Available online at www.sciencedirect.com

SCIENCE DIRECT®

Learning and Motivation 35 (2004) 221–241

Learning and Motivation

www.elsevier.com/locate/l&m

Peak shift but not range effects in recognition of faces

Marcia L. Spetch,[a,*] Ken Cheng,[b] and Colin W.G. Clifford[c]

[a] *Department of Psychology, University of Alberta, Edmonton, Alta., Canada T6G 2E9*
[b] *Macquarie University—Sydney, Australia*
[c] *University of Sydney, Australia*

Received 8 August 2003; received in revised form 4 November 2003
Available online 21 February 2004

Abstract

University students were trained to discriminate between two gray-scale images of faces that varied along a continuum from a unique face to an average face created by morphing. Following training, participants were tested without feedback for their ability to recognize the positive face (S+) within a range of faces along the continuum. In Experiments 1 and 4, the range of stimuli presented during testing was manipulated. In Experiment 2, participants viewed different ranges of faces during an adaptation period that followed training and preceded testing. In all experiments, generalization functions revealed peak shifts or area shifts (fewer "yes" responses to novel faces on the negative side of the S+), but no systematic effects of the test or adaptation range. Peak shift was found both for upright and inverted faces and occurred even if the orientation of the face was reversed between training and test. Using similar methods, either an area shift or range effect (but not both together) was demonstrated for line tilt stimuli (Experiment 3), and the appearance of these effects depended on instructions. It appears that peak shift and area shift are robust across many different kinds of stimuli, but range effects may not readily occur with complex multidimensional stimuli.

The processes of discrimination and generalization have typically been investigated using unidimensional stimuli such as wavelengths, brightnes of a light, or tilts of a line (e.g., Guttman & Kalish, 1956, Thomas, Mood, Morrison, & Wiertelak, 1991). Typically, subjects are trained to perform a particular response to one

---

[*] Corresponding author. Fax: +780 492-1768.
E-mail address: mspetch@ualberta.ca (M.L. Spetch).

doi:10.1016/j.lmot.2003.11.001

stimulus (S+) and then are tested with a range of stimuli that includes the S+. Usually the test stimuli vary along a single dimension, and an orderly generalization gradient is found in which responding decreases systematically as the stimuli become less similar to the S+ (e.g., Guttman & Kalish, 1956). Some discrimination training can produce systematic deviations from the typical generalization gradient. Intra-dimensional discrimination training in which subjects are rewarded for responding to an S+ and not rewarded for responding to another stimulus (S−) that vary on a single dimension can produce asymmetrical generalization gradients during testing. These gradients may be characterized by "peak shift" or "area shift." In peak shift, subjects respond more to a test stimulus that is different than the S+ in the direction away from the S− (hence shifting the peak of the distribution in the direction opposite to the S−). In area shift, the peak is not shifted but subjects respond more to stimuli on the S+ side of the distribution than to stimuli on the S− side. Peak shift was first identified by Hanson (1959) and has since been shown in many dimensions of experience with nonhuman animals (for reviews, see Ghirlanda & Enquist, 2003; Honig & Urcuioli, 1981; Purtle, 1973; Rilling, 1977).

In studies with animals, peak shift effects have most often been interpreted in terms of an interaction of excitatory and inhibitory gradients. According to Spence's classic theory (1937), an excitatory gradient around the S+ and an inhibitory gradient around the S− form during discrimination training, and the interaction of these gradients is responsible for the shift in the peak during generalization testing. In humans, by contrast, any peak shift has mostly been attributed to response range effects, in which the location of the peak of responding depends on the range of stimuli used in testing (e.g., Thomas, 1974, 1993, review; Thomas et al., 1991). Specifically, a range that includes mostly values on the S+ side of the distribution results in a shift away from the S+ to a value on the S+ side of the distribution (just like the standard peak shift effect). However, if the test range includes mostly values on the S− side of the distribution, then an opposite shift in the peak is found, toward the S− side of the distribution. Thomas (1993) accounted for response range effects with a model based on Helson's (1964) "adaptation level" theory. Thomas proposed that subjects code stimuli relatively, with respect to an adaptation level. The adaptation level is determined by the average of all stimuli encountered. For instance, if S+ is brighter than S−, S+ is coded as brighter than the adaptation level, while S− is coded as dimmer. As tests are given, the adaptation level is determined more and more by the range of stimuli found on tests, and thus the range influences where the peak of responding is found. Thomas (1993) concluded that most, if not all, peak-shift-like effects in humans are actually range effects. It should be noted that the standard test for peak shift, which is to present an equal range of stimuli on either side of the S+, cannot distinguish between a true peak shift effect and one that is produced by range effects. Specifically, during training with the S+ and S−, the adaptation level would fall between these two stimuli, whereas during testing with a symmetrical range around the S+ the adaptation level would fall at the S+. Thus, learning that the S+ is a certain distance away from the adaptation level would lead to a shift in the peak during testing.

Although most studies of peak-shift and range effects have used simple stimuli that vary along a single dimension, a few studies have demonstrated peak-shift effects with more complex stimulus dimensions. For example, Honig and Stewart (1993) trained pigeons to discriminate between stimulus displays that differed in relative numerosity (e.g., proportion of red dots in a display of red and green dots). When tested with a range of displays that varied in relative numerosity, the typical peak shift effect was found. Range effects were not tested. Cheng and Spetch (2002) recently found peak shift in human spatial generalization, with range effects and adaptation level controlled. People were trained to discriminate between the same stimulus presented at two different locations on a computer screen and then during testing the stimulus was presented in a range of locations. The tests were given only occasionally, interspersed among training trials. In contrast, in most studies with humans, tests were given as a block after sufficient training. The occasional testing served to keep the adaptation level at its training level. Peak shift and area shift were found, while range effects were not evident. One might consider spatial location to be a complex or multidimensional stimulus because the location can be encoded on the basis of various distance and direction cues.

Recently, Lewis and Johnston (1999) demonstrated a peak shift effect in humans after discrimination training involving faces, another complex stimulus dimension. The processes by which people discriminate between and generalize among faces has recently received considerable experimental attention (Leopold, O'Toole, Vetter, & Blanz, 2001), with modern graphic techniques transforming faces into a series forming a complex dimension. Lewis and Johnston (1999) created a series of morphed faces that varied from a unique face to an average face (created by morphing a set of unique faces) and then trained participants to discriminate between two faces that differed along a range from the average to a unique face. Interestingly, when participants were subsequently tested with a range of faces, a peak shift effect was found.

The occurrence of a peak shift effect in face recognition is interesting because it suggests that a process as complex as facial recognition is subject to some of the same processes that underlie discrimination and generalization with much simpler stimulus dimensions. Moreover, as discussed by Lewis and Johnston (1999), the peak shift effect may be related to the caricature effect, which is the finding that caricatures of human faces are often easier to identify than normal (uncaricatured) faces, which are in turn more identifiable than anticaricatures (Lee, Byatt, & Rhodes, 2000). A caricature is further from average than a normal face while an anticaricature is closer to average. Caricatures may thus be more identifiable because they are 'peak shifted' stimuli.

Thus the findings by Lewis and Johnston (1999) suggest a link between research on peak shift effects and the extensive literature on face recognition. Their research also raises some interesting questions. First, is the discrimination of faces also subject to range effects? Thomas's suggestion that most peak-shift like effects in people are due to range effects would lead one to expect that range effects would occur in face discrimination. The study by Lewis and Johnston used the standard procedure for assessing peak-shift effects and did not vary the test range so did not specifically test for range effects. There are, however, reports in the literature of highly transient adaptation effects in face recognition. Webster and MacLin (1999) found after-effects from adaptation with distorted faces, on the judgment of what looks normal.

Leopold et al. (2001) found brief adaptation effects in face identification induced by faces with opposing features and configurations ('antifaces'). These adaptation effects are analogous to range effects. The norm or average, or adaptation level in Thomas's (1993) terminology, becomes shifted toward the adaptation stimuli. However, the time frame is very different, with the adaptation effect reported for faces lasting on the order of seconds, but the range effects influencing judgments that are made several minutes later. Thus, the relationship between range effects and the adaptation after-effects seen with faces is unclear.

Second, does inversion of the face alter generalization and the occurrence of peak shift or range effects? Several studies of face recognition have found that inverted faces are more difficult to recognize than upright faces (see Valentine, 1988), and it has been suggested that recognition of upright faces involves both configuration and featural processing whereas recognition of inverted faces relies primarily on featural processing (e.g., Carey & Diamond, 1977). Thus, we wondered whether the peak-shift effect with upright faces reported by Lewis and Johnston (1999) would also occur with inverted faces. If the peak-shift effect arises from the processing of features, it might be expected to occur in both cases, but, if it arises from the processing of configuration, it may be restricted to upright faces.

Our studies were designed to replicate the findings of Lewis and Johnston (1999) and to extend their work to address the above questions. The first two experiments simultaneously tested for both peak-shift and range effects. Because we failed to find range effects, the third experiment was designed to ensure that we could replicate Thomas's finding of range effects with the simpler stimulus dimension of line tilt. The fourth experiment tested for peak shift and range effects with both upright and inverted faces, using the procedure shown to produce range effects with line tilts.

Experiment 1

In this experiment we explored both range effects and peak shift in people's ability to distinguish one face from other highly similar faces. After training to discriminate between one positive (S+) face and one negative (S−) face, subjects were tested with a range of faces. Two test procedures were used, to promote either peak shift or range effects. To promote peak shift, test faces were interspersed as occasional probes without feedback among ongoing training trials with feedback. To encourage range effects, test faces were presented in a block without feedback. In both procedures, test range was manipulated (within subjects, with different faces). One range was skewed toward the S+ side while the other was skewed toward the S− side.

Methods

Participants

Introductory Psychology students at the University of Alberta ($n = 69$) participated for course credits. Data from five were excluded due to failure to acquire the task (1), a request to terminate early (1) or equipment problems (3).

Design

Participants were assigned randomly to one of 32 conditions (2 per condition). Four factors were varied: (1) type of test procedure (probe or blocked), (2) range of test stimuli (toward the average or toward the unique face), (3) stimuli (two male or two female faces, each gender in two orders), and (4) whether the positive face in the first set was toward the average or unique end of the distribution (the positive face was in the opposite end of the distribution for the second set). Test range (Table 1) was a within subject variable: One of two stimulus faces was tested on a range skewed toward S+; the test range for the other face was skewed toward S−.

Stimuli

Stimuli were 8-bit grayscale images of faces. Models were photographed in front of a light colored, uniform background with overhead lighting. Average faces for each gender were produced from 20 individual faces by a morphing procedure based on custom software written in Matlab (Mathworks). Facial configuration was specified by manually plotting grid co-ordinates of 27 points indicated by anatomical landmarks onto each face. The points were: 4 around each eye, 3 across the nose, 4 around and 1 in the centre of the mouth, 4 around the outline of the head, 3 on each eyebrow and 1 central to the eyebrows (see Fig. 1). Configural information was stored as the co-ordinates of the 27 landmark points. Morphing of configuration was achieved by calculating a weighted average between two co-ordinate sets.

Each set of landmark co-ordinates defines a triangular mesh linking specified points on each face. Correspondence between pixels in different images was defined relative to facial configuration. Pixels within each original image were mapped onto the corresponding position within the corresponding triangular region of the target configuration. Grayscale values from the original images were then averaged pixel-by-pixel, using the same weighting.

In all, 41 faces were created in each series, with face 41 being the average face, and face 1 the original unique face. The remaining faces within each series were formed by constructing weighted morphs between the average face and the original unique face. For face 2, the weighting was 2.5% average face and 97.5% original face. For face 3, the respective weightings were 5.0 and 95.0%. The weights for subsequent faces differed in increments of 2.5%, up to 97.5% average face and 2.5% original face for face 40. S+ and S− were faces 15 and 21 (counterbalanced). The faces used as S+ and S− are shown in Fig. 2. The full range of stimuli may be found on the web (http://www.psych.usyd.edu.au/staff/joelp/Pics4_Marcia/).

Procedure

Participants viewed stimuli on a 17-in monitor at approximately 70 cm distance. The face stimuli measured approximately 6 cm wide by 8.5 cm high. Responses were recorded via a two-button response box. The experiment began with the presentation of the S+ face for the first set. Participants were given a name for the face (Person X, being either Bill, Tom, Mary, or Susan) and told to press the YES button whenever they saw Person X. The S− face was then presented. Participants were told that this was someone who looked similar to but was not Person X and that whenever they saw

Table 1
Stimuli used on tests in the experiments

Experiment 1														
Test faces	31	29	27	25	23	21	19	17	15	13	11	9	7	5
S+ range	S + 5	S + 4	S + 3	S + 2	S + 1	S+	BS+	BS−	S−	S − 1	S − 2			
S− range				S + 2	S + 1	S+	BS+	BS−	S−	S − 1	S − 2	S − 3	S − 4	S − 5
Experiment 2														
Test faces	27	25	23	21	19	17	15	13	11	9				
All participants	S + 3	S + 2	S + 1	S+	BS+	BS−	S−	S − 1	S − 2	S − 3				
Experiment 3														
Test line tilts	20	25	30	35	40	45	50	55	60	65	70			
Wide range	S + 7	S + 6	S + 5	S + 4	S + 3	S + 2	S + 1	S+	BS+	BS−	S−			
Narrow range					S + 3	S + 2	S + 1	S+	BS+	BS−	S−			
Experiment 4														
Test faces	39	36	33	30	27	24	21	18	15	12	9			
Wide range	S + 7	S + 6	S + 5	S + 4	S + 3	S + 2	S + 1	S+	BS+	BS−	S−			
Narrow range					S + 3	S + 2	S + 1	S+	BS+	BS−	S−			

In Experiments 1 and 2, stimuli and ranges were counterbalanced, so that for some participants, Face 15 was S+ and Face 21 was S−. Stimuli were not counterbalanced in Experiments 3 and 4, replicating procedures used by Thomas et al. (1991).

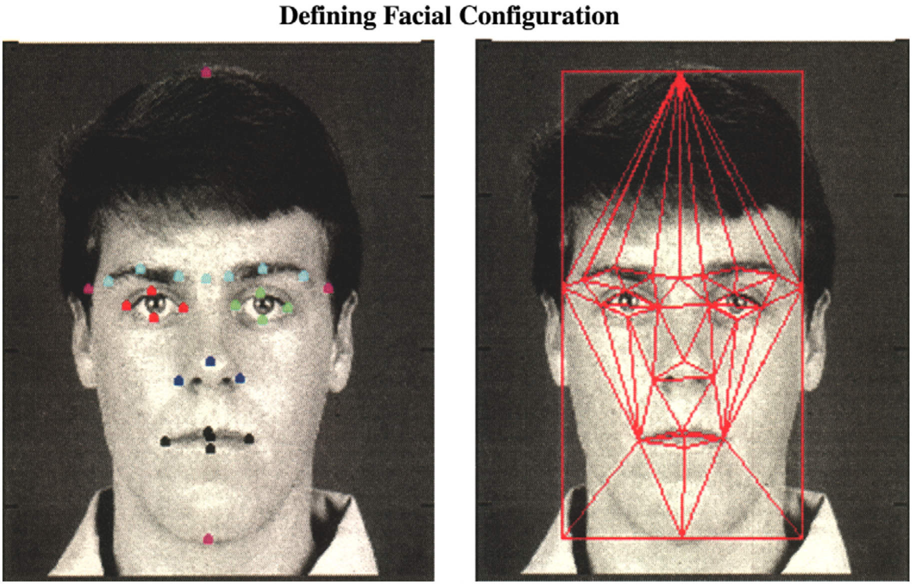

Fig. 1. Example of a face showing the 27 anatomical landmarks used for the morphing procedure. See text for details.

someone who was not Person X they were to press the NO button. The participant then received a series of trials in which the S+ or S− face was presented. Correct choices (pressing YES for the S+ face and NO for the S− face) were followed by the auditory feedback "Correct" and incorrect choices were followed by the auditory feedback "Wrong." Each participant received a minimum of 2 blocks of 10 training trials (5 S+ faces and 5 S− faces presented in random order within each block) and training continued until accuracy on the preceding block equaled or exceeded 80% correct.

Tests were then given. The blocked group received 16 blocks of 11 tests. Each block contained one presentation of each test face in random order. No feedback was given. For the probe condition, the test phase consisted of 4 blocks of 45 trials. Each block had 17 S+ trials and 17 S− trials (with feedback), and 11 tests (one of each test face, without feedback). These tests were preceded with the label "no feedback trial."

A second set with a different series of faces, and a different name, then followed. The new faces were the same gender as the first set but the location of the S+ and S− in the distribution were reversed (e.g., if the S+ was on the unique end of the distribution in the first set, then it was on the average end of the distribution for the second set). The range of faces presented during testing (i.e., skewed toward the S+ or the S− side of the distribution) was also reversed for the second half.

Data analysis

Face stimuli were first recoded in relation to their distance from S+ and S− (from S + 5 to S − 5). The two values between the S+ and S− were denoted as BS+ (the one closer to the S+) and BS− (the one closer to the S−). Inferential statistics were done

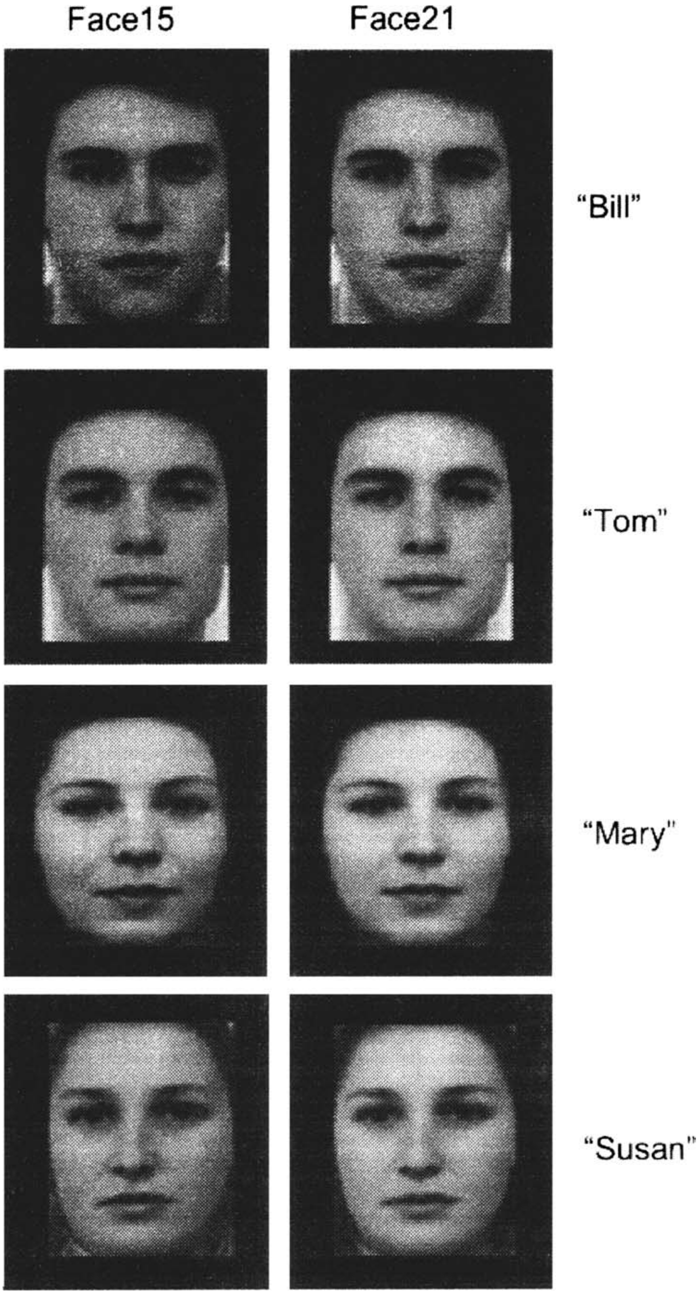

Fig. 2. Examples of face stimuli used in the experiments. See text for the morphing methods used to produce them. The morphed faces were ordered in a series which ranged from 1 (unique face) to 41 (average face). S15 refers to the 15th stimulus in the series of faces, and S21 refers to the 21st stimulus in this series.

on the range of stimuli common to both test ranges ($S+2$ to $S-2$), separately for blocked and probe test groups. A mixed-model analysis of variance (ANOVA) was run on the proportion of YES responses. Within-subjects variables were test range (skewed toward S+ or S−) and test stimulus ($S+2$ to $S-2$). The between-subjects variables were gender of face stimuli (2), assignment of face stimuli (face 15 or 21) to test ranges (2), and order of test ranges (2). We ignored the counterbalancing of faces within each gender. An area shift or peak shift is indicated by higher responding to $S+1$ and $S+2$ than to BS+ and BS− (an a priori contrast). A range effect is indicated by an interaction of test range and test stimulus. An α level of $p = 0.01$ was adopted throughout.

Results and discussion

Evidence of area shift was found in both methods of testing (Fig. 3), with stimuli on the S+ side receiving more YES responses than stimuli on the S− side. For blocked tests, a test range around S+ produced lower YES responding around S+, but no shift in the peak, contrary to the predictions of adaptation theory.

For probe tests, the mixed ANOVA found only a significant main effect of face stimulus ($F(7,168) = 213.43$). An a priori contrast shows that YES responding was higher for $S+2$ and $S+1$ than for BS+ and BS− ($F = 188.02$), indicating area shift. For blocked tests, the mixed ANOVA found significant main effects of test stimulus ($F(7,168) = 54.42$) and test range ($F(1,24) = 27.95$), with the range

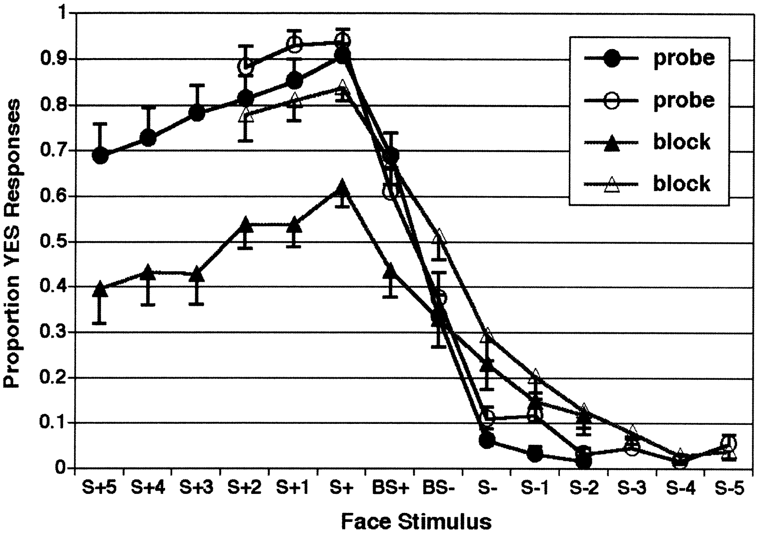

Fig. 3. Proportion of YES responses in Experiment 1 (M ± *SE*) under block and probe testing conditions. Results from two blocks of testing have been combined. Filled symbols indicate test range skewed toward S+ and open symbols indicate test range skewed toward S−. The face stimuli are labeled according to their distance from the positive (S+) and negative (S−) face, with BS indicating faces between the S+ and S−.

centered on S− producing more YES responding, and a significant interaction of test stimulus and range ($F(7, 168) = 2.81$). A priori contrasts were made separately for the two test ranges. For each test range, subjects responded YES more to $S+2$ and $S+1$ than to BS+ and BS− (range around S+: $F = 13.38$, range around S−: $F = 22.40$), indicating area shift.

To test whether the peak differed between the test ranges for blocked tests, a peak was calculated over the range $S+2$ to $S-2$ for each subject in each test range. Integers were assigned to the eight test stimuli, with $8 = S+2$ and $1 = S-2$. A weighted average was calculated over the range $S+2$ to $S-2$, weighted by the proportion of YES answers for each stimulus. The peaks thus calculated did not differ significantly between the two ranges by a paired t test. Thus, the stimulus × range interaction was not caused by the difference in peaks predicted by adaptation theory. Instead it may be the case that with more stimuli on the S+ side (test range centered on S+), subjects were less confident about which specific stimulus was S+, and hence responded YES at lower levels.

In sum, area shift in face perception was found with both testing methods. No significant effects of assignment of face stimuli was found, indicating that area shift was not dependent on whether the positive face was toward the average or unique end of the distribution. A range × stimulus interaction was found for block testing, but this was not a difference in peak between the two test ranges as predicted by adaptation level theory.

Experiment 2

Having failed to find adaptation effects in Experiment 1, Experiment 2 was designed to provide a more conclusive test by using stronger adaptation manipulations. Perhaps the test range manipulation used in Experiment 1 was not strong enough to alter the adaptation level established in training. In Experiment 2, subjects were first trained as in Experiment 1. A 10-min adaptation phase then followed, with one of two extreme ranges of face stimuli being presented. Without further training, a common test phase then followed. Adaptation level theory predicts an adaptation range × test stimulus interaction that stems from different peaks for the different adaptation ranges.

Methods

Participants ($n = 32$) recruited from the same pool were divided into two groups, differing in the range of faces that they were exposed to in an adaptation phase (S+ side or S− side). Each participant went through three phases and was only exposed to one face series. Phase 1 was S+/S− discrimination training and followed the training methods of Experiment 1. Phase 2 was an adaptation phase. Each participant saw 147 faces (21 different faces, each shown seven times) from the series of faces from which S+ and S− were chosen. Participants were exposed to every face within the range of either faces 35–15 or faces 21–1. Each face was presented for 2500 ms, with a 500 ms interstimulus interval of blank screen. Occasionally (seven trials in total), a face of the opposite gender was presented. The task in this phase was to press the space bar whenever this occurred. Phase 3 was a test phase. Participants were shown faces ranging from S + 3 to S − 3 (10 in all, Table 1), in a random order. Their task was to indicate whether each face was S+. Ten rounds of tests (thus 10 for each stimulus) were given. For analysis, adaptation ranges were classified as S+ side or S− side. We analyzed the first 1, 3, 5, and 10 tests to look for possible range effects, but since we found none, only results from the first five tests are presented. Data analysis had between-subjects factors of adaptation range (2), S+ (face 15 or face 21), and gender of face. The within-subjects factor was test stimulus, ranging from S + 3 to S − 3 (10).

Results and discussion

Results from the first five tests (Fig. 4) show an area shift, but adaptation range did not seem to affect the pattern of results. The mixed ANOVA revealed a significant main effect of test stimulus ($F(9, 216) = 28.77$), as well as significant interactions of stimulus × face gender ($F(9, 216) = 2.58$) and stimulus × S+ × face gender ($F(9, 216) = 3.96$). Both of these interactions were caused by the fact that the usual area shift was not found for female faces with S+ = face 15; in fact the pattern was reversed, with more YES responding to BS+, BS−, and S− than to S + 1, S + 2, and S + 3. We have no explanation for this finding. No other effects were significant. In particular, the stimulus × range effect was not close to significance, giving no evidence

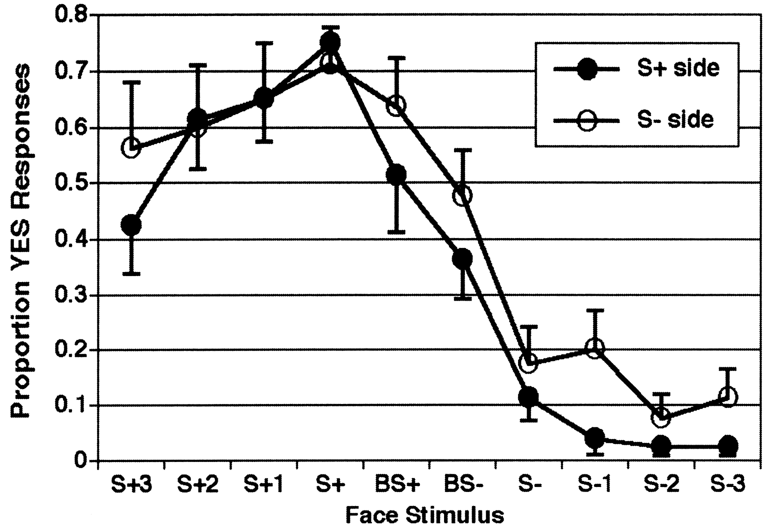

Fig. 4. Proportion of YES responses in Experiment 2 of participants adapted to faces on the S+ side or the S− side (M ± *SE*). The face stimuli are labeled according to their distance from the positive (S+) and negative (S−) face, with BS indicating faces between the S+ and S−.

for range effects. Turning to the stimulus main effect, an a priori contrast of S + 2 and S + 1 against BS+ and BS− came out significant ($F = 7.09$), indicating an area shift.

On the possibility that a transient range effect (significant stimulus × range interaction) might be found, we looked for the effect by also examining the first 1, 3, and 10 tests. For the analysis with the first test only, the stimulus factor was reduced to 2 levels by pooling. Number of YESes to S + 3, S + 2, and S + 1 formed the S+ side, while number of YESes to BS+, BS−, and S− formed the S− side. For the first 3 and 10 tests, the standard ANOVA with 10 levels of stimulus was performed. In no case was the stimulus × range interaction close to significance. In short, the data again provide some evidence for area shift, but no evidence at all for adaptation effects.

Experiments 3a and 3b

It looks like line tilts and brightness are subject to range effects but not faces. There are, however, some differences besides the nature of the stimuli between our experiments and those which have demonstrated range effects. In particular, our stimuli were presented on a computer monitor, while Thomas et al. (1991) and most earlier research presented visual stimuli on a tachistoscope or a screen. We thus attempted to replicate the procedure that Thomas et al. (1991) used in their Experiment 2, to ensure that we could get range effects with line stimuli presented on a computer monitor.

Methods

The methods used were highly similar to those used in Experiment 2 of Thomas et al. (1991). The stimuli consisted of a black line, approximately 45 mm long and 2.5 mm wide, presented against a white background. The line was centered within a black circle, 8 cm in diameter, with a line width of approximately 1 mm. The orientation of the line ranged from tilt angles of 20° to 70° counterclockwise from horizontal. S+ was always 55°, and S− was always 70°. In both experiments, two groups were tested with two different ranges, a narrow group from S + 3 to S−, and a wide group from S + 7 to S− (Table 1). As in the Thomas et al. study, participants were instructed that they would be trying to identify a target line, which they would see on the first trial. On subsequent trials, they were to determine whether the presented line was the same or different from the target line. Whenever the target line was presented, they should press the YES key while the stimulus was on. When the presented line did not match the target line, they should withhold responding. The stimuli stayed on the screen for 1 s, and, on initial trials, the participants received feedback. Participants heard "correct" if they pressed the YES key within 1 s on S+ presentations and if they withheld responding on S− presentations; they heard "wrong" if they withheld responding on S+ trials or responded on S− trials. Testing began after a minimum of 16 training trials if 6 or more correct responses occurred in the last 8 trials. Testing included 77 trials without feedback. The narrow group received 11 tests with each of the 7 line tilts, and the wide group received 7 tests with each of the 11 line tilts. This number of trials was similar to that used by Thomas et al., who gave 63 tests to the narrow group and 66 tests to the wide group, and thus should provide adequate exposure for adaptation effects. All trials, during both training and testing, were separated by a 5-s ITI during which the screen was black.

In Experiment 3a, participants were not monitored during experimentation. Of 40 participants, 12 were excluded for reporting the use of strategies to mark S+ (such as touching the screen to leave a smudge or holding their finger up to the screen), leaving 28 who contributed data. Including the excluded participants did not change the pattern of results. Experiment 3b was a replication with 15 subjects in the narrow group and 16 in the wide group. Subjects were told not to point to or touch the monitor, and they were watched during experimentation. Range effects in each case are assessed by a test for condition × stimulus interaction.

Results and discussion

The results showed an area shift but no range effects in Experiment 3a (Fig. 5A), but clear range effects in Experiment 3b (Fig. 5B). We again analyzed the stimulus range common to both groups, from S + 3 to S−. In Experiment 3a, the mixed ANOVA revealed only a significant main effect of stimulus ($F(6, 156) = 25.20$). Comparing S + 2 and S + 1 against BS+ and BS− in an a priori contrast revealed a significant effect ($F = 50.61$), indicating area shift. The stimulus × range interaction was not close to significance.

The ANOVA on Experiment 3b, however, revealed both a significant main effect of stimulus ($F(6, 174) = 9.57$) and a significant interaction of condition × stimulus ($F(6, 174) = 4.30$). The nearly symmetrical function about S+ for the narrow group indicates a lack of area shift. These results, which resemble those found by Thomas et al. (1991), indicate range effects but not peak shift. The striking difference between the results of Experiments 3a and 3b indicates that instructions and/or monitoring of participants can dramatically alter the generalization functions seen for tilt stimuli presented on a computer screen.

Experiments 4a, 4b, and 4c

Experiment 3b showed that with proper instructions and monitoring, range effects will occur for line tilt stimuli on a computer screen. Experiments 4a, 4b, and 4c provided additional tests for range effects with face stimuli, this time using a procedure that was designed to be as similar as possible to that used in Experiment 3b. A single series of male faces served as stimuli. After discrimination training, one group was tested without feedback on the range S + 3 to S−. The other group was tested on the range S + 7 to S−.

These experiments also examined the effect of stimulus orientation on generalization functions for faces. Many studies have found that inverted faces are more difficult to recognize than upright faces, a difference referred to as the "face inversion effect" (for review see Valentine, 1988). In addition, evidence suggests that perception of spatial relationship information is impaired in inverted faces, suggesting a qualitative difference in the processing of upright and inverted faces (Murray, Young, & Rhodes, 2000). Cabeza and Kato (2000) examined the effect of face

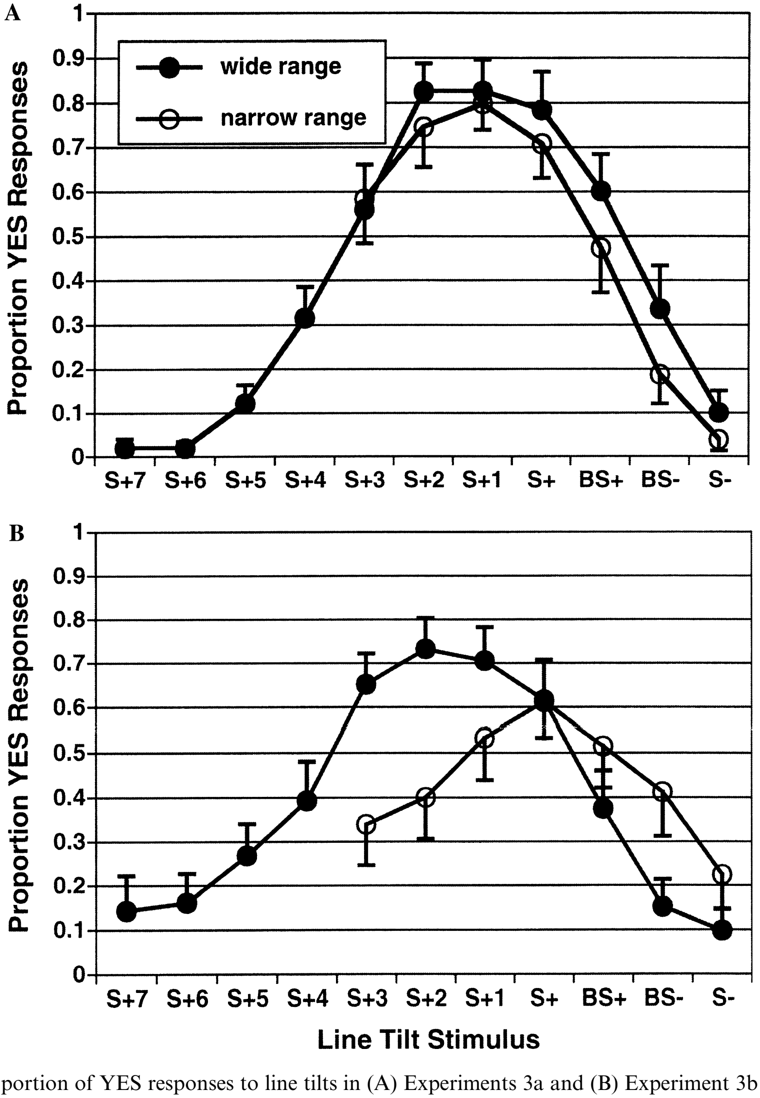

Fig. 5. Proportion of YES responses to line tilts in (A) Experiments 3a and (B) Experiment 3b (M ± *SE*). Participants in Experiment 3b (but not participants in Experiment 3a) were told not to use hands or other props to mark the computer screen on which stimuli appeared. The line orientations are labeled according to their distance from the positive (S+) and negative (S−) orientations, with BS indicating orientations between the S+ and S−.

inversion on the false recognition of two kinds of prototypes. Configural prototypes were formed by morphing the shape and colour/texture information in a set of four faces. Featural prototypes were created by averaging the features in the set of four faces. Participants studied the original faces in the set and then were tested for false memory of the unstudied prototypes. The interesting result was that false recognition of configural prototypes was eliminated by inversion of the faces but false recognition of featural prototypes was not. These results are consistent with a dual-mode view proposed by Carey and Diamond (1977) in which processing of upright faces is thought to be based on both configuration (spatial relationship) and features (components), whereas processing of inverted faces is based primarily on component information. Rhodes and Tremewan (1994) manipulated both orientation and caricature. Inverted faces were reliably more difficult to recognize than upright faces, and anticaricatures were more difficult to recognize than normal or caricature faces. In this study, caricature faces were not reliably easier to recognize than normal faces. Interestingly, orientation and caricature effects did not interact. In light of these findings, it would be interesting to determine whether peak shift or range effects would be found with inverted face stimuli.

In Experiment 4a, standard upright faces were used during both the training and test phase. In Experiment 4b, subjects were assigned to one of two conditions that differed in the orientation of the face during training and testing. Condition Upright–Upright (U–U) was identical to Experiment 4a and presented upright faces during training and testing. In Condition Inverted–Inverted (I–I), the faces were upside down during both training and testing. In Experiment 4c, faces were flipped between training and testing. In condition Upright–Inverted (U–I) the participants were trained with upright faces and tested with upside down faces; whereas in condition Inverted–Upright (I–U), they were trained with upside down faces and tested with upright ones.

Methods

For stimuli, S+ was face 18 and S− was face 9 in the series of "Bill" faces (see Fig. 6). These values provided stimuli that were a bit easier to discriminate and thus made the discrimination task more similar to the line discrimination in Experiment 3. In addition, the stimuli were slightly smaller (4.5 cm wide by 6 cm high) than in the first two experiments, making them closer in size to the line stimuli presented in Experiment 3. The instructions and procedure were identical to those of Experiment 3b, except that

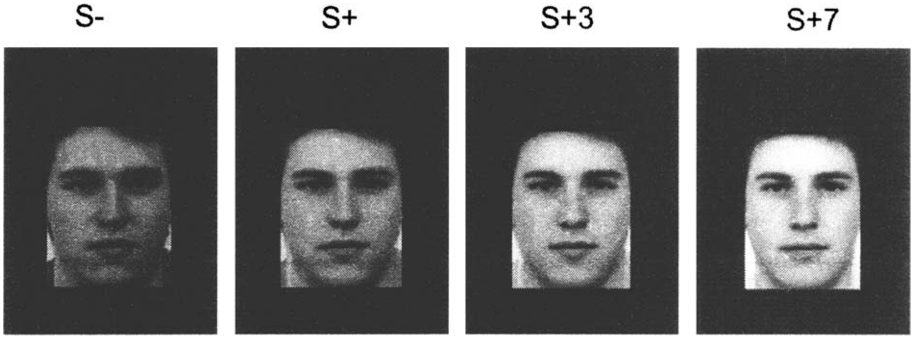

Fig. 6. Face stimuli used as the positive (S+) and negative (S−) stimuli *n* Experiment 4 (shown in upright orientation only). S + 3 represents the top end of the test range for the narrow group and S + 7 represents the top end of the test range for the wide group.

faces were substituted for the lines, and in Experiments 4b and 4c all subjects were informed that the faces would be seen either in the upright or upside down orientation. Unlike in the first two experiments, no names were provided for the faces. Participants were again recruited from the Psychology subject pool at the University of Alberta. In Experiment 4a, the two groups (wide range and narrow range) each had 18 participants. In Experiment 4b, participants were assigned at random to the U–U ($N = 24$) and I–I conditions ($N = 41$). (Fewer participants were assigned to U–U than to I–I because U–U was a replication of Experiment 4a.) In Experiment 4c, participants were assigned to the U–I or I–U conditions ($N = 17$ each). Within each condition, participants were assigned to the wide and narrow range in as equal numbers as possible.

Participants were trained with the S+ (face 18) and S− (face 9) according to the procedure described for Experiments 3a and 3b. During testing, participants in the wide groups were presented with faces S + 7 to S−; participants in the narrow groups were presented with faces S + 3 to S− (Table 1).

Results and discussion

The results of Experiment 4a (Fig. 7) again showed clear evidence of area shift, but no range effects. The mixed ANOVA was conducted on the range common to both groups (S + 3 to S−). It revealed only a significant main effect of stimulus ($F(6, 204) = 26.58$). The condition × stimulus interaction was not near significance. An a priori comparison contrasting S + 2 and S + 1 against BS+ and BS− was significant ($F = 92.01$), indicating area shift.

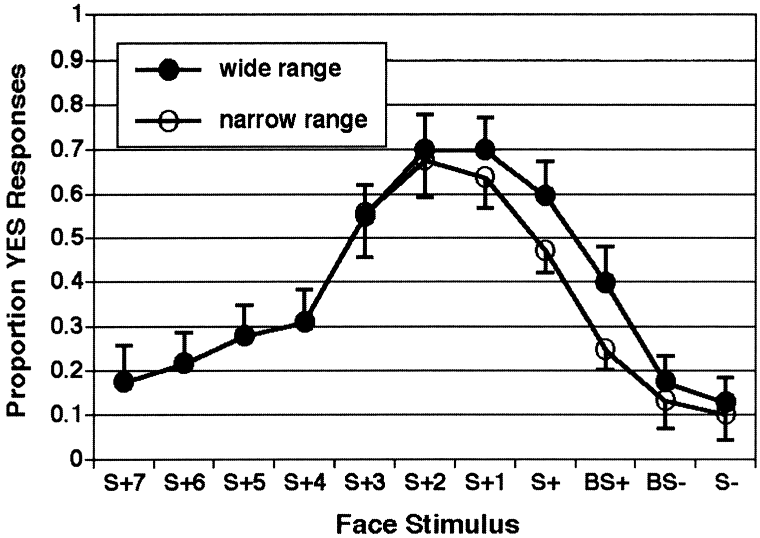

Fig. 7. Proportion of YES responses to face stimuli presented in Experiment 4a (M ± *SE*), with upright faces in training and testing. The face stimuli are labeled according to their distance from the positive (S+) and negative (S−) face, with BS indicating faces between the S+ and S−.

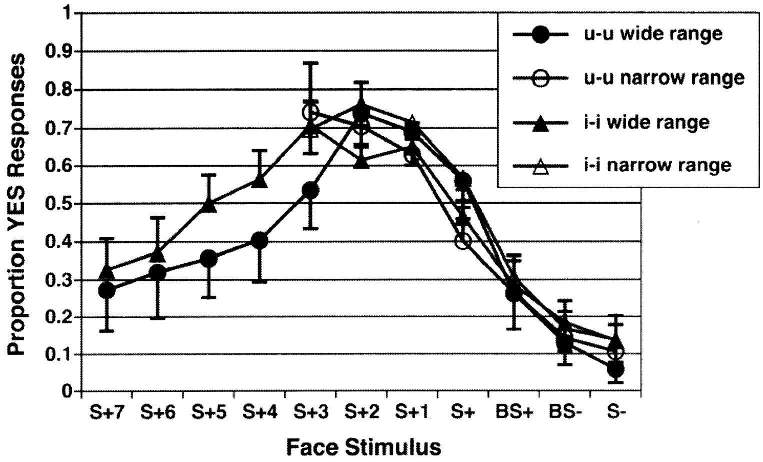

Fig. 8. Proportion of YES responses to face stimuli presented in Experiment 4b (M ± *SE*) for conditions with upright faces in training and testing (u–u), and with inverted faces in training and testing (i–i). The face stimuli are labeled according to their distance from the positive (S+) and negative (S−) face, with BS indicating faces between the S+ and S−.

In Experiment 4b (Fig. 8), clear peak shift effects were found, but no range effects. The mixed ANOVA conducted on the range S + 3 to S− revealed only a main effect of stimulus ($F(6, 366) = 56.73$). Results were similar for U–U and I–I conditions, and no other effects were close to significance. An a priori comparison contrasting S + 2 and S + 1 against BS+ and BS− was significant ($F = 190.18$). In addition, a (non-orthogonal) comparison of S + 3, S + 2, and S + 1 against S+ was significant ($F = 22.01$). Together, these contrasts indicate peak shift.

In Experiment 4c (Fig. 9), clear peak shift effects were found. The peaks were at the extreme left for both narrow and wide ranges, and results were similar for U–I and I–U conditions. The mixed ANOVA conducted on the range S + 3 to S− revealed a significant main effect of stimulus ($F(6, 180) = 35.00$). The range main effect and the stimulus × range interaction were right at the significance level (range: $F(1, 30) = 7.44$, $p = 0.011$; stimulus × range: $F(6, 180) = 2.94$, $p = 0.009$). Because the peaks of both distributions were at the left end (Fig. 9), this interaction did not represent a true range effect (i.e., a difference in peak between the ranges) but instead seemed to reflect less confidence in the wide-range conditions in responding "YES" at the S+ end.

Separate ANOVAs were then run on the wide and narrow ranges in order to do comparisons between stimuli. For the wide range, an a priori comparison of S + 2 and S + 1 against BS+ and BS− was significant ($F = 18.10$). A comparison of S + 3, S + 2, and S + 1 against S+ was also significant ($F = 9.18$). For the narrow range, a comparison of S + 2 and S + 1 against BS+ and BS− was significant ($F = 86.48$). A comparison of S + 3, S + 2, and S + 1 against S+ was also significant ($F = 17.57$). These contrasts indicate clear peak shifts.

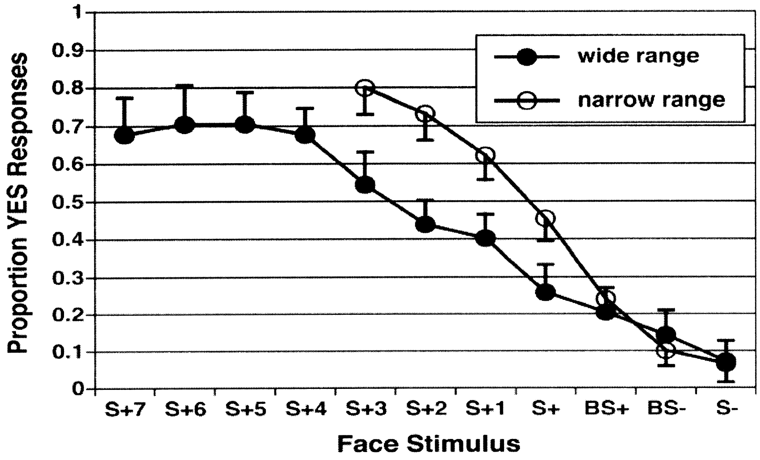

Fig. 9. Proportion of YES responses to face stimuli presented in Experiment 4c (M ± *SE*). Conditions with upright faces in training and inverted faces in testing, and with inverted faces in training and upright faces in testing, have been combined. The face stimuli are labeled according to their distance from the positive (S+) and negative (S−) face, with BS indicating faces between the S+ and S−.

General discussion

These experiments tested peak shift and range effects in the recognition of faces and line tilts. Participants were trained to recognize one stimulus as S+, and to discriminate a similar stimulus as S−. In Experiments 1, 3, and 4, they were then tested with different ranges of stimuli without feedback. In Experiment 2, an adaptation phase of exposure to different ranges of faces was interposed between training and testing. For faces, peak shift or area shift effects were consistently found, but range effects were not found. For line tilts, a range effect was found, but no independent area or peak shift was found when subjects were prevented from using their hands or other props to mark or indicate locations on the monitor. When subjects were not monitored, some reported using the strategy of physically indicating on the screen where S+ was. Most likely, many subjects did this, with some of them not reporting it. Under these conditions, generalization functions for line tilts showed area shift but no range effects. Interestingly, we did not find peak or area shift and range effects together in any experiment.

The peak shift and area shift effects replicated the results of Cheng and Spetch (2002) on spatial location. They show that with some procedures and stimuli, area shift can be consistently found in humans. Because our experiments controlled for range effects, the results with faces also indicate that the peak shift effect obtained earlier with faces (Lewis & Johnston, 1999) is indeed peak shift and not a range effect.

The similar peak-shift effect found for upright and inverted faces in Experiment 4b is interesting in light of the evidence suggesting that upright faces are processed differently than inverted faces (Carey & Diamond, 1977). Although the literature suggests that inverted faces are more difficult to recognize than upright faces (e.g., Valentine, 1988) and that perception of configural information is impaired by inversion of faces (Murray et al., 2000), our results indicate that peak shift effects occur equally with both upright and inverted faces. This implies that the peak shift effect is unlikely to reflect the processing of configural information. Moreover, the independence of the peak shift effect from orientation of the face stimuli is congruent with the finding by Rhodes and Tremewan (1994) that caricature effects and face orientation do not interact.

Perceptual after-effects of viewing distorted faces also occur with both upright and inverted faces (Watson & Clifford, 2003; Webster & MacLin, 1999; Zhao & Chubb, 2001). This perceptual after-effect is substantially weaker if the orientation of the face changes between adaptation and test (Webster & MacLin, 1999), particularly if the adaptor is inverted and the test is upright (Watson & Clifford, 2003). Interestingly, we observed exceptionally strong peak-shift effects in our groups that experienced a shift in orientation between training and test (Experiment 4c). This finding is consistent with the notion that the discrimination training procedure that induces peak shift may operate on processing of component rather than configural information.

The results of Experiment 3 indicate that methods of testing can influence whether range effects are found with line tilts. The use of physical props seems to abolish range effects, but interestingly, area shift was still found. While range effects have been frequently found (Thomas, 1993), they prove malleable. Even 'mental' props established by instructing participants to imagine line tilt stimuli on a clock face (Thomas & Thomas, 1974) or verbal labelling of wavelength stimuli (Tomie & Thomas, 1974) can abolish or alter range effects.

Even with a paradigm that produced robust range effects in line tilts, we failed to see range effects in faces. We therefore conclude that range effects are difficult to find in face recognition. In this case, type of stimulus seems to matter more than method of testing. Although Leopold et al. (2001) reported adaptation effects in face identification, which are analogous to range effects, their methods were very different from ours, and the time span that they examined was much briefer. Indeed, they report adaptation effects induced by a single 5-s exposure to an 'antiface' as measured by the perception of a brief (200-ms) presentation of a subsequent test face. They also report that in pilot work, longer test face presentations led to confused percepts because "the adaptation effect was fleeting, disappearing in a few hundred milliseconds" (Leopold et al., 2001, p. 94). Thus, the adaptation effects that have been found in face perception appear to occur under very different temporal parameters than the range effects that occur in line discriminations and likely reflect different processes.

We end by suggesting speculatively that the more complex and multidimensional the stimuli, the more likely area shift rather than range effects will be found. The typical stimuli showing range effects form simple dimensions, for example, line tilts (Thomas et al., 1991), lifted weights (Hébert, Origlio, & McGuirk, 1972), or gray levels (Hébert, 1970). On the other hand, Cheng and Spetch (2002) found peak shift and area shift with spatial position on a monitor, but no range effects, and our results show peak and area shift but no range effects in face recognition. These stimuli exemplify multiple dimensions and many descriptions.

Acknowledgments

This research was supported by an NSERC research grant to M.L.S., a Visiting Scholar grant from Macquarie University for a visit by M.L.S., and a Queen Elizabeth II Fellowship from the Australian Research Council to CWGC. We are grateful to Joel Pearson for help in generating the face stimuli and to Blair Dent, Emily Gray, Anne Ferrey, and Angela Nguyen for their assistance and contributions to various aspects of the research.

References

Cabeza, R., & Kato, T. (2000). Features are also important: Contributions of featural and configural processing to face recognition. *Psychological Science, 11*, 429–433.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195*, 312–314.

Cheng, K., & Spetch, M. L. (2002). Spatial generalization and peak shift in humans. *Learning and Motivation, 33*, 358–389.

Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour, 66*, 15–36.

Guttman, N., & Kalish, H. I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology, 51*, 79–88.

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

Hébert, J. A. (1970). Context effects in the generalization of a successive conditional discrimination in humans. *Canadian Journal of Psychology, 24*, 271–275.

Hébert, J. A., Origlio, D. P., & McGuirk, F. D., III (1972). Training and testing effects in the generalization of a voluntary response. *Psychonomic Science, 26*, 209–210.

Helson, H. (1964). *Adaptation level theory*. New York: Harper & Row.

Honig, W. K., & Stewart, K. E. (1993). Relative numerosity as a dimension of stimulus control: The peak shift. *Animal Learning & Behavior, 21*, 346–354.

Honig, W. K., & Urcuioli, P. J. (1981). The legacy of Guttman and Kalish (1956): Twenty-five years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior, 36*, 405–445.

Lee, K., Byatt, G., & Rhodes, G. (2000). Caricature effects, distinctiveness, and identification. *Psychological Science, 11*, 379–385.

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

Lewis, M. B., & Johnston, R. A. (1999). Are caricatures special? Evidence of peak-shift in face recognition. *European Journal of Cognitive Psychology, 11*, 105–117.

Murray, J. E., Young, E., & Rhodes, G. (2000). Revisiting the perception of upside-down faces. *Psychological Science, 11*, 492–496.

Purtle, R. B. (1973). Peak shift: A review. *Psychological Bulletin, 80*, 408–421.

Rilling, M. (1977). Stimulus control and inhibitory processes. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 432–480). Englewood Cliffs, NJ: Prentice-Hall.

Rhodes, G., & Tremewan, T. (1994). Understanding face recognition: Caricature effects, inversion, and the homogeneity problem. *Visual Cognition, 1*, 275–311.

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

Thomas, D. R. (1974). The role of adaptation-level in stimulus generalization. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 91–145). San Diego, CA: Academic Press.

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

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

Thomas, D. R., & Thomas, D. H. (1974). Stimulus labeling, adaptation-level, and the central tendency shift. *Journal of Experimental Psychology, 103*, 896–899.

Tomie, A., & Thomas, D. R. (1974). Adaptation level as a factor in human wavelength generalization. *Journal of Experimental Psychology, 103*, 29–36.

Valentine, T. (1988). Upside-down faces. A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*, 471–491.

Watson, T. L., & Clifford, C. W. G. (2003). Pulling faces: an investigation of the face distortion aftereffect. *Perception, 32*, 1109–1116.

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

Zhao, L., & Chubb, C. (2001). The size-tuning of the face-distortion after-effect. *Vision Research, 41*, 2979–2994.