97.

Huttenlocher, J., Hedges, L.V. and Duncan, S. (1991). Categories and Particulars: Prototype Effects in Estimating Spatial Location. Psychological Review, 98(3), 352-376.

Neuberg, S.L. and Fiske, S.T. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. Journal of Personality and Social Psychology, 53(3), 431-444.

Schmidt, T., Werner, S. and Diedrechsen, B. (2003). Spatial distortions induced by multiple visual landmarks: How local distortions combine to produce complex distortion patterns. Perception and Psychophysics, 65(6), 861-873.

Sheth, B. and Shimojo, S. (2001). Compression of space in visual memory. Vision Research, 41, 329-341.

Tversky, B. and Schiano,D.J. (1989). Perceptual and Conceptual Factors in Distortions of Memory for Graphs and Maps. Journal of Experimental Psychology: General, 118(4), 387-398.

Chapter 4: Conclusion

Conclusion

This thesis has presented two sets of studies that demonstrate the effects of frames of reference on two different types of memory tasks. The first (Ch. 2) demonstrated that facial discrimination and recognition may sometimes be affected by test stimulus range. Under particular experimental parameters, a biased range of test stimuli can result in a response bias in the same direction as the range bias, if the training task was of sufficient difficulty, and if the extent of the range bias is large enough. Theoretically, the effect stems from a shift in the adaptation level of the respondent, which is used as a reference point for remembering the target face, to which they learned to respond. Thus, by biasing test range, participant's memory for the target face is systematically distorted in the same direction as the range bias. The second study (Ch. 3) demonstrated that sectioning lines within a circle lead to a functional use of the lines as landmarks, rather than as categorical boundaries. This is unusual in that without lines, people spontaneously produce spatial categories corresponding to the four Cartesian quadrants. Subsequently, their memories of stimulus dot locations are biased towards the centers of the four spatial categories as replicated in this thesis. When sectioning radial lines were present, in particular during the response phase, participants' memory of the dot locations were biased towards the lines themselves, rather than towards the center the categories as defined by the lines.

While the first study showed that test stimulus distributions may affect memory for a target, the second showed that additional features of the stimulus display may likewise change the participants' memories. Both of these effects thus show that the context or frame of reference in which one remembers a stimulus can distort one's memory to another relatively likely candidate, though neither effect is independent of the conditions under which the target stimulus is first experienced. In particular, in the third experiment of Chapter 2, we found that difficult training discriminations led to significant range effects, but that easier discriminations seemed to hinder the development of the effects. In the first experiment of Chapter 3, we demonstrated that when the stimulus was presented in a sectioned circle, and the response was made in a blank circle, that there was no particular strategy that seemed dominant over the others; participants seemed to be influenced by both the natural categorization scheme, as well as by a strategy that uses the section lines as landmarks. In both chapters, the focus of our investigations was on the stimulus context during recall, but further research might expand investigation in these paradigms to include further manipulation of the context during encoding. In fact, our findings suggest the importance of the context of stimulus encoding. In particular, when the stimulus and response circles were both blank, participants returned to using the natural categorization scheme

In both chapters we also dealt with singular points of reference: the adaptation level in Chapter 2, and the category prototype and landmark in Chapter 3. The adaptation level is theoretically an average stimulus value over a period of time that an organism maintains as a reference point for stimulus generalization (Thomas and Jones, 1962). Similarly, a category prototype serves as a point of reference for category memories, as a mnemonic device (Huttenlocher, Hedges and Duncan, 1991). The similarity between the reference points doesn't go much further though. Theoretically, people bias their memories of location toward spatial category prototypes. On the other hand, response gradients follow the adaptation level, but they are not biased inward towards it per se. Also, adaptation level is derived over time from a sample of stimuli, whereas the prototype does not appear to be affected by stimulus distributions as shown by Huttenlocher, Hedges, Corrigan and Crawford (2004). The prototype is distinct from an exemplar, which is specifically in reference to experience (Ch. 8 of Reisberg, 1997), and which is thus similar to an adaptation level to some extent. Further comparison and contrast of different frames of reference such as prototypes and adaptation levels is possible. Continued investigation of their roles in cognition, in comparison to one another, will more specifically explain how organisms deal with the world around them.

References

Huttenlocher, J., Hedges, L., Corrigan, B., and Crawford, E. (2004). Spatial categories and the estimation of location. Cognition, 93, 75-97.

Huttenlocher, J., Hedges, L.V. and Duncan, S. (1991). Categories and Particulars: Prototype Effects in Estimating Spatial Location. Psychological Review, 98(3), 352-376.

Reisberg, D. (1997). Cognition, Exploring the Science of the Mind. New York, NY: W. W. Norton & Company, Inc.

Thomas, D.R. and Jones, C.G. (1962). Stimulus generalization as a function of frame of reference. Journal of Experimental Psychology, 64(1), 77-80.