Differentiation for Novel Dimensions

Stephen A. Hockema (shockema@indiana.edu)
Psychology and Cognitive Science, Indiana University
1101 E. Tenth Street, Bloomington, IN 47405 USA

Mark R. Blair (mrblair@indiana.edu)
Psychology and Cognitive Science, Indiana University
1101 E. Tenth Street, Bloomington, IN 47405 USA

Robert L. Goldstone (rgoldsto@indiana.edu)
Psychology and Cognitive Science, Indiana University
1101 E. Tenth Street, Bloomington, IN 47405 USA

Abstract

Two experiments are reported that provide evidence for perceptual differentiation between a pair of novel, integral dimensions, in contrast to previous attempts that failed to differentiate these same two dimensions (Op de Beeck, Wagemans, & Vogels, 2003). In Experiment 1, an acquired distinctiveness effect was created on the category-relevant dimension through a categorization training regimen that gradually increased in difficulty. Response times for correct trials were faster across the category boundary. This effect was replicated in Experiment 2 using a new training procedure where participants had to predict category boundaries while watching an animation in which shapes transformed along the category-relevant dimension. Furthermore, the accuracy results of Experiment 2 also indicated that discriminability was changed on the category-relevant dimension relative to the irrelevant dimension across the entire range of the dimension, not just at the category boundary.

Keywords: visual perception; perceptual learning; categorical perception; selective attention

Introduction

Our ability to categorize depends on our ability to selectively attend to dimensions and unitary features. To tell a cat from a dog, we will probably pay attention to certain morphological (e.g., ear and nose shape) and behavioral (e.g., barking vs. meowing, tail wagging) features while deemphasizing others (e.g., the fact they both have eyes); to tell a Beagle from a Basset Hound we will probably also attend to dimensions such as height, weight, color and ear length, while perhaps ignoring the shininess of their fur. This ability to analyze our perceptions in terms of features and dimensions appears to improve gradually through development (Smith, 1989). Furthermore, as we become more expert in a domain, we improve our ability to perceive and appreciate previously difficult domain-relevant distinctions (Burns & Shepp, 1988). Indeed, much work has shown that our perceptual systems are not fixed—they can be tuned by our experiences in the world (see Goldstone (1998) for a summary). But just how malleable are they? This paper examines processes by which experience might change the visual dimensions along which we perceive the world.

In previous work, categorization training has been shown to affect the dimensions underlying visual perception. For easily separable dimension pairs[1], such as size and brightness (Goldstone, 1994) or aspect ratio and curvature (Op de Beeck et al., 2003), the discriminability of either dimension was independently improved if it was relevant for categorization. For more integral dimension pairs, such as color saturation and brightness, it has been shown that, while categorization training helped improve the discriminability of both dimensions, there was more improvement in the discriminability of the category-relevant dimension (Goldstone, 1994). Furthermore, Goldstone and Steyvers (2001) went on to show with more abstract dimensions (created by arbitrary combinations of images of faces) that after categorization training where only one dimension is relevant, it is easier to learn a category rule based only on the irrelevant dimension, but not on a combination of the relevant and irrelevant dimensions. These findings were taken to support a hypothesis that category training could lead to dimension *differentiation*, increasing the separability of integral dimensions.

However, recently Op de Beeck et al. (2003) challenged this conclusion with a set of novel, perceptually-integral dimensions. With these dimensions, there was no evidence that categorization led to a greater improvement in discriminability along the category-relevant dimension relative to the irrelevant dimension. From this, they speculated that category learning is only capable of biasing already separable dimensions, but not of making integral dimensions more separable. In other words, only if we can already separate two dimensions can we learn to selectively attend to one of them at the expense of the other. However, Op de Beeck et al. (2003) only examined one method of category learning. It remains open whether different training methods might lead to dimension differentiation with this challenging stimulus set. The goal of the present experiments was to test this possibility.

This question – can the degree of separability of perceptually-integral dimensions be changed – is an important one because it bears on theories of conceptual development. Are our concepts grounded on a universal set of perceptual primitives? And, if not, how do our experiences change these building blocks? The present studies shed light on these questions by utilizing novel shape dimensions that are unlikely to correspond to any independent built-in perceptual primitives.

[1] We say two dimensions are *separable* when we have the ability to attend to one dimension without attending to the other (cf. Garner & Felfoldy, 1970). To the degree that variation along one dimension influences discrimination along the other, we say the two dimensions are *integral*. This is not a binary classification; the separability/integrality of two dimensions is a matter of degree.

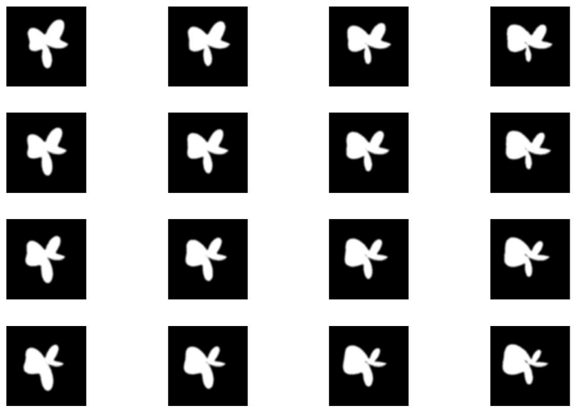

Figure 1: RFC shapes used for discrimination tests in Experiments 1 and 2. The "horizontal" dimension varied along the rows; the "vertical" dimension varied along the columns.

The Present Studies

In Experiment 1, a standard categorization training procedure with feedback was used . The main difference between this procedure and that employed by Op de Beeck et al. (2003) was that the order of training trials was carefully controlled to start with easy categorizations—shapes far from the category boundary—and gradually increase in difficulty to include shapes nearer to the boundary. In Experiment 2, a nonstandard type of training was employed, inspired by Hockema (2004). Subjects were exposed to animations showing the gradual transformation of a shape from one category to the other along the category-relevant dimension and taught to predict the point when the category boundary was crossed. In both experiments, the training phase was followed by a test phase to assess the discriminability of both the category-relevant and category-irrelevant dimension.

The stimuli used by Op de Beeck et al. (2003) were created by combining seven sinusoidal functions (each with three parameters: frequency, phase, and amplitude), referred to individually as radial frequency components (RFCs) into a single, complex curve and then bending these to create closed contours. While five of the seven RFCs remained fixed, two were chosen to have their amplitudes varied to define a two-dimensional space of "blobs". Op de Beeck et al. (2003) showed that these dimensions are relatively integral. Examples are shown in Figure 1. One of these dimensions was arbitrarily chosen to be referred to as the "Horizontal" dimension and the other as the "Vertical" dimension. (This labeling was consistent across experiments.) All stimuli were created using Matlab code provided by Op de Beeck as used in (Op de Beeck et al., 2003), where more details can be found about their specification and characteristics. (The only difference between what is reported there and the stimuli used here is that our stimuli were slightly blurred using a Gaussian filter, in order to remove the pixelated edges.)

Experiment 1

The hypothesis of this study was that training on a categorization task in which one of the RFC dimensions was relevant and one was not would lead to selective sensitization of the relevant dimension, allowing participants to discriminate along that dimension either more rapidly, more accurately, or both. To test this hypothesis we trained participants to discriminate between two categories of RFC stimuli, such that only one of the two varying dimensions was relevant for the task. Participants then performed a delayed-match-to-sample task, wherein they decided which of a displayed pair of stimuli was displayed immediately prior. For this discrimination task, trials were grouped into four types: within-category discriminations in the first row/column of stimuli parallel to the category boundary (W_1), within-category discriminations in the second, more distant, row/column of stimuli parallel to the category boundary (W_2), within-category discriminations along the relevant dimension between the first and second rows (W_{12}), and boundary discriminations (B), where the relevant-dimension distractor was in the other category. (See Figure 2.)

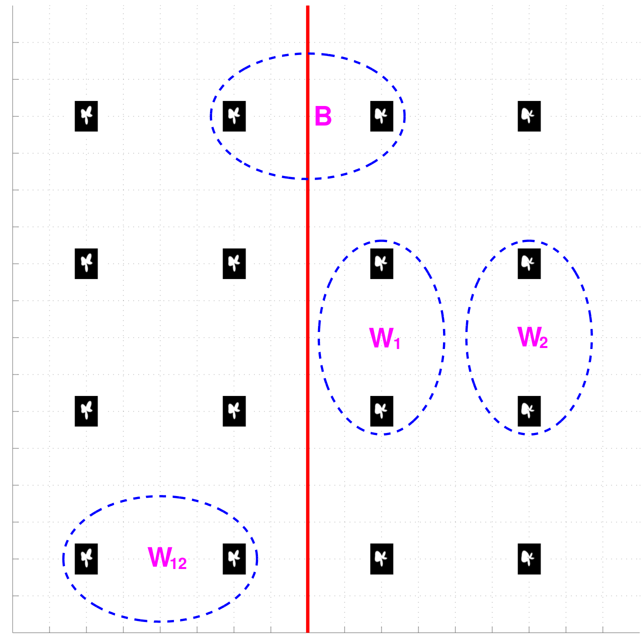

Figure 2: Examples of the four equivalence classes of trial type on the discrimination task in the test phase: B – distractor is across category boundary; W_{12} – distractor is perpendicular to boundary but within same category; W_1 – within-category distractor is parallel with and near to boundary; W_2 – within-category distractor is parallel with and far from boundary.

There were several possible univocal outcomes. First, there could be no effect of category training on the discrimination of RFC stimuli. This result would be consistent with the idea that there are limits on the influence of experience on perceptual processing. In particular, although experience may serve to differentiate perceptual dimensions that are already coded by the perceptual system, other not so privileged dimensions, such as the RFC dimensions, cannot be so flexibly used. Second, there could be faster or more accurate discrimination of stimuli along the relevant dimension (W_{12} and B trials) than along the irrelevant ones (W_1 and W_2 trials). This would occur if training sensitized the relevant RFC dimension along the entire trained range of its values. Finally, there might be an acquired distinctiveness effect, i.e. a faster or better discrimination for trials where the distractor is in another category (B

trials) than when the distractor is in the same category (W_1, W_2, W_{12}; Goldstone, 1994).

Method

Participants Ninety-five students at Indiana University participated in partial fulfillment of a class requirement. Participants were randomly assigned to either the horizontal boundary or the vertical boundary condition. Forty-nine were in the horizontal boundary condition, and 46 were in the vertical boundary condition.

Stimuli Stimuli were complex, closed contours composed of seven RFCs as described above and in Op de Beeck et al. (2003). There were 16 different values along each of the two dimensions, making a square stimulus space. The stimuli used during categorization training were two units apart on each dimension. The stimuli used during the discrimination task were four units apart and distinct from the stimuli used during the training phase. Illustrations of the stimuli used for the discrimination phase are shown in Figure 1.

Procedure There were 368 categorization trials followed by 192 discrimination trials. On each categorization trial, a stimulus was displayed on a black background near the center of the screen. The exact position of the stimulus was varied randomly as much as 25 pixels to either side of the center, both horizontally and vertically. For each stimulus the participant pressed either the '1' key or the '0' key on the keyboard, indicating the category to which they believed the stimulus belonged. The correct response was presented to the left of the stimulus, either in white, or if the participant responded incorrectly, in red. The next trial began after a black screen was displayed for 500 m.s. There were 23 trial blocks with 16 stimuli each and no breaks between trial blocks. The 16 stimuli in each of the blocks consisted of four from each quadrant of the space. The order of presentation of the stimuli within a trial block was randomized for each participant. Initially, the stimuli presented were those in the center of the irrelevant dimension and perpendicular to the category boundary. As training progressed, the stimuli presented were those closer to the boundary. In the final trial blocks, the stimuli presented were those closest to, and parallel with, the category boundary. (See Figure 3.) Both the horizontal and vertical conditions were shown identical stimuli with the same frequency, but in the opposite order.

After all the training trials were completed, instructions for the discrimination test appeared on the screen. On each discrimination trial, a stimulus was shown in the center of the screen, near the bottom, for 2000 m.s., followed by a mask for 500 m.s. Finally, two stimuli were shown near the top of the screen. These two stimuli remained, along with instructions for the participant to choose which of them was identical to the initial stimulus, until the participant responded. We used a forced-choice task rather than the same/different task that Op de Beeck et al. (2003) and Goldstone and Steyvers (2001) used because the same/different task is not bias free. (Participants may have a bias to respond "same" or "different" in the previously-used task and this bias could potentially influence the d' measure (Balakrishnan, 1998).)

The experiment lasted approximately 30–40 minutes for most participants.

Results

Overall, there were no differences across the horizontal and vertical boundaries, and so for all analyses we collapsed across conditions. Training, which progressed from simpler to more difficult categorizations, produced good performance early in training, greater than 80% correct, and dropped slowly to 60% by the final trial block, which is significantly better than chance, $t(94) = 6.03, p < .0001$. Accuracy during the discrimination phase did not differ by condition, $F \approx 1$, and participants were 67% correct overall. In the analysis of participants' reaction times, we used only correct trials[2] for which the reaction time was between 300 m.s. and three standard deviations above the mean (8012 m.s.). In addition, while most participants had few if any responses that did not meet the inclusion criteria, there were seven participants with less than 80% of their responses meeting the inclusion criteria. These participants were excluded from this analysis. An Analysis of Variance using discrimination trial type (W_1, W_2, W_{12}, B) as a within-subjects variable yielded a significant main effect, $F(3,279) = 8.85, p < .001$. Pair-wise comparisons revealed that participants were significantly faster ($p < .01$) to respond to B trials ($M = 1470$ m.s.) than any of the other trial types (1572-1593 m.s.) and that there were no significant differences among the other groups.

That participants are faster to reject distractors near the category boundary when they are along the relevant dimension than when they are along the irrelevant dimension suggests selective sensitization of the sort for which Op de Beeck et al. (2003) failed to find evidence. That participants are faster at making accurate discriminations at the category boundary than farther back on the relevant dimension makes this result a case of acquired distinctiveness, echoing earlier results by Goldstone (1994).

Experiment 2

Although Experiment 1 successfully showed an effect where the category-relevant but not the irrelevant dimension improved at the category boundary, we will next consider whether it could be possible to differentiate the two dimensions in other regions of the space as well. In order to do this, we devised a new testing procedure designed to make the category-relevant dimension as salient as possible *throughout* the two-dimensional space: participants were shown animations of change along the category-relevant dimension.

Method

Participants Ninety-seven students at Indiana University participated in partial fulfillment of a class requirement. Participants were randomly assigned to either the horizontal animation or the vertical animation condition. Forty-eight were in the horizontal animation condition, and 49 were in the vertical animation condition.

Stimuli The stimuli used in the discrimination phase were identical to those used in Experiment 1. The animations for the training phase were created by linearly interpolating between the points on this grid, one animation each per row and

[2] On incorrect trials, reaction time may or may not correlate with discriminability. For example, the trial might be incorrect because the subject has lost interest and did not try, but rather clicked a choice immediately.

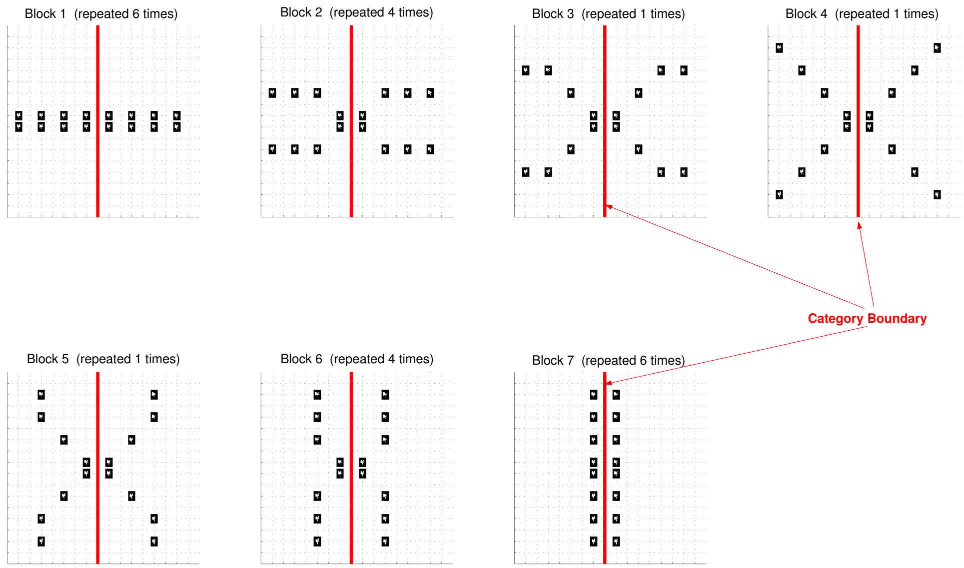

Figure 3: Training Phase progression for participants in the Vertical Boundary condition. (For participants in Horizontal Boundary condition, the block sequence would be in reverse order.) Within a block, the order was randomized.

column of the 16x16 space. Each animation consisted of 120 frames and lasted 8 seconds.

Procedure In the training phase, participants were first shown an animation where a man's face morphed into a woman's as an example of a transition from an instance of one category to another. They were told that they would be watching similar transitions with two novel categories and had to learn to identify the boundary point. On each trial, an animation across a single row or column (depending on the condition) was shown and participants were to press the space bar at the point they guessed the boundary was crossed. When they pressed the space bar a beep would sound (and the precise point in time would be recorded), but the animation always continued to its conclusion. At the end of the trial, feedback was given as to the position of the participant's guess relative to the category boundary by displaying what the shape looked like at the point of the guess and pointing to where this was on an axis abstractly representing the dimension (the category boundary was always shown as the midpoint of this axis). Figure 4 shows a screen shot of this. To prevent participants from learning to predict the boundary based simply on how much time had elapsed in the animation, the starting point was chosen randomly on each trial to be somewhere within the first 3.2 seconds of the animation. This was reflected on the feedback axis by shortening the portion prior to the midpoint by the appropriate amount.

The animations proceeded twice through the space in order from the first row (or column) consecutively up to the sixteenth and then back down to the first. For each row (or column), there was one trial where the animation was shown normally and one trial where it was shown in reverse, making 128 training trials in total. By showing the animation going both directions along the dimension, participants had to learn to predict the category boundary from both sides. For each of the two consecutive trials on a row (or column), the direction to show first (for example, 'left' or 'down' vs. 'right' or 'up') was randomly chosen.

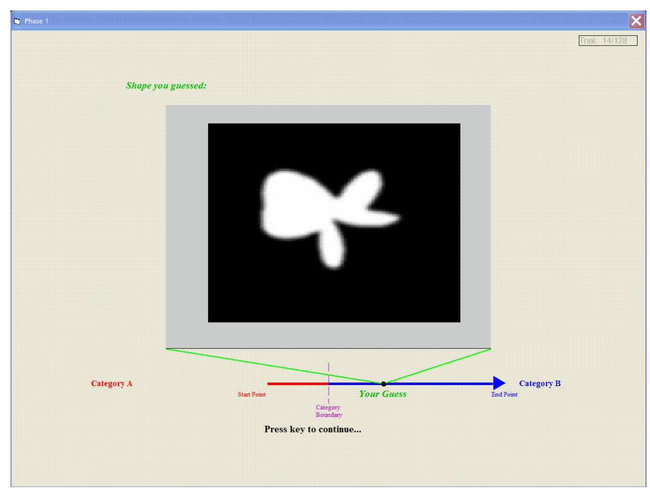

Figure 4: Screen shot of feedback given during Training phase. On this trial, the button was pressed too late, after the boundary had been passed.

The animation phase was then followed by a discrimination phase exactly the same as in Experiment 1 with the exception that, in the interest of time, there were only 96 trials. Figure 5 illustrates the procedure.

Results

For each training trial, Δ_t, the absolute value of the distance in seconds between a participant's guessed boundary point and the actual boundary, was calculated. If participants are improving at this task, then as training progresses, Δ_t should decline. Thus, a negative correlation between Δ_t and the number of the training trial is evidence of learning in this task. This was true for seventy-six of the 97 participants, with the average correlation across all participants being $r = -.754$ ($p \ll 0.0001$). The average Δ_t started around 1.4 seconds and decreased to about 1.0 second.

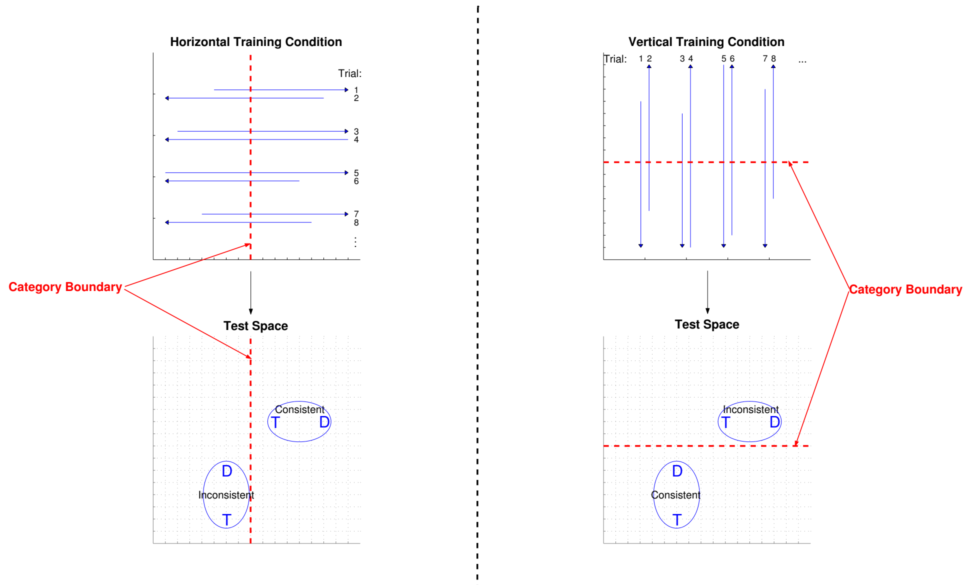

Figure 5: Training Phase animations proceeded across/down one row/column, depending on the condition (top two panels), and then back again on the next trial. The starting point was randomized as shown. Only four rows/columns are shown here, but in the experiment there were 16. Examples of two Testing Phase trials are shown for each condition (bottom two panels), one of each type. On a trial, a target ('T') was shown, then a mask, and then the target with the distractor ('D').

In the discrimination phase, the mean accuracy was 71.3% (S.D. = 9.5%) and the mean response time over all trials was 1694 m.s. (S.D. = 1175 m.s.). The overall error rates in the horizontal and vertical conditions were not significantly different. Any trial with a response time less than 300 m.s. or greater than 8012 m.s. was discarded, and one participant was not included because this was true for more than 20% of their trials. Additionally, six more participants were excluded because the pattern of their results indicated that they had responded randomly. Thus, 90 participants were included for the remaining analyses, 44 in the horizontal condition and 46 in the vertical.

Each trial in the discrimination phase was classified as *consistent* with the animation if the distractor was displaced on the category-relevant dimension (the dimension along which the animation took place) and *inconsistent* if it was displaced along the irrelevant dimension. The mean error rates relative to the total number of errors for each subject on the consistent and inconsistent dimension (globally, .522 and .478 respectively) were compared using a paired-samples t-test and found to be significantly different ($t(89) = 2.13, p < .04$). This is evidence that the animation differentially changed performance on one dimension more than the other. The fact that the consistent-dimension distractor caused more errors than the inconsistent dimension is consistent with the results found by Hockema (2004), where viewing animations along a dimension caused decreased discrimination along that dimension.

When location in the space was taken into account by dividing the discriminations into four categories based upon whether or not they occurred on a horizontal or vertical edge of the grid, discriminations in the middle of the space (both B and W_1 in Figure 2) were more accurate ($p < .001$) and faster ($p < .001$) than those on the borders (W_2 and W_{12}). This reflects the fact that participants were exposed more to these stimuli in general, given both the random starting offsets of the animations on each trial and the probability that they did not watch the animation as closely after pressing the space bar on a trial. However, there were no significant interactions between condition and accuracy by grid location. Thus, the differential accuracy effect appears to have affected the whole dimension.

Finally, similar to Experiment 1, while the average response times for consistent and inconsistent trials in general were not significantly different, participants were significantly faster at making accurate (consistent) discriminations across the boundary (B trials) than they were at making accurate discriminations elsewhere ($p \ll .0001$), including W_1 trials ($p < .02$). Thus, the acquired distinctiveness effect found in Experiment 1 was replicated here as well.

General Discussion

Experiments 1 and 2 successfully showed that categorization training can lead to acquired distinctiveness of the relevant dimension. Experiment 2 went even further to show that the differential effects are not just limited to occurring at the category boundary. Taken together, these results provide strong evidence that novel, perceptually-integral dimensions can indeed be differentiated with appropriate categorization training.

For Experiment 2, the fact that the error rate for discrimination trials with the distractor differing on the animated (category-relevant) dimension was higher than the irrelevant dimension is consistent with Hockema (2004), where it was modeled using a recurrent connectionist network. One theory as to why this might be the case is that because the pathway between the two adjacent shapes—the target and the distractor—on this dimension has been explicitly elucidated, they might be seen as more similar to one another. That is, the animation might change the underlying similarity space.

Another theory is that the animation could make the system more sensitive along the category-relevant dimension, making subsequent comparisons more sensitive to noise inherent in the low-level perceptual system. In any event, the main point here is that the category-relevant and irrelevant dimensions were affected *differently* by the animation throughout the two-dimensional space.

We believe that the discrepancy between the results reported here and those reported by Op de Beeck et al. (2003) are probably due to differences in the training procedure. In Experiment 1, the procedure may have helped them learn the category boundary better, or it may have helped them learn the two categories *as* categories better (as opposed to merely memorizing some of the stimuli with their associated labels through repetition). If the latter, this would be consistent with work like that of Zaki and Homa (1999) that variations in the order of the training sequence can lead to more or less robust categories. Indeed, in their work they found that if the training sequence progressed through the category in a systematic way to highlight the category-relevant dimensions, this led to "transformational knowledge" of the category, in turn making it more robust. This idea also underlies the training we employed in Experiment 2.

Furthermore, the animations appear to have been an especially effective way to selectively highlight the category-relevant dimension and trigger perceptual learning. Again, this is consistent with the results found by Hockema (2004) where two dimensions were made more integral by a similar animation technique in which they were co-varied. Both cases rely on the underlying principle that our perceptual systems are oriented around *transformations*. Transformations contain rich information about the structure of the world. Indeed, it is through this temporal structure that we perceive atemporal structure. (For example, movement is very important for the detection of occluding edges.) This idea has a long history in psychology, going back through Gibson (1979) at least as far as James (1890) and Helmholtz (1866). In Experiment 2, this principle was successfully applied to trigger a perceptual adaptation that lasted long enough to affect the discrimination phase accuracy results. In future work, we will explore just how long this adaptation can last, for example, by testing to see if the differential discriminability effects wear off and, if so, how easily they can be renewed.

The present results raise at least three other interesting questions for future research. First, what are the neural mechanisms that underlie dimension differentiation? Second, what are the constraints on the dimensions that can be differentiated? For example, obviously, at a minimum, they have to be built out of our available sensory streams: we have to already be sensitive to changes on both dimensions (i.e. have the ability to perceive them integrally). For both of these questions, are the answers the same for different dimensions and across modalities, for different ages? Finally, can the techniques employed here be developed into training procedures for teaching hard-to-learn distinctions in the real world? As one example, it might be possible to better train Japanese adults learning English as a second language to discriminate and produce English /*r*/ and /*l*/ phonemes (something that has proven to be very difficult (Yamada, 1995)) with a suitable training procedure. There has already been some success training this distinction by using minimal-pair exemplars of each phoneme from multiple speakers and in differing phonetic contexts in a forced-choice categorization task (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999). In the context of the work presented here, this is similar to training block 7 in Figure 3: lots of variation on irrelevant dimensions but minimal, across-boundary pairs along the relevant dimensions. Thus, there are practical applications for dimension differentiation techniques that provide additional incentive to understanding the process.

Acknowledgments

The first author of this paper was supported through a grant from the U.S. National Institute for Child Health and Human Development (T32 HD07475). The second author was supported through a grant from the U.S. National Institute for Mental Health (F32 MH068218). Thank you to Andrew Long for helping collect the data for this paper.

References

Balakrishnan, J. D. (1998). Some more sensitive measures of sensitivity and response bias. *Psychological Methods*, *3*, 68–90.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /*r*/ and /*l*/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, *61*(5), 977–985.

Burns, B., & Shepp, B. E. (1988). Dimensional interactions and the structure of psychological space: The representation of hue, saturation, and brightness. *Perception and Psychophysics*, *43*, 494–507.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, *1*, 225–241.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Hillsdale, NJ, USA: Lawrence Erlbaum Associates.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, *123*(2), 178–200.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585–612.

Goldstone, R. L., & Steyvers, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General*, *130*(1), 116–139.

Helmholtz, H. L. F. von. (1866). Concerning the perceptions in general. In *Treatise on physiological optics* (Vol. III, pp. 1–37). (Translated by J. P. C. Southall, 1925, *Optical Society of America*. Reprinted in 1962, New York: Dover Publications, Inc.)

Hockema, S. A. (2004). Perception as prediction. In *Proceedings of the 2004 Cognitive Science conference.*

James, W. (1890). The Perception of Time. In *The Principles of Psychology* (Vol. 1, pp. 605–642). Cambridge, MA, USA: Harvard University Press.

Op de Beeck, H., Wagemans, J., & Vogels, R. (2003). The effect of category learning on the representation of shape: Dimensions can be biased but not differentiated. *Journal of Experimental Psychology: General*, *132*(4), 491–511.

Smith, L. B. (1989). From global similarities to kinds of similarities: The construction of dimensions in development. In *Similarity and analogy* (pp. 146–178). Cambridge, MA, USA: Cambridge University Press.

Yamada, R. A. (1995). Age and acquisition of second language speech sounds: Perception of American English /*r*/ and /*l*/ by native speakers of Japanese. In *Speech perception and linguistic experience: Issues in cross-language research* (pp. 305–320). Baltimore, MD, USA: York Press.

Zaki, S. R., & Homa, D. (1999). Concepts and transformational knowledge. *Cognitive Psychology*, *39*, 69–115.