The Sensitization and Differentiation of Dimensions During Category Learning

Robert L. Goldstone
Mark Steyvers
Indiana University

Correspondence Address: Robert Goldstone
                             Department of Psychology
                             Indiana University
                             Bloomington, IN. 47408

Other Correspondences: rgoldsto@indiana.edu
                             (812) 855-4853

Abstract

The reported experiments explore two mechanisms by which object descriptions are flexibly adapted to support concept learning: selective attention and dimension differentiation. Arbitrary dimensions were created by blending photographs of faces in different proportions. Consistent with learned selective attention, positive transfer was found when initial and final categorizations shared either relevant or irrelevant dimensions. Unexpectedly good transfer was observed both when irrelevant dimensions became relevant and relevant dimensions became irrelevant, and was explained in terms of participants learning to isolate one dimension from another. This account was further supported by experiments indicating that conditions expected to produce positive transfer via dimension differentiation produced better transfer than conditions expected to produce positive transfer via selective attention, but only when stimuli were composed of highly integral and spatially overlapping dimensions.

The Sensitization and Differentiation of Dimensions During Category Learning

People's ability to learn new concepts is a critical part of their ability to flexibly accommodate to their world and tasks. Concept learning allows children to develop similar conceptions of their world to the adults of their community (Gershkoff-Stowe, Thal, Smith, & Namy, 1997), individuals to coordinate their shared understanding of a situation (Markman & Makin, 1998), and experts to organize their world in useful ways (Estes, 1994). At times, this conceptual flexibility can be achieved by simply combining and rearranging existing perceptual features, such as when concepts are acquired by learning logic-based rules involving pre-established features (Nosofsky, Palmeri, & McKinley, 1994). At other times, conceptual flexibility must be accompanied by flexibility from perceptual and attentional processes

(Goldstone, 1998; Hock, Tromley, & Polmann, 1988; Schyns, Goldstone, & Thibaut, 1998). That is, the perceptual features that are used as inputs to concept learning processes are, themselves, adapted to the concepts being learned. The experiments reported here examine how perceptual descriptions of objects are influenced by the acquired categories that make use of the descriptions. Evidence for both selective attention and differentiation of dimensions is found, with additional constraints placed on the mechanisms that these processes use.

Selective Attention During Category Learning

When people learn to make a new categorization, they often have to selectively attend to some features of the objects to be categorized and ignore other features. To categorize an object as a book, color is irrelevant but shape must be attended. To categorize an object as a banana, both shape and color are relevant. One can selectively attend to color and ignore shape to distinguish ripe avocados from those that are not ready to eat. Effective categorization depends on our ability to flexibly attend to different features on different occasions. The first reported study explores our ability to flexibly attend arbitrary dimensions, and how learned attention to dimensions transfers positively and negatively to a subsequent a categorization task.

One way that perception becomes adapted to tasks and environments is by increasing the attention paid to perceptual dimensions and features that are important, and/or by decreasing attention to irrelevant dimensions and features. A feature is a unitary stimulus element, whereas a dimension is a set of linearly ordered values. "3 centimeters" and "gray" are features; length and brightness are dimensions. Despite the apparently clear-cut definitional difference between dimensions and features, deciding whether a psychological structure (such as *shape*) is a feature or dimension is often difficult and perhaps even arbitrary. In the following

discussions, evidence for selective attention (and later, differentiation) to dimensions and features will be combined together, even though the two types of selective attention may dissociate. For example, Kersten, Goldstone, and Schaffert (1998) find that training on a particular dimension value may cause attention to be subsequently increased to the dimension while at the same time decreased to the particular dimension value.

Most successful theories of categorization and learning incorporate some notion of selective attention. In Sutherland and Mackintosh's (1971) analyzer theory, learning a discrimination involves strengthening the tendency to attend to relevant "analyzers." In Nosofsky's (1986, 1991) exemplar model of categorization, the categorization of an object depends on its similarity to previously stored category members in a multidimensional space (see also Medin & Shaeffer, 1978). Critically, distances between objects are compressed and expanded along dimensions in this space depending on the categorization required. Dimensions that are relevant for a categorization are expanded while the distances between objects on irrelevant dimensions are compressed. For example, Nosofsky finds that if participants are given a categorization where the angle of a line embedded in a circular form is relevant while the size of the circular form is irrelevant, then distances between objects on this dimension are increased. This process will be called "attention weighting" and refers to the flexible allocation of attention to stimulus analyzers, features, or dimensions. Further work in this line has shown how neural networks acquire task-appropriate weights for stimulus dimensions (Kruschke, 1992).

Researchers in animal learning and human categorization have described shifts toward the use of dimensions that are useful for tasks (Blough & Blough, 1997; Nosofsky, 1986) or have previously been useful (Lawrence, 1949). Lawrence describes these situations as examples of stimulus dimensions "acquiring

distinctiveness" if they have been diagnostic in predicting rewards. The stimulus aspects that are selectively attended may be quite complex; even pigeons can learn to selectively attend to the feature "contains human" in photographs (Herrnstein, 1990). In addition to important dimensions acquiring distinctiveness, irrelevant dimensions also acquire equivalence, becoming less distinguishable (Honey & Hall, 1989). For example, in a phenomena called "latent inhibition," stimuli that are originally varied independently of reward are harder to later associate with a reward than those that are not initially presented at all (Lubow & Kaplan, 1997; Pearce, 1987). Haider and Frensch (1996) find that improvements in performance are frequently due to reduced processing of irrelevant dimensions. Thus, there is evidence that learning involves both increasing attention to relevant dimensions, and decreasing attention to irrelevant dimensions.

The above studies illustrate shifts in the use of dimensions as a function of their task relevance, but from these studies it is not clear how mandatory or strategic these shifts are. One source of evidence that allocation of attention is not completely under voluntary control is that attentional highlighting of information occurs even if it is to the detriment of the observer. When a letter consistently serves as the target in a detection task, and then later becomes a distracter (a stimulus to be ignored) it still automatically captures attention (Shiffrin & Schneider, 1977). The converse of this effect, negative priming, also occurs. Targets that were once distracters are responded to more slowly than never-before-seen items (Tipper, 1992). In the negative priming paradigm, the effect of previous exposures of an item can last upwards of two weeks (Fox, 1995), suggesting that a relatively permanent change has taken place. In addition to suggesting that attention is not completely determined by short-term strategic demands, these studies also show negative transfer effects due to acquired attentional weights. When the learned attention

weights to a feature or dimension are inappropriate for a subsequent task, then performance on this task will tend to suffer (see also Goldstone, 1994-a).

Dimension Differentiation During Category Learning

Attention weighting is a critical component of categorization learning, but it may not be the only process that dynamically alters the description of an object in a categorization task. A second candidate process is dimension differentiation, by which dimensions that are originally psychologically fused together become separated and isolated. Attention weighting presumes that the different dimensions that make up a stimulus can be selectively attended. To increase attention to size but not color, one must be able to isolate size differences from color differences. In his classic research on stimulus integrality and separability, Garner argues that stimulus dimensions differ in how easily they can be isolated or extracted from each other (Garner, 1976, 1978; Garner & Felfoldy, 1970). Dimensions are said to be separable if it is possible to attend to one of the dimensions without attending to the other. Size and brightness are classic examples of separable dimensions; making a categorization on the basis of size is not significantly slowed if there is irrelevant variation on brightness. Dimensions are integral if variation along an irrelevant dimension cannot be ignored when trying to attend a relevant dimension. The classic examples of integral dimensions are saturation and brightness, where saturation is related to the amount of white mixed into a color, and brightness is related to the amount of light coming off of a color. For saturation and brightness, it is difficult or impossible to attend to only one of the dimensions (Burns & Shepp, 1988; Foard and Kemler, 1984; Garner, 1976; Melara, Marks, & Potts, 1993).

From the above work distinguishing integral from separate dimensions, one might conclude that attention weighting processes can proceed with separable but not integral dimensions. However, one interesting possibility is that category learning can, to some extent, change the status of dimensions, transforming dimensions that were originally integral into more separable dimensions. Experience may change the underlying representation of a pair of dimensions such that they come to be treated as relatively independent and non-intefering sources of variation that compose an object. Seeing that stimuli in a set vary along two orthogonal dimensions may allow the dimensions to be teased apart and isolated, particularly if the two dimensions are differentially diagnostic for categorization. There is developmental evidence that dimensions that are easily isolated by adults, such as the brightness and size of a square, are treated as fused together for four-year old children (Kemler & Smith, 1978; Smith and Kemler, 1978). It is relatively difficult for children to decide whether two objects are identical on a particular dimension, but relatively easy for them to decide whether they are similar across many dimensions (Smith, 1989a). Children show considerable difficulty in tasks that require selective attention to one dimension while ignoring another, even if the dimensions are separable for adults (Smith & Evans, 1989). When given the choice of sorting objects by their overall similarity or by selecting a single criterial dimension, children tend to use overall similarity whereas adults use the single dimension (Smith, 1989b). Whereas older children and adults tend to organize objects into groups by single dimensions (Regehr & Brooks, 1995), children under the age of five often organize objects by overall similarity (Shepp, Burns, & McDonough, 1980). Finally, adjectives that refer to single dimensions are learned by children relatively slowly compared to nouns (Smith, Gasser, and Sandhofer, 1997).

One commonly used method for diagnosing whether a participant is treating an object as composed out of well or poorly differentiated dimensions has been the "triad sorting task." In this task, participants are shown three objects, and are asked to group the objects together in a way that makes sense to them. If shown a blue circle (A), an orange

circle (B), and a red ellipse (C), a participant's grouping B with C has been taken as evidence of responding to relatively differentiated overall similarity, because both color and shape dimensions are used to create the groups, and because B and C are overall similar across both dimensions but have no single dimension in common (Kemler & Smith, 1978; Shepp et al., 1980; Smith & Kemler, 1978). By a similar logic, grouping A and B together is taken as evidence for analytic, dimensionalized processing, because the grouping selectively focuses on the shape identity while ignoring the large color difference. Although shape and color need not be completely undifferentiated for children, the evidence from the triad sorting and selective attention tasks suggests that color and shape dimensions are less likely to be isolated for children than adults. There have been concerns with respect to the methodology and generality of the results from triad sorting tasks (Aschekenasy & Odom, 1982; Cook & Stephens, 1995; Cook & Odom, 1992; Ward & Vela, 1986). Still, the hypothesis that there is a developmental trend from integral to separable dimensions has received support. In many cases, perceptual dimensions seem to harder to isolated for children than adults, such that children cannot easily access the individual dimensions that compose an object.

The developmental trend toward increasingly differentiated dimensions (dimensions that can be isolated) is echoed by adult training studies. Under certain circumstances, color experts (art students and vision scientists) are better able to selectively attend to dimensions (e.g. hue, chroma, and value) that comprise color than are non-experts (Burns & Shepp, 1988). Researchers have found that it is possible to train participants to categorize on the basis of saturation rather than brightness (or vice versa) (Foard & Kemler, 1984), and once learned, categorization judgments are well explained by attention being selectively placed on the relevant color dimension (Nosofsky, 1987). Research in our laboratory (Goldstone, 1994-a) has shown that people who learn a categorization in which saturation is relevant and brightness is irrelevant (or vice versa) can learn to perform the categorization accurately, and as a result of category learning, they develop a selectively heightened sensitivity at making saturation, relative to brightness, discriminations. That is, categorization training that makes one dimension diagnostic and another dimension nondiagnostic can serve to split apart these dimensions, even if they are traditionally considered to be integral dimensions. These training studies show that to know how integral two dimensions are, one has to know something about the observer's history. Originally fused dimensions can become at least partially split apart with training. Melcher and Schooler (1996) provide suggestive evidence that expert, but not non-expert, wine tasters isolate independent perceptual features in wines that closely correspond to the terminology used to describe wines.

Research on the role of comparison in concept learning also points to people's increasing appreciation for the dimensional organization underlying objects. Gentner and her colleagues (Gentner & Markman, 1997; Gentner & Namy, in press; Kotovsky & Gentner, 1996; Markman & Gentner, 1996; Medin, Goldstone, & Gentner, 1993) have shown that the act of comparing objects causes their dimensional organization to become more transparent. Dimensions that have clear correspondences between compared objects become selectively emphasized. This work generally suggests that both dimension values and relations between dimensions are more efficiently isolated with age and experience.

There is a relation between the integrality of dimensions and their salience, but these notions can and should be kept distinct. A categorization involving a dimension may be difficult either because the dimension is psychologically fused with another dimension, or because it is hard to detect the dimension

because of its lack of salience. Mathematical analyses have suggested that what can be modeled by assuming dimensional integrality can frequently also be modeled by including noise that interferes with the detection of small differences along separable dimensions (Ennis, 1988). In addition, one reason why a dimensional difference may be difficult to detect is because of the noise added by the intrusion of other dimensions that cannot be ignored. However, the integrality of a dimension is not equivalent to its subtlety. Dimension salience is a property of the dimension itself, but integrality is always a relation between two or more dimensions.

Concerns have been raised about some of the most common methods for assessing the degree of differentiation between dimensions. One technique diagnosing undifferentiated dimensions has been to compare a category learning situation in which the two categories can be differentiated on the basis of a single dimension (a horizontal or vertical categorization rule) to a situation in which both dimensions must be considered (a diagonal categorization rule). If the former categorization is as accurate and fast as the rotated, diagonal categorization, then there is some grounds for believing that the dimensions are not psychologically necessary descriptions of the objects. Conversely, if an upright, horizontal or vertical categorization is easier than the rotated categorization, then the dimensions are viewed as subjectively represented or "psychologically privileged" (Grau & Nelson, 1988). However, by this measure, even apparently integral dimensions such as saturation and brightness show evidence of being psychologically privileged (Foard & Kemler, 1984; Melara, Marks, & Potts, 1993) in that the rotated categorization is more difficult than the upright categorization. This measure may still be useful in a relative, not absolute, sense. That is, the more differentiated two dimensions are, the greater the difference between rotated and upright categorizations

(Kemler-Nelson, 1993). This interpretation is consistent with the existing literature suggesting that multiple-dimension categorization rules are especially harder than single-dimension categorization rules when the dimensions are separable rather than integral (Kruschke, 1993; Maddox, 1992).

A second technique for assessing the degree of differentiation between dimensions has been to find best fitting values of r in the distance formula:

$$D_{i,j} = \left[\sum_{k=1}^{n} \left| X_{ik} - X_{jk} \right|^r \right]^{\frac{1}{r}},$$

where $D_{i,j}$ is the subjective dissimilarity between objects i and j, n is the number of dimensions, X_{ik} is the value of item i on dimension k, and r is a parameter that allows different spatial metrics to be used (if r=1, then the distance between items is equal to the sum of their dimensional differences; if r=2, then the distance is the length of shortest line that connects the items). A typical finding is that dissimilarities between stimuli composed out of undifferentiated dimensions are best fit by letting r equal 2, whereas dissimilarities between stimuli composed out of differentiated dimensions are best fit by letting r equal 1 (Handel & Imai, 1972; Maddox, 1992; Melara, 1989). However, difficulties arise with this technique as well. The value of r that best fits human similarity assessments depends on participants' strategies as manipulated by instructions (Melara, Marks, & Lesko, 1992). Stimuli that are composed of dimensions with very small value differences are often better fit with r=2 than r=1 even if the dimensions are separable (Nosofsky, 1987). Furthermore, the relation between model/data fit and r value is often times non-linear (even nonmonotonic) and extremely noisy.

In summary, there is a large body of evidence that dimension differentiation occurs with learning and may characterize child-to-adult development and novice-to-expert

training. At the same time, there are methodological reasons to search for new ways to asssses dimension differentiation. The current experiments (particularly, Experiments 2, 3, and 4) were designed to explore dimension differentiation with a new paradigm based on transfer between category learning tasks.

The Current Investigation

The current experiments explore both processes of dynamic selective attention to dimensions and differentiation of dimensions. Rather than using previously established and possibly innate dimensions such as hue and size, the first three experiments use arbitrary dimensions. Other researchers have also studied concept learning using arbitrary dimensions (e.g. Schyns & Rodet, 1997). For example, in classic studies of dot pattern classification, randomly generated prototypes are often created by assigning random locations to a set of dots (Posner & Keele, 1968). Hock, Tromley, & Polmann (1988) argue that arbitrary configurations of dots can be learned if they are diagnostic for a learned categorization. Arbitrary dimensions are used in the current experiments because learned dimension differentiation should only be observed when participants do not initially organize the stimuli into the dimensions in question. If subjects possess the differentiated dimensions before the start of the experiment, as is the case for brightness and size, then dimension weighting, but not dimension differentiation, should be observed. Saturation and brightness, the dimensions used in Experiment 4, are more integral dimensions, but Grau and Nelson (1988; see also Melara, Marks, & Potts, 1993) have argued that even these dimensions are not genuinely arbitrary. By creating arbitrary dimensions in the first experiments, we can hopefully obtain stimuli that participants do not originally organize according to these dimensions. Instead, the dimensional organization may be learned during category learning. It is possible, however, that participants pre-experimentally possess

dimensions that are correlated with the arbitrary dimensions. This possibility will be discussed in the General Discussion.

Arbitrary dimensions are created by taking two randomly selected bald heads, and creating a morph sequence between them. For example, one dimension is created by morphing between Faces 1 and 2 in Figure 1, and a second dimension is created by morphing between Faces 3 and 4. Using a technique described by Steyvers (1999), a 4 by 4 matrix of faces can be created from these two dimensions such that each face is defined half by its value on Dimension A and half by its value on Dimension B. Although participants may not originally characterize faces by their values on these two dimensions because they have been arbitrarily constructed, with sufficient practice people may use these dimensions if they are relevant for a categorization, or if they vary systematically within a set of faces. By creating these arbitrary dimensions, we can explore whether selective attention processes can operate on arbitrary dimensions, and whether people can learn to organize stimuli according to arbitrary dimensions. Faces were used as the endpoints for our dimensions because of evidence suggesting that faces are processed in a particularly configural manner (Farah, 1992; Tanaka & Farah, 1993). We should have the greatest chance of finding evidence for dimension differentiation with stimuli that are not naturally decomposed into delineated parts.

Our dimensions dimensions requires a closer examination of what is meant by a psychological dimension and arbitrariness. Routinely, people use the term "psychological dimension" when there is a relatively straightforward connection between an internally represented stimulus aspect and a relatively easily specified physical characteristic of an external object. For example, the psychological dimension of brightness is related, although not necessarily linearly, to the amount of light coming off of an object, which can be measured in candelas per meter squared.

In our experiments, we do not know the exact physical properties that are related to the psychological dimensions that are created by morphing between faces. When one object is gradually morphed into another object, many changes occur simultaneously throughout the object and these changes are perfectly correlated with each other. For example, as Person A's lips become wider to match Person B's lips, Person A's eyes may simultaneously become smaller to match B's eyes. An observer may represent the morph-based dimension in terms of either of these changes, or both. However, we still consider the dimensions that are sensitized and differentiated to be psychological dimensions despite the difficulty in associating them with simple physically measurable properties. Certainly other psychological dimensions, including "beauty," "shyness," and "abstractness" are similarly difficult to directly link to specific physical properties.

We will adopt a functional rather than extrinsic approach to identifying psychological dimensions. A psychological dimension is any ordered set of mutually exclusive values that can be selectively attended. That is, stimulus elements are processed as a unified psychological dimension to the extent that they can be processed independently of other stimulus properties (Goldstone, Steyvers, Spencer-Smith, & Kersten, 2000). Thus, while saturation and brightness are not strong psychological dimensions for most people, they may be for color experts (Burns & Shepp, 1988) if the color expert can demonstrate that they can attend to saturation without being very influenced by brightness. For our purposes, the status of brightness and saturation as psychological dimensions does not depend on their having direct physical descriptions because rotated versions of these same dimensions would still count as psychological dimensions as long as people could selectively attend to the rotated dimensions.

In calling the morph-generated dimensions arbitrary, we mean simply that the dimension that is relevant for categorization is selected at random from a large set of potential dimensions (see Goldstone, 2000). That is, our categorization rules are randomly selected, and for different categorizations, different dimensions would be relevant. As suggested above, dimensions need not be arbitrary in the sense of having no relation at all to pre-experimentally possessed dimensions. Rather, the important notion is that no matter which faces are selected as endpoints of a dimension, some aspect will be relevant for the categorization, and the issue is whether this aspect can come to be isolated and selectively weighted. Given the large number of dimensions that can be created by pairing each face from a set with every other face (there are 62 faces in our sample), if an arbitrary dimension from this set can be isolated and weighted, then these processes will need to be highly flexible.

If people can develop an ability to selective attend to arbitrary, laboratory-constructed dimensions, then learning that a particular dimension is relevant should facilitate learning subsequent categories for which this same dimension is relevant, and may interfere with the learning of categories for which the dimension is irrelevant. If people learn not only to selectively attend to arbitrary dimensions, but also learn to differentiate arbitrary dimensions, then this may also be reflected in transfer across categorizations. In particular, positive transfer is not only expected when the same dimension is relevant for two categorizations, but is also predicted when the two categorizations require the same differentiation of the stimuli into dimensions. The current experiments go beyond previous studies in that we are not only interested in how to tell whether two dimensions are treated as unitary or are differentiated, but we are also interested in the effect that category learning has on dimension differentiation. Thus, Experiment 1 was designed to test the influence of category learning on the isolation and selection of dimensions. These processes were

explored by transferring participants from one categorization rule to another. Positive transfer by selective attention were predicted when the initial and final categorization rules shared relevant or irrelevant dimensions, and negative transfer was predicted if initially relevant dimensions became irrelevant or vice versa. Positive transfer by dimension differentiation was predicted when the initial and final categorization required the same isolation of dimensions, and negative transfer by dimension differentiation was predicted when the two categorizations required incompatible organizations of stimuli into dimensions.

Experiment 1

The primary goal of the Experiment 1 was to explore selective attention toward relevant, and away from irrelevant, arbitrary dimensions during a category learning task, and to observe the transfer of attention to subsequent categorizations. By comparing transfer performance on the second categorization as a function of the first categorization, we can discover whether learned relevance and learned irrelevance both arise, and whether one type of learning is significantly stronger than the other. In addition, we can observe whether selective attention to arbitrary dimensions results in positive transfer when the learned attention weights are appropriate, negative transfer when learned attention weights are inappropriate, and if both occur, which type of learning is stronger.

In this experiment, participants learned two categorizations, an initial and transfer categorization. The transfer categorization was identical for all participants, and involved a categorization in which Dimension A was relevant and Dimension B was irrelevant. The initial and transfer categorizations were related to each other in one of seven ways. A dimension that was initially relevant could continue to be relevant, could become irrelevant, or could be replaced altogether. The same alterations were applied to the irrelevant dimension. Thus, by observing transfer

categorization performance, we can assess how appropriate the initial attention weights to dimensions were for subsequent categorization tasks. By giving all seven groups of participants the same transfer categorization, we can be confident that any systematic differences between the groups on final categorization performance must be due to differences in how well the initial categorization prepared them for this final categorization.

Method

Participants. 199 undergraduate students from Indiana University served as participants in order to fulfill a course requirement. The students were split approximately evenly into seven conditions.

Materials. The stimuli were faces that were generated by morphing between photographs of bald heads selected from Kayser (1997). Sample photographs that were used in generating the 16 morphs of a set are shown on the sides of Figure 1 and are labeled Faces 1, 2, 3, and 4. Each presented face varied along two arbitrary dimensions, where dimensions were generated by creating negative contingencies between two faces — the more of Face 1 that was present in a particular morphed face, the less of Face 2 there was. The horizontal dimension shown in Figure 1, Dimension A, might be called the "The proportion of Face 1 relative to Face 2" dimension, where the two faces were randomly paired. The vertical dimension, Dimension B, might be called the "The proportion of Face 3 relative to Face 4" dimension. The two dimensions were orthogonal - the proportion of Face 1 relative to Face 2 was independent of the proportion of Face 3 relative to Face 4. Thus, each face was a blend between different faces, and different faces only varied in the proportion of each of the four faces. Each face consisted of equal proportions of the horizontal and vertical dimensions.

A subset of eight from the full 4 X 4 array of faces was used for stimuli. This subset was selected so as to form an octagon around the

center of the array, and is exemplified in Figure 2. For the half of the face that represents the horizontal Dimension A, the four columns possess 100% Face 1 (0% Face 2), 67% Face 1 (33% Face 2), 33% Face 1 (67% Face 2), and 0% Face 1 (100% Face 2). For the half of the face that represents the vertical Dimension B, the four rows possess 100% Face 3 (0% Face 4), 67% Face 3 (33% Face 4), 33% Face 3 (67% Face 4), and 0% Face 3 (100% Face 4). Thus, the leftmost face on the top row of Figure 2 is a blend of Face 1 (67% of 50%, or .335), Face 2 (33% of 50%, or .165), and Face 3 (50%). An advantage of using 8 rather than the full 4 X 4 set of faces is that the structure of the faces does not strongly suggest a dimensional organization. As such, if evidence for a dimensional organization exists, it can likely be traced to the categorization feedback provided to participants rather than the unsupervised two-dimensional array structure for the faces shown in Figure 1.

In the experiment, seven different 4 X 4 arrays of faces were created by pairing four different dimensions together. Each of the four dimensions was generated by creating a continuum between two unique faces. The eight original faces were selected from a larger database of sixty-two bald faces. The subjective similarity between each pair of these sixty-two faces was obtained by a method described by Goldstone (1994-b). The eight faces were selected because each possible pair from this set of faces received an average subjective similarity rating that was within 15% of any other pair. When faces were randomly paired together to form dimensions, the dimensions were likely to be of roughly equal salience because the endpoint faces were roughly equally similar to each other.

Each of the morphs was automatically generated using a morphing technique described by Steyvers (1999). Applying this technique, the main contours in the face images were delineated by 127 control lines. These control lines served to align the features of the four faces. In the warping phase of this morphing algorithm, correspondences were calculated between the pixels of all the images to be morphed. Then, in the cross-dissolving phase, the gray scale values of corresponding pixels were blended to create the gray scale values of the resulting morph image.

Each face was displayed in grayscale with 256 possible brightness values per pixel (one pixel = .034 cm), and measured 14.48 cm tall by 11.68 cm wide. Each face was photographed against a dark background and displayed on a white Macintosh II SI computer screen. The average viewing distance was 46 cm.

Procedure. Learned sensitivity to morph-based dimensions was tested using a category learning paradigm in which participants received an initial and transfer category learning task. Category rules involved dividing a set of eight faces either horizontally or vertically into equal halves. For a vertical categorization boundary that divided the faces into left and right sets, Dimension B was irrelevant for the categorization and Dimension A was relevant. On each trial of both intial and final categorizations, participants saw a face and categorized it by pressing either "A" or "B" on the keyboard, with feedback on each trial from the computer indicating with a check or an "X" whether or not the participant was correct, and also indicating the correct category assignment for the face. Participants were instructed: "Your task will be to categorize faces as accurately as possible into different categories (clubs) that are based purely on physical appearance. If you believe the face belongs to Club A, press the "A" key. Press the "B" key for Club B. You will received feedback indicating whether your guess is correct."

Participants completed 200 trials in each of the two category learning tasks. The initial and transfer categorizations were related to each other in one of seven ways. In the representation used in Table 1, the dimension left of the dividing line is relevant and the dimension to the right of the line is irrelevant. A vertical category boundary for the set of faces

in Figure 1 would be represented as A|B; Dimension A is relevant and Dimension B is irrelevant.

Across the seven conditions, there were four different dimensions, and for each of these dimensions, two unique faces were used as the anchoring endpoints. As shown in Table 1, the seven different conditions involved the same transfer condition, in which Dimension A was relevant and B was irrelevant. Dimensions that were relevant or irrelevant during the initial categorization could become relevant, irrelevant, or absent altogether during the transfer categorization. This produces seven different initial categorization rules because there are three possible dimension types (relevant, irrelevant, and novel) and two dimensions (yielding 2^3 conditions) but two of these (A|A, and B|B) are impossible because the same dimension would have to be both relevant and irrelevant. First, starting with the control condition (C|D), the initial and transfer stages involved completely different generating faces for their dimensions. Dimensions C and D were created by using completely different faces as endpoints than were used for Dimensions A and B, and thus the actual displayed stimuli and categorizations were unrelated across the two phases of the experiment. The next three conditions might be expected to produce beneficial transfer relative to the control condition. The first involves identity transfer, in which the dimension that was relevant continued to be relevant, and the dimension that was irrelevant continued to be irrelevant. Even here, the initial and transfer stages did not involve exactly the same categorization. Whenever the relevant dimension was the same during initial and transfer stages, the category labels were reversed from A to B and from B to A, so that any observed transfer could be attributed to changes in selective attention to dimensions rather than acquiring specific stimulus-to-category associations. Thus, the identity transfer condition was equivalent to a "reversal shift" (Tighe & Tighe, 1969) in which

successive categorization rules are based on the same categories, but the labels given to those categories have been reversed.

In the next condition A|C, the same dimension A was relevant in the initial and transfer conditions, but different dimensions were irrelevant. This condition is called the "acquired distinctiveness" condition because if participants are better in the transfer condition when it is preceded by A|C than by the control condition, then it is likely because participants learned to attend the relevant dimension A and this helped them to distinguish subsequent categories that differed on the same dimension. The next condition (C|B) is the complement of this; the irrelevant dimension B was the same between initial and transfer conditions, but the relevant dimension changed. This is called an "acquired equivalence" condition because it tests whether participants can learn to ignore an irrelevant dimension, and if so, whether this learning transfers to ignoring the same dimension when it is later irrelevant. Evidence for learning to ignore irrelevantly varying dimensions would be found if C|B produces better transfer to A|B than does C|D.

If these three conditions produce better transfer than the control condition, it may simply be because their faces are, overall, more similar to the transfer faces due to the shared dimensions. However, it is possible to design conditions that have the same number of overlapping dimensions between the two categorization stages, but with a potentially detrimental effect. With the fourth condition (B|C), the dimension B that was initially relevant becomes irrelevant. This condition can be labeled "attentional capture," making reference to Shiffrin and Schneider's (1977) results showing that when participants are trained to respond to a particular letter as a target, performance is quite poor when that letter later becomes a distracter to be ignored. Transfer performance is predicted to be poorer for B|C than the control if participants continue to attend the formerly relevant dimension even

after it has become irrelevant. The complement of this condition is C|A, in which Dimension A, which was formerly irrelevant, becomes relevant. This situation is reminiscent of negative priming, in which items that were distracters on an earlier trial and then become targets are responded to more slowly than items that were not previously distracters. Participants should be hindered at learning the transfer categorization if they continue to ignore Dimension A because it was previously irrelevant. Finally, both of the transferred dimensions of the negative priming and attentional capture conditions are combined together to create the B|A condition. The irrelevant dimension becomes relevant and the relevant dimension becomes irrelevant. Another way of conceptualizing this condition is that the same 8 faces were used as stimuli in the two categorization phases, but the categorization rule is rotated 90 degrees. For this 90 degree condition, half of the faces from the initial categorization phase received the same category label on the final phase, and half received the opposite label.

The 200 categorization trials within each phase of the experiment were divided into 25 repetitions of each of the eight faces. Within each block of 8 faces, the order in which the faces were presented was randomized. Participants received short breaks every 50 trials. During these breaks, the computer displayed the participants' accuracy and average response time on the previous block of trials. At the end of the first block of 200 trials, participants were explicitly warned that the categories were going to change for the second categorization task, and that they would have to learn new categories that might not have any relation to their previously learned categories.

Results

The average accuracy over all of the initial categorization trials was 73.2%. When transferred to the categorization rule A|B, the average accuracy over all categorization trials was 80.2%. For these transfer categorization trials, participants differed significantly in their categorization accuracy over the seven conditions, $F(6,198) = 6.02$, MSE = 0.0717, $p < 0.001$, as shown in Figure 3. The most appropriate comparison for the first six groups in Figure 3 is the neutral control condition C|D in which none of the faces underlying the stimulus dimensions in the transfer task were used for the initial categorization dimensions. The condition A|B produced better transfer than the control condition, $p<0.001$, (the significance of all post-hoc comparisons are calculated using a two-tailed version of Fisher's LSD), as well as producing better transfer than any other condition. This is not surprising given that for the A|B condition, the initial and final categorizations involve the same faces and dimensions, and only differ in their assignment of labels to faces. The conditions A|C and C|B also produced better transfer to the A|B condition than did the control condition, ($p< .05$ for each comparison), but did not significantly differ from each other. These conditions reveal both categorization advantages when relevant dimensions continue to be relevant (A|C) and when irrelevant dimensions continue to be irrelevant (C|B).

The next three conditions in Figure 3 were designed to reveal potentially negative transfer effects from the initial to final categorization. In Condition C|A, the dimension that was initially irrelevant later became relevant. This condition did not yield significant worse transfer performance than the control condition, $p>0.1$. Compared to the control condition, marginally significantly worse performance for Condition B|C was found, $p=0.054$. Relative to the condition C|A, the performance in condition B|C was not significantly worse, $p>0.1$. The Condition B|A did not significantly differ in its transfer to the A|B categorization relative to the control group, $p>0.1$. Finally, the condition B|A did produce significantly better transfer than did condition B|C, $p <.05$, but did not produce significantly better transfer than condition C|A, $p>.1$.

Discussion

With the exception of one of the seven conditions, the results from Experiment 1 are generally consistent with the role of selective attention in promoting beneficial and detrimental transfer across categorizations. Selective attention can be directed toward arbitrary dimensions created by randomly pairing faces, and once a participant has learned to direct selective attention toward or away from a dimension, this learning extends to new categorizations involving different stimuli. Two conditions that show strong transfer effects (A|C, C|B) do not repeat any of the same faces across initial and transfer conditions, but rather only involve the same dimensions. For example, none of the faces belonging to the A|B set are the same as faces from the C|B set. The only similarity between these sets is that Dimension B is irrelevant for both sets. Thus, our results suggest sensitization of dimensions rather than simply sensitization of particular faces.

Experiment 1 also allowed the contributions of learning to ignore irrelevant and attend relevant dimensions to be separately assessed. Somewhat surprisingly, the positive transfer associated with a shared irrelevant dimension (C|B) is just as strong as the transfer due to a relevant dimension (A|C). In an informal post-experimental interview, most of our participants indicated that during category learning they learned to attend to diagnostic aspects of the faces that allowed them to discriminate between the two categories. No participant spontaneously reported learning to ignore irrelevant dimensions. Yet, given the strong transfer from C|B to A|B, two categorizations that only share an irrelevant dimension, it appears that our participants do in fact learn what dimension not to attend, and that this learning transfers to ignoring the same dimension when it is later irrelevant. Given the equivalent objective transfer for relevant and irrelevant dimensions despite the intuition that learning generally involves attending to relevant dimensions, an interesting speculation is that learning not to attend to irrelevant dimensions is a relatively implicit skill compared to learning to attend to relevant dimensions.

An alternative account for the above positive transfer effects that does not involve learned selective attention at all is that transfer is based on the similarity between initial and transfer faces. That is, participants develop facility with processing particular faces, and show beneficial transfer to subsequent categorizations to the extent that the new faces are similar to the familiarized faces. By this account, the A|B categorization yields the best transfer because the faces are identical across the two categorizations. The A|C and C|B conditions produce better transfer than the control C|D categorization because the faces in these two sets are similar to the A|B faces. This similarity is due to the component dimension that these two conditions have in common with the A|B faces.

However, the conditions testing negative transfer do not support this "transfer based on familiarized faces" account. The conditions C|A and B|C have the same similarity to the A|B faces as the conditions A|C and C|B, but showed no evidence of positive transfer. In fact, if one averages across the four conditions that have one common dimension shared with the A|B condition, accuracy for the A|B condition is not any higher than it is for the control condition ($F(1,176)=0.414$, MSE=0.0055, p>0.1). This suggests that there is no significant transfer based on face similarity per se.

The one condition that is problematic for an account that relies solely on learned selective attention is the B|A condition. The results from this condition seem quite surprising at first. In this condition, the relevant dimension becomes irrelevant and the irrelevant dimension becomes relevant, and yet performance is no worse than it is for the control condition and is significantly better than for the B|C condition. One might have expected performance to be worse for this

condition than it was for either the C|A or the B|C conditions given that it combines both of their disadvantages.

It is for this B|A condition that we believe a dimension differentiation account may be useful. Our explanation for the beneficial transfer, relative to the negative transfer conditions, from B|A to A|B categorizations rests on the observation that they both involve the same set of 16 faces. The categorization rules are orthogonal (separated by 90 degrees), splitting the stimuli horizontally in one case and vertically in the other. As such, both categorization rules depend on separating the horizontal dimension from the vertical dimension in order to selectively attend only one of these dimensions. Effective performance on the A|B categorization requires isolating Dimension A from B, and once accomplished, this may be useful in acquiring the B|A categorization because this categorization also requires the same differentiation of dimensions, albeit for opposite purposes. Although A|B and B|A have opposite requirements as far as selective attention to component dimensions, they are consistent with each other in requiring that Dimensions A and B be isolated from each other. The opposing transfer results predicted from selective attention and from dimension differentiation may in fact both be operating. Negative transfer driven by selective attention may be canceling out most of the positive transfer driven by dimension differentiation, leaving only an insignificant positive transfer from B|A to A|B relative to the control condition.

Experiment 2A

One of the most surprising results from Experiment 1 was that when a formerly irrelevant dimension became relevant and at the same time a relevant dimension became irrelevant, transfer performance was better than when only one of these changes occurred, significantly so in one case. Our account of this result is that there is a tendency for B|A learning

to facilitate an A|B categorization because both categorizations are facilitated by learning to successfully isolate the two Dimensions A and B. Another potential account for the surprisingly good transfer from B|A to A|B is that these conditions involve the same eight faces, and that participants become familiarized with the faces during B|A training, and can then easily learn new assignments to the familiar faces during the second categorization. This account is cast into some doubt by Experiment 1, which found no evidence for transfer based on face similarity. Conditions with one overlapping dimension in common with the transfer task did not produce better transfer, on average, than did the condition with no overlapping dimensions. However, it is possible that the case of exactly identical faces across categorizations is a special case that promotes transfer based on familiarity even though intermediate levels of similarity do not.

In Experiment 2, we controlled for the familiarity of the faces by using the same faces during initial and transfer categorizations for all conditions. We manipulated whether the initial and transfer categorizations encouraged the same organization of stimuli into dimensions, or whether incompatible, cross-cutting dimensionalizations were required. Specifically, we compared a situation where the initial and final categorization rules were separated by 90 or 45 degrees. We created categorization rules for the 8 faces in Figure 2 that divided these faces into two categories either horizontally, vertically, or diagonally. The straight lines in Figure 4 illustrate the four categorization rules used, and also provide two example pairs of initial and final categorization rules. In one example, the initial categorization rule has a horizontal category boundary such that the top four faces belong to Category A and the bottom four faces belong to Category B. In the 45 degree rotation condition, the transfer categorization rule following this horizontal rule would involve a diagonal categorization, in either a forward-slash "/" or backward-slash "\"

form. In the 90 degree rotation condition, the transfer rule would be a vertical category boundary. In the second example, the initial categorization is a diagonal rule such that the upper-right faces belong to Category A and the lower-left faces belong to Category B. The transfer rule in the 45 degree rotation condition would either be vertical or horizontal. The transfer rule in the 90 degree rotation condition would be the diagonal rule facing the opposite direction as the initial rule. The point of the examples is to stress that the difference between the 45 and 90 degree rotations is unrelated to the actual orientation of the line that defines the transfer categorization rule. As shown in Figure 4, whether the transfer diagonal rule counts as a 45 or 90 degree condition depends on the orientation of the initial categorization rule. The difference between the 45 and 90 degree conditions involves the relation between two categorizations, not the absolute orientation of a categorization rule.

The 45 and 90 degree rotation conditions have complementary costs and benefits with respect to transfer to the final categorization. The conditions can be considered with respect to transfer based on selective attention demands and dimensional organization. Considering selective attention demands, two categorizations are compatible to the extent that the weights given to the dimensions for optimal categorization efficiency are similar. Categorizations related by 90 degrees are completely incompatible as far as selective attention demands because for one categorization Dimension X should receive an attention weight of 0 and Dimension Y a weight of 1, and for the other categorization X should receive an attention weight of 1 and Y a weight of 0. The average difference between attentional weights would be 0 if two categorizations were completely compatible. Categorizations related by 45 degrees are partially compatible by selective attention, because for one categorization Dimension X should receive a weight of 0 and Dimension Y a weight of 1, and

for the other categorization both Dimensions X and Y should receive weights of 0.5. These attentional similarity relations depend only on the relation between the two categorization rules, holding irrespective of the rules' specific orientations.

Considering dimensional organization, two categorizations are compatible to the extent that the dimensions that they require are independent and therefore able to co-exist. For example, in terms of dimensional organization, categorizing rectangles on the basis of height is compatible with categorizing them on the basis of width because these two dimensions can each be separately registered and do not interfere with each other. Someone who thought about rectangles in terms of height would also be likely to think about them in terms of width. Organizing rectangles in terms of shape (ratio of width to height) and area is an alternative dimensional organization, as shown in Figure 5. A person who thinks in terms of rectangle shape (e.g. the ratio of the height to the width) might also be expected to think in terms of area because this is the remaining dimension along which rectangles vary once shape has been extracted. However, organizing rectangles in terms of height is incompatible with organizing them in terms of area because area is partially dependent on height. In fact, results have been equivocal on whether rectangles are properly viewed as being psychologically represented in terms of height and width, or in terms of area and shape (Feldman & Richard, 1998; Shoenemann, 1977). However, researchers agree that if a person tends to view rectangles in terms of height, then they would also view rectangles in terms of width rather than area or shape. In assessing the compatibility of dimensional organization, two dimensions are compatible if they involve independent sources of variation, and are incompatible if they do not. If dimensions involve related sources of variation, then some of the variation that is accounted for by one dimension need not be accounted for by the other dimension, making

that second dimension less pertinent to the task of fully describing a stimulus set.

If the only thing that participants learn during the initial categorization is how much to selectively attend to pre-existing and pre-differentiated dimensions, then we would expect the 45 degree rotation to produce better transfer to the final categorization than the 90 degree rotation. In the 45 degree rotation condition, the dimension that is relevant during the initial categorization is partially relevant for the final categorization, and similarity in terms of attentional weights is 0.5. In the 90 degree condition, the dimension relevant during the initial categorization is completely irrelevant for transfer, and the similarity in terms of attentional weights is 0. However, in terms of dimensional organization, the 90 degree rotation may show better transfer, because the initial and transfer categorization rules are compatible in promoting the same differentiation of dimensions. In half of the categorization conditions, this differentiation will imply an organization of stimuli into horizontal and vertical dimensions. In the remaining conditions, this differentiation will imply an organization of the stimuli that cuts across this organization, organizing the faces into two dimensions that are separated by 45 degrees from the horizontal-vertical dimensionalization. Both dimensional descriptions are possible ways of viewing the set of stimuli and it is likely that neither is *a priori* privileged due to the arbitrary manner in which the dimensions were originally created.

Method

Participants. 131 undergraduate students from Indiana University served as participants in order to fulfill a course requirement. The number of students in the 45 and 90 degree rotation conditions were 65 and 66 respectively.

Procedure. The stimuli were faces constructed in the same manner as in Experiment 1. A 4 X 4 matrix of faces was obtained by blending values on two dimensions, with each dimension defined by two faces. The four faces selected as dimension endpoints were chosen to be roughly equally similar to each other in a multidimensional scaling solution (Goldstone, 1994-b). From this 4 X 4 matrix, the eight faces shown in Figure 4 were selected as stimuli.

Experiment 2A used the same general procedure as was used in Experiment 1. Participants received two categorizations that were related to each other by a 45 or 90 degree rotation. During each trial of a categorization task, participants saw a face, guessed its category ("club membership") and received feedback from the computer indicating their accuracy. After the first categorization task was completed, participants were warned that the categories had changed and that they would have to learn new categories in the second phase of the experiment. Each phase contained 56 trials, consisting of 7 repetitions of eight faces.

For the initial categorization phase, one of four categorization rules was randomly chosen for each participant: horizontal, vertical, forward diagonal, or backward diagonal. The category boundary line split the eight faces into two categories, and the assignment of the two sets to Categories A and B was randomized. In the transfer categorization, the category boundary was rotated by either 45 or 90 degrees, depending upon the participant's condition. Whether a rotation was counter-clockwise or clockwise was randomly determined. Thus, the 90 degree rotation of an initial diagonal boundary was always a diagonal boundary facing the other direction, and the 45 degree rotation was either a horizontal or vertical boundary. Assigning the two sets to category labels was again randomized. As such, in the 90 degree rotation condition, four out of the eight faces received the same label in the initial and transfer phases, and the remaining four faces received a label of "Category A" during one phase and "Category B" during the other phase. However, in the 45 degree rotation condition, there are two equally probable ways that the labels may be assigned. In the first way, six out of eight faces received the same label during the two phases. In the second way, two out of eight faces received the same label.

Results

Categorization accuracy for the two groups did not significantly differ in the initial categorization phase of the experiment, $F(1, 129) = 0.744$, MSE = 0.0342, p > 0.1. The primary result of interest concerns the categorization accuracies for the transfer portion of the experiment, and are shown on the left side of Figure 6. For comparison purposes, a summary of the results from Experiments 2-4 are shown in Table 2. As described in the methods section, rotating a category rule by 45 degrees produced categorizations with either two or six labels in common across initial and transfer conditions. However, with either method of assigning categories to labels, transfer performance was worse for the 45 degree condition than the 90 degree condition. Overall, the 90 degree condition produced better transfer than the 45 degree condition, $F(1, 129) = 4.013$, MSE = 0.185, p < 0.05. Within the 45 degree condition, the average categorization accuracies when two and six labels were in common across categorizations were 59.7% and 63.2% respectively, which did not significantly differ, T (63)=0.97, p > .1. A second way in which common category assignments across the categorization phases might influence accuracy is through assignments to specific items rather than through the overall number of shared category assignments. When an item received the same category assignment across the two phases of the experiment, the average categorization accuracy for it was 66.0% in the second phase, and this was significantly different from the 61.1% categorization accuracy for an item the received different category assignments, paired T (1,130)=2.3, p < .05.

Initial categorization performance was not significantly better for horizontal and vertical boundaries (63.5%) than diagonal boundaries (66.3%), $F(1, 130)=1.48$, MSE = 0.0248, p > .10. These accuracy rates are significantly below what they were in Experiment 1, but this difference may largely be due to the relatively small number of trials per categorization phase. Across the two categorization phases, there was no significant difference in categorization accuracy for the 90 degree rotation condition, T(65)= 0.39, p>0.1 and for the 45 degree condition, categorization accuracy decreased from the initial to transfer condition, although this effect was only marginally significant, T (64)=1.65, p=.10.

Discussion

Transfer across categorizations was better when the categorizations were related by 90 rather than 45 degrees. This is a surprising result from the perspective of learned selective attention, given that categorizations that are related by 45 degrees partially overlap in the dimensions that are relevant, whereas there is no overlap at all in the 90 degree condition. It would, however, be premature to conclude that our results are evidence against Thorndike's (1903) principle of common elements, which states that there is positive transfer between skills to the extent that the two skills involve the same procedural elements. Our results are consistent with transfer by common elements if one includes elements that extend beyond selective attention processes. More specifically, our results are reconciled with transfer by common elements if the notion of dimension differentiation is incorporated. Categorizations that are separated by 90 degrees encourage the same isolation of dimensions from each other, whereas categorizations that are separated by 45 degrees require cross-cutting dimensional organizations.

An account of the advantage of 90 over 45 degree transfer in terms of dimension differentiation could take two forms. One form emphasizes the hindrance caused by inconsistent dimensional organizations from initial to transfer phases. The other form emphasizes the facilitation caused by congruent dimensional organizations. Our results lend more support to the hindrance account. Relative to initial categorization performance, transfer performance was worse in the 45 degree rotation

condition, and was not significantly better in the 90 degree rotation condition. As such, when dimensional organizations are incompatible with each other across tasks, negative transfer occurs. Furthermore, this negative transfer is apparently strong enough to counteract the presumed positive transfer from the partially shared relevant dimensions in the 45 degree rotation condition.

An important result from Experiment 2A was that the diagonal categorization rules were not any harder than the vertical or horizontal rules. This confirms that our morphing technique produced genuinely arbitrary dimensions. In situations where stimuli have privileged dimensions, categorization is faster and more accurate when the categorization rule is orthogonal to these dimensions rather than at a 45 degree angle (Grau & Nelson, 1988; Kruschke, 1993; Melara et al., 1993). Our results do not reveal a similar pattern of privileged dimensions. This is also consistent with our participants' informal reports. They did not experience a clear dimensional organization to the faces. We seem to have successfully developed a set of materials that are intrinsically ambiguous as far as their dimensional structure. If the eight faces are grouped according to a diagonal categorization rule, then participants apparently do not show a bias to interpret them as being composed out of the horizontal and vertical dimensions shown in Figure 4. Rather, the perceived dimensional organization is influenced by the categorization rule itself. This pattern of results contrasts markedly to what is found with stimuli composed out of clearly delineated, separable dimensions. For these materials, diagonal categorization rules are notably more difficult than horizontal or vertical rules (Ashby & Maddox, 1994; Kruschke, 1993), as one might expect if the diagonal categorization rules involve integration across two dimensions that are normally separated while the horizontal and vertical rules involve only one dimension. The failure to find this difference implies that our

diagonal categorization rules are not more complex than the vertical and horizontal rules. The dimensional status of horizontal, vertical, and diagonal axes is equal prior to categorization training. Although we, as experimenters, can provide simpler descriptions for the horizontal and vertical axes than for the diagonal axes, these descriptions are apparently not serving as the basis for participants' categorizations (see also Cheng and Pachella, 1984).

Experiment 2A lends support to Experiment 1's conclusion that dimension differentiation may explain the surprisingly good transfer from B|A to A|B. In supporting this conclusion, the experiment eliminates an explanation of the transfer (more precisely, the lack of negative transfer) that is based on stimulus familiarity. All of the faces in the second phase of Experiment 2 were exposed to participants in the initial phase, regardless of the participants' transfer condition. Furthermore, accounts in terms of transfer by common labels can also not fully explain Experiment 2. Better transfer was observed in the condition that preserved 4 category assignments (the 90 degree condition) than in conditions that preserved either 2 or 6 category assignments.

Experiment 2B

Experiment 2A found better transfer between category rules related by 90 than by 45 degrees, a surprising fact from the perspective of selective attention given that rules related by 90 degrees are completely non-overlapping in their attentional requirements whereas rules related by 45 degrees partially agree on what dimensions are relevant and irrelevant. Our account for the advantage of a 90 over 45 degree rule rotation is that only the former rotation encourages a compatible dimensional interpretation of the stimuli. If the good transfer in the 90 degree rotation condition is because the two categorizations encourage the same differentiation of the faces into dimensions rather than cross-cutting dimensions, then we

should not expect the 90 degree rotation condition to produce better performance when dimension differentiation is not required -- that is, with more separable stimuli. With more separable dimensions, participants should be able to selectively attend one dimension without much need to differentiate fused dimensions. If dimension differentiation was the cause of the relatively good transfer observed in the 90 degree rotation condition, then using relatively separable dimensions should reduce or eliminate the advantage of the 90 degree rotation condition. Experiment 2B explored this prediction by testing whether the 90 degree rotation advantage can be eliminated by using stimuli with more cleanly delineated dimensions.

Experiment 2B used the same face stimuli of the previous experiments, but generated more separable dimensions by isolating and morphing specific face parts. Dimension A morphed from the mouth of one face (call him "Joe") to the mouth of another face ("David"). Dimension B morphed from the eyes of Joe to the eyes of David. Thus, the two dimensions involved the appearance of separated face parts rather than superimposed, overlapping whole-face dimensions. If these stimulus dimensions are more differentiated than the whole-face dimensions of the previous experiments, then we would not predict as large an advantage of the 90 over 45 degree transfer.

Figure 7 depicts the stimuli used in Experiment 2B, in which the mouth (vertical dimension) or eyes (horizontal dimension) of one face were selectively morphed into the mouth or eyes of another face while preserving the rest of the face features. The face part is seamlessly joined with the rest of the face using a morphing technique described by Steyvers (1999). With this technique, faces can be constructed that vary on two spatially separated dimensions. Such stimuli allow us to compare perceptual categorization tasks that involve overlapping (Experiment 2A) and separated (Experiment 2B) dimensions, controlling for the nature of the stimuli. Although mouth and eye features may be processed configurally (Farah, 1992), a study using Garner interference

(Garner, 1976, 1978) from our laboratory has revealed that the dimensions of Figure 4 are more strongly integral than are the mouth and eye dimensions of Figure 7.

The logic of the Garner interference measure of integrality is that if two dimensions are relatively integral, then it will be relatively difficult to attend to one of the dimensions while ignoring irrelevant variation on the other dimension. This is operationalized by comparing the efficiency of two categorizations – one which requires a binary classification along a single dimension (the control task), and one which requires a binary classification along a single dimension while irrelevant variation exists along on a second dimension (the filter task). When dimensions are relatively separable, the filter and control tasks are approximately equally fast and accurate. When dimensions are integral, the filter task is much more difficult than the control task, consistent with the notion that participants have a hard time ignoring variation along an irrelevant dimension if the two dimensions are psychologically fused. With the face dimensions shown in Figure 4, categorization responses based on one of the dimensions increased from 740 to 1125 milliseconds when irrelevant variation on the other dimension was introduced, whereas for the mouth and eye dimensions of Figure 7, variation on an irrelevant dimension only slowed responses from 717 to 778 milliseconds. Thus, based on response times for categorizations, it is much more difficult for participants to selectively attend to the dimensions of Figure 4 than Figure 7. The logic for Experiment 2B does not require us to claim that faces are perceived as composed out of separable dimensions. Indeed, a large body of research suggests that faces are perceived configurally. Face parts influence the perception of other face parts, and selective attention to individual face parts is difficult (Farah, 1992; Tanaka & Farah, 1993). However, our claim, supported by the observed interference between dimensions, is that prior to extensive training, the dimensions of Figure 7 are more isolated and pre-differentiated for our participants than are the dimensions of Figure 4.

Method

Participants. 120 undergraduate students from Indiana University served as participants in order to fulfill a course requirement. The

number of students in the 45 and 90 degree rotation conditions were 62 and 58 respectively.

Procedure. The face dimensions were constructed in the same manner as in Experiment 2A, but were composed of spatially separated, non-overlapping face parts. Thus, a subset of 8 faces was selected from a 4 X 4 matrix of faces varying on two dimensions. A standard face was generated by blending the two original faces shown on the sides of Figure 7. These were two of the same faces (Faces 3 and 4) used to generate the stimuli in Experiment 2A. Then, dimensions were created by selectively altering the mouth or eye regions of the standard face. Four values of the mouth dimension were used: 100% Face 3, 67% Face 3 and 33% Face 4, 33% Face 3 and 67% Face 4, and 100% Face 4. The analogous values were used in constructing the vertical, eye dimension. From the 4 X 4 matrix of faces, the 8 faces that were chosen to be stimuli all had a 67% contribution of Face 3 or 4 along the two dimensions. Thus, the faces are arranged in the octagonal form shown in Figure 7.

The same two-phase category learning procedure used in Experiment 2A was used again. Initial categorization rules were selected at random from one of four possibilities: horizontal, vertical, diagonal-/, and diagonal-\. Final categorizations were again related to these initial rules by either 45 or 90 degrees.

Results

Like Experiment 2A, categorization accuracy varied as a function of the form of the categorization boundary. Unlike Experiment 2A, the initial diagonal category boundaries were significantly harder (62.6% accuracy) than the horizontal and vertical boundaries (69.7%), T (118)= 3.086, $p < 0.01$. As such, in addition to the results cited for Garner interference, the categorization results gave independent grounds for believing that the two dimensions of Experiment 2B were more separable than those of Experiment 2A. Consistent with previous studies (e.g. Kruschke, 1993; Shepard, Hovland, & Jenkins, 1961), when two dimensions are

psychologically separable, it is difficult to attend to both dimensions while learning a categorization.

Overall, categorization accuracy did not vary significantly as a function of transfer condition, with the 45 and 90 degree rotation conditions producing 65.2% and 65.0% accuracies, respectively, on the final categorization, T (118) = 0.082, $p > .5$. Going from the initial to transfer categorizations, there was no significant difference in categorization accuracy for the 90 degree rotation condition, T(57)= 0.282, $p>0.5$, or for the 45 degree condition, T (61)=1.41, $p>.10$. As with Experiment 2A, we conducted a more detailed analysis by considering how many items received the same category assignment across the initial and transfer conditions. This number must always be 4 out of 8 in the 90 degree rotation condition, but will be either 2 or 6 in the 45 degree rotation condition. Final categorization accuracy did not significantly depend on the number of unchanged category assignments, F(2, 119)=0.039, MSE=0.00073, $p > 0.5$. Similarly, categorization accuracy was not significantly different for items that received the same category assignment across the two phases relative to items that received different assignments, paired T(1, 118)=0.4, $p> 0.1$.

Discussion

Experiment 2B eliminated some possible counter explanations of Experiment 2A that do not involve the notion of compatible and incompatible dimensional organizations. In Experiment 2B, an advantage of the 90 over 45 degree rotation condition was not found. The faces of Experiments 2A and 2B were highly similar. The major difference between the experiments was that the experimental dimensions used to construct the faces in Experiment 2B were more psychologically separable than those used in Experiment 2A. This difference was confirmed both by a pilot experiment showing greater Garner interference for the dimensions used in Figure 4 than Figure 7, and also by the significantly worse

categorization accuracies for rules that were diagonal rather than orthogonal to the dimensional axes for Experiment 2B but not Experiment 2A. For Experiment 2B, the horizontal and vertical axes shown in Figure 7 were psychologically privileged axes, and categorization rules involving rotations of these axes were difficult to learn.

Dimensional separability is predicted to influence categorization accuracies for the 45 and 90 degree rotation conditions. If the advantage of 90 over 45 degree rotations observed in Experiment 2A is due to dimension differentiation, then the advantage should not persist if dimension differentiation is not required. Exactly this result was obtained in Experiment 2B. The dimensions used in Experiment 2B were not completely separable, but they were apparently more separable than the dimensions of Experiment 2A.

The lack of an advantage of the 90 degree rotation condition in Experiment 2B argues against general learning accounts of Experiment 2A that do not depend on the nature of the dimensions. For example, it is possible to give post-hoc accounts of Experiment 2A whereby if categorization rules are too similar (2 identical category assignments) or too different (6 identical category assignments) then transfer performance suffers. It might be that if initial and final rules are too similar, participants have a tendency to continue using their imperfect initial rule during transfer. If the two rules are too dissimilar, then participants might either try to reverse the rule or might generally give up on the possibility of transferring anything learned from the initial phase. Apart from the inelegant and post-hoc nature of this nonmonotonic account, it also does not correctly predict the outcome from Experiment 2B. General learning accounts for the results from Experiment 2A that depend solely on the nature of the initial and final rules cannot account for the discrepancy between Experiments 2A and 2B.

Our results suggest that the degree of transfer between categorization depends not just on the relation between the categorization rules, but also on the nature of the dimensions underlying the stimuli. Experiments 2A and 2B are the first results to our knowledge that suggest that the strength of transfer from one categorization to another depends on the perceptual integrality of the stimulus dimensions. Such a dependency is predicted if transfer between categorizations is based not only on their similar requirements for selective attention, but also on the compatibility of their dimensional organizations. Categorization rules related by a 45 degree rotation are incompatible in the dimensional organizations that they require. Rules related by 90 degrees are incompatible in terms of their selective attention requirements, but are compatible in encouraging the same isolation of one dimension from another. This compatibility in terms of dimensional organization is expected, and found, to be of particular importance when the dimensional organization of the category members is ambiguous.

Experiment 3

The dimension differentiation process observed in Experiment 2A could be directed by either supervised or unsupervised information. Unsupervised information is provided by the stimuli themselves, and their statistical properties across the entire set of stimuli. For example, if stimuli such as AAQ, ARA, SAA, BBT, BUB, VBB are presented, even without any information about the categories from which these stimuli are drawn, one might deduce that there are two categories and that the first three objects belong to one category and the last three objects belong to the other category. The basis for this induction is present in the stimuli themselves, given their natural clustering into two categories. Previous research has shown that people do in fact form categories on the basis of such unsupervised information (Clapper & Bower, 1994). Supervised information is the feedback given to a person once they have categorized the object. When

different patterns of perception occur when the same stimuli are used but are categorized according to different rules, then we can conclude that supervised feedback is affecting perceptual change (Goldstone, 1994-a; Goldstone et al., 2000).

Experiment 2A provides suggestive but not definitive evidence that dimension differentiation is caused by supervised rather than unsupervised information. In this experiment, all of the categorization groups were presented with the same eight faces, and better transfer was found between categorizations related by a 90 degree rotation. However, the stimuli themselves weakly suggest a dimensional organization, because each face shares exactly the same value along the horizontal and vertical dimensions of Figure 4 with one other face. If people were sensitive to these shared identical values on a dimension and if they assumed that the underlying dimensions of the stimuli should have identical values for different stimuli, then even without any category feedback, they would be able to reconstruct the horizontal and vertical dimensions in Figure 4. This cannot explain the results from Experiment 2A because we found dimension differentiation for diagonal as well as horizontal and vertical categorization rules separated by 90 degrees. Still, the stimuli in Figure 4 do potentially have privileged dimensions, and it is possible to create equivalent face sets that do not. As such, the purpose of Experiment 3 was to replicate Experiment 2A's results using stimuli with no unsupervised information that could bias the formation of dimensional structures. The faces for Experiment 3 were created by using dimension values that were circularly arranged. The octagonal structure of Experiment 2A approximates such a circular arrangement, but Experiment 3 used trigonometric relations to more precisely capture the circular arrangement. With such an arrangement, any of the eight linear categorization rules should produce equally difficult categorizations, and there is no bias toward one of these categorizations from unsupervised information.

Method

Participants. 161 undergraduate students from Indiana University served as participants in order to fulfill a course requirement. The students were randomly assigned to the 45 and 90 degree conditions.

Materials. The stimuli were faces that were generated by morphing between the same photographs used in Experiment 2A. The set of 16 faces that all participants saw are shown in Figure 8. These faces, rather than being selected from a 4 X 4 matrix of faces as was true in the previous experiments, were arranged along a circle. The horizontal and vertical dimensions in Figure 8 both represent the relative proportion of two faces, similar to the axes shown in Figure 2. A variable D was created that was assigned 16 different values for degrees from 0 to 360 in 22.5 degree steps. For each value assigned to D, the vertical dimension value for a face was equal to cosine(D) and the horizontal dimension value was sin(D). In this manner, we can be assured that the statistical structure of the set of faces does not itself suggest a dimensional organization. The circularly arranged stimuli have no privileged dimensional axes until the categorization feedback is given. In all other respects, the appearance of the faces was identical to that used in Experiment 2A.

Procedure. The experimental procedure closely followed that of Experiment 2. For each participant, a randomly constructed initial category boundary was selected by choosing a random number between 1 and 16 inclusively. This number determined the first face from Figure 8 that belonged to Category A. The block of eight faces, starting with this first face and going around the circle clockwise, were all assigned to Category A, and the remaining faces were assigned to Category B. For the participants in the 45 degree condition, the final categorization rule shifted the initial categorization clockwise by two faces. For the

participants in the 90 degree condition, the final categorization rule shifted the initial categorization boundary clockwise by four faces. For both the 45 and 90 degree conditions, the category labels were swapped from "A" to "B" and from "B" to "A" for half of the participants, randomly determined. Thus, the 90 degree condition always led to eight faces receiving the same categorization and eight receiving the opposite categorization across the two category learning stages, whereas the 45 degree conditions led to either four or twelve faces receiving the same categorization across the two stages.

During each of the two category learning stages, participants were given 4 repetitions of the 16 faces. The display durations, feedback, and randomizations were identical to that of Experiment 2. Participants received rest breaks after every 32 trials.

Results

Categorization accuracy for the 45 and 90 degree conditions did not significantly differ in the initial categorization phase of the experiment, and were 58.6% and 59.4% respectively, $F(1, 160) = .177$, MSE $= 0.00253$, $p > 0.5$. The primary result of interest concerns the categorization accuracies during the transfer portion of the experiment. Overall, categorization performances in the second stage of the experiment were 53.7% and 58.9% for participants in the 45 and 90 degree rotation conditions, respectively, $F(1, 160) = 10.25$, MSE $= .113$, $p < 0.01$. Going from the initial to transfer categorizations, there was no significant difference in categorization accuracy for the 90 degree rotation condition, $T(78) = 0.135$, $p > 0.5$, and for the 45 degree condition categorization accuracy was significantly lower in the transfer phase, $T(81) = 3.75$, $p < .001$. Transfer in the 90 degree condition surpassed either of the labeling conditions of the 45 degree condition. Within the 45 degree condition, categorization accuracies were 53.6% and 53.5% when four and twelve faces received the same categorization, $T(80) = 0.016$, $p > .5$.

Categorization accuracy was not significantly different for items that received the same category assignment across the two phases relative to items that received different assignments, paired $T(1, 160) = 0.3$, $p > 0.1$.

For both the initial and transfer categorizations, and both the 45 and 90 degree conditions, there was a strong influence on accuracy of the distance of a face from the category boundary, $F(3,643) = 20.39$, MSE $= .274$, $p < 0.001$, but no higher order interactions involving distance from category boundary. Overall, for faces that were located one, two, three, and four positions away from the boundary, categorization accuracies were 52.6%, 56.3%, 59.6%, and 62.0% respectively.

Discussion

Experiment 3 replicated the results from Experiment 2A, the other experiment that involved arbitrary and completely overlapping dimensions. In both experiments, performance on the transfer categorization was better when its category boundary was related by 90, rather than 45, degrees to an earlier categorization. In Experiment 3, the circular arrangement of the faces prevents any source of information other than categorization feedback from producing the superior transfer performance in the 90 degree condition. The advantage of the 90 degree over 45 degree condition in Experiment 3 was just as large as the difference found in Experiment 2A, suggesting that the small amount of unsupervised statistical information present in Experiment 2A was not responsible for the superiority of the 90 degree condition in this experiment either. This conclusion is consistent with our participants' and our own introspection that it is virtually impossible to tell whether two faces have the same value on one of the arbitrary dimensions in Figure 4.

The advantage of the 90 degree over 45 degree condition was apparently largely due to negative transfer caused by the 45 degree condition rather than positive transfer in the 90 degree condition. This is suggested by the significantly worse categorization performance

in the transfer than initial categorization stage for the 45 degree condition. Negative transfer in the 45 degree condition was also found in Experiment 2A.

Experiment 4

Until now, the dimensions under inquiry have been dimensions created by blending arbitrarily paired faces in different proportions. As indicated in Experiment 2, these dimensions appear to pass the Garner tests for integrality. For example, responding to one of two dimensions is slowed by the presence of irrelevant variation along the other dimension. However, we would have a more compelling case for dimension differentiation if the results from Experiment 3 could be replicated with other materials. There are at least four reasons why a replication is desirable.

First, results from previous experiments were statistically reliable, but weak. One reason for relatively small effect sizes may be that overall categorization accuracy was fairly poor. For example, in Experiment 3, categorization accuracy in the initial category learning phase was only 59%. Evidence for strong transfer between categorization conditions is only expected if substantial learning occurs during the initial categorization phase. Second, if dimension differentiation exists, it should be possible to obtain evidence for it using classically integral dimensions. While the dimensions of Experiments 2A and 3 are apparently integral, far more empirical evidence exists that saturation and brightness are integral (Foard & Kemler, 1984; Nosofsky & Palmeri, 1996; Shepp, Burns, & McDonough, 1980). Third, the face-based dimensions used in the previous experiments may themselves have components. Participants who learn to accurately categorize faces grouped by an arbitrary dimension may not be basing their judgments on the entire dimension. Instead, they may be responding to a local region of the face that covaries with the dimension. Some of the local regions may not have arbitrary

characterization, and in fact may correspond to pre-existing dimensions such as "happiness," "age," and "adiposity." This is true of any morph-based dimension given that many parts simultaneously change as the dimension value changes. As such, we can only conclude that participants learn to differentiate between the morph-based dimensions or between correlates of these dimensions. Fourth, given that several authors have suggested that face processing may be an important special case of perception with unique processes associated with it (Farah, 1992), it is important to know whether the dimension differentiation observed with face dimensions generalizes to other objects, or whether it is a peculiarity restricted to faces. Face dimensions were used because of the integral nature of facial dimensions, but faces are not the only materials that possess this property.

Experiment 4 used the same circular arrangement of stimuli used in Experiment 3, with 45 and 90 degree transfer conditions. In order to address the four points above, the morph-based dimensions were replaced with two classic examples of integral dimensions, saturation and brightness. Saturation is related to the amount of white mixed into a color, and brightness is related to the amount of light coming off of a color, but to most people they are fused together to form a unified perception of color. If the results from Experiment 3 replicate in Experiment 4, then this is evidence that the advantage found for dimensionally consistent categorizations does not depend on a special property of faces per se. Unlike morph-based dimensions, saturation and brightness do not involve correlated changes across many different kinds of parts. People may examine the saturation or brightness of a small region of a stimulus, but an accurate assessment requires the same kind of evidence anywhere on the stimulus, as long as the stimulus has a uniform color. To increase the overall accuracy of participants' categorization, they were given a

larger number of trials in both the initial and final categorization phases.

Method

Participants. Fifty-six undergraduate students from Indiana University served as participants in order to fulfill a course requirement. The students were randomly assigned to the 45 and 90 degree conditions, with 28 participants in each group.

Materials. The stimuli were squares with slightly different colors. Sixteen colors were created in a circular arrangement, with approximately equal similarity between every pair of neighboring colors. A variable D was created that was assigned 16 different values for degrees from 0 to 360 in 22.5 degree steps. For each value assigned to D, the brightness dimension value for a square was equal to cosine(D) and the saturation dimension value was sin(D). These values were passed to internal routines on an Apple computer (PowerMac 7200) for color production. The saturation and brightness values for the sixteen colors were subsequently confirmed with a Spectra Scan 714 chromometer. The resulting Commission Internationale de L'Eclairage (CIE) 1976 model color coordinates are presented in Table 3. Although the resulting saturation and brightness values no longer positioned on a perfect circle once converted to CIE coordinates, the average deviation from values that would lie on a circle in CIE space was less than 5%.

Each color was presented on a 6 cm^2 square patch with a red hue with primary wavelength of 680 nm. The colors were presented in darkened cubicle illuminated only by a 7.5 watt overhead light bulb. Colors were presented on a white screen background.

Procedure. The experimental procedure closely followed that of Experiment 3. For each participant, a randomly constructed initial category boundary was selected by choosing a random number between 1 and 16 inclusively. This number determined the first colors from the

circularly arranged colors that belonged to Category A. The block of eight colors, starting with this first colors and going around the circle clockwise, were all assigned to Category A, and the remaining colors were assigned to Category B. For the participants in the 45 degree condition, the final categorization rule shifted the initial categorization clockwise by two colors. For the participants in the 90 degree condition, the final categorization rule shifted the initial categorization boundary clockwise by four colors. For both the 45 and 90 degree conditions, the category labels were swapped from "A" to "B" and from "B" to "A" for half of the participants, randomly determined. Thus, the 90 degree condition always led to eight colors receiving the same categorization and eight receiving the opposite categorization across the two category learning stages, whereas the 45 degree conditions led to either four or twelve colors receiving the same categorization across the two stages.

During each of the two category learning stages, participants were given 15 repetitions of the 16 colors, yielding 240 trials. The display durations, feedback, and randomizations were identical to that of Experiment 3. Participants received rest breaks after every 48 trials.

Results

Categorization accuracies for both initial and final phases of the categorization are shown in Figure 9. Categorization accuracies did not significantly differ in the initial categorization phase of the experiment, and were 79.1% and 80.2% for participants in the 45 and 90 degree rotation conditions respectively. Categorization accuracies in the second stage of the experiment were 74.4% and 82.1% for participants in the 45 and 90 degree rotation conditions respectively, $F(1, 54) =9.2$, MSE = .083, $p < 0.01$. Going from the initial to transfer categorizations, there was not a significant difference in categorization accuracy for the 90 degree rotation condition, $F(1,54)= 0.54$, MSE=0.0057, $p>0.1$, and for the 45 degree condition categorization accuracy was marginally significantly lower in the transfer

phase, F (1,54)=2.98, MSE=0.03, $\underline{p}$=.09. Transfer in the 90 degree condition surpassed either of the labeling conditions of the 45 degree condition. Within the 45 degree condition, categorization accuracies were74.1% and 74.8% when four and twelve faces received the same categorization, F(1,80) =0.016, MSE=0.011, $\underline{p}$ >.5. Categorization accuracy was not significantly different for items that received the same category assignment across the two phases relative to items that received different assignments, paired T(1, 80)=0.5, $\underline{p}$> 0.1.

For both the initial and transfer categorizations, and both the 45 and 90 degree conditions, there was a strong influence on accuracy of the distance of a face from the category boundary, F(3, 220) = 71.3, MSE = 0.972, $\underline{p}$ < 0.001, but no higher order interactions involving distance from category boundary. Overall, for faces that were located one, two, three, and four positions away from the boundary, categorization accuracies were 61.0%, 79.6%, 88.2%, and 89.6% respectively during the initial categorization, and were 60.8%, 74.4%, 85.8%, and 88.0% during the second categorization.

As with Experiment 3, categorization rules based on saturation or brightness were not significantly easier than rules based on rotated versions of these dimensions. From the initial categorization phase, we identified four categorization rules that were predominantly based on saturation, four based on saturation, and eight that were predominantly based on 45 degree rotations of saturation and brightness. The average categorization accuracy for saturation and brightness rules (79%), was not significantly different from categorization accuracy involving 45 degree rotations of these rules (81%), F(1, 49), MSE = 0.0098, $\underline{p}$>0.1. As such, our results do not suggest that brightness and saturation are privileged axes of dimensional organization for colors.

Discussion

Using classic examples of integral dimensions rather than face-based dimensions,

the results from Experiment 4, shown in Figure 9, largely replicated the major trends found in Experiment 3. Specifically, better transfer was found for categorization rules related by 90 rather than 45 degrees. As with the earlier experiments, the difference between the conditions seems to be at least as much due to difficulties in the 45 degree condition as to beneficial transfer in the 90 degree case. That is, the final categorization rule in the 45 degree condition was learned significantly less successfully than was the initial rule, even though the initial and final categorization rules sampled equally from the same set of 16 rules for defining category boundaries. This result is consistent with the notion that categorization rules related by 45 degrees are incompatible in the dimensional organizations that they require, and consequently negative transfer exists between them. Rules related by 90 degrees require very different selective attention to dimensions, but can be successively learned without changing the essential description of stimuli in terms of dimensions.

General Discussion

The reported experiments explored how the processing of perceptual dimensions is adapted to category learning tasks. The source of evidence for adaptation was participants' performance on a category learning task that was preceded by different initial categorization tasks. The results provide support for both the sensitization and differentiation of dimensions. For both mechanisms, the dimensions were created by arbitrarily pairing and blending faces. As such, these mechanisms can be observed even when the relevant dimensions were not easily described and *a priori* , as they are in the case of brightness, size, and orientation. By sensitization, participants learn to selectively attend to dimensions based on their category relevance. Extending beyond previous demonstrations of category-induced selective attention (Kruschke, 1992; Nosofsky, 1986), the transfer conditions of the current experiments

distinguish between sensitization of relevant dimensions and desensitization of irrelevant dimensions. Evidence for both effects were found, and were found to be equally strong in the positive transfer conditions.

Results from all four experiments suggest that selective attention is not the only process of adaptation. In addition to learning to selectively attend to dimensions, participants also apparently learn to differentiate between dimensions in the first place. Differentiation of dimensions involves the isolation of dimensions, but is not equivalent to selectively attending to these dimensions. Rather, isolation of dimensions precedes efficient selective attention. Some selective attention may be possible to integral dimensions such as saturation and brightness, given that the best fitting categorization model requires relatively high attention weight for the relevant dimension (Nosofsky, 1987). On the other hand, Nosofsky and Palmeri (1996) find that with integral dimensions, even when a categorization rule specifies the relevance of only one of the dimensions, the best fitting attention weights for the two dimensions are approximately equal. A partial reconciliation of these results is that selective attention to one of a pair of integral dimensions may be possible, but it is not highly efficient. Even in the Nosofsky (1987) study, the attention weight given to the irrelevant dimension was not zero, as would be expected if the dimension could be completely ignored. To selectively attend to Dimension X and ignore Dimension Y with efficiency, one must be able to isolate Dimension X in the stimulus. Once isolated, one can then choose to either attend or ignore it. In Experiments 2 and 3, conditions expected to produce positive transfer via dimension differentiation produced better transfer than conditions expected to produce positive transfer via selective attention, but only when stimuli were composed of highly integral, overlapping dimensions. For such materials, particularly strong positive transfer was found in situations in which originally Dimension X was

relevant and Dimensions Y was irrelevant, and subsequently Dimension X was irrelevant and Dimension Y was relevant. This positive transfer can be explained in terms of learning to isolate Dimensions X and Y from each other.

Selective Attention During Category Learning

There is strong evidence that learning new categorizations entails altering selective attention to the features or dimensions that comprise stimuli (Kruschke, 1996; Nosofsky, 1991; Sutherland & Mackintosh, 1971). The current investigation extends this research in three ways. First, we have shown that the pattern of attention learned during one categorization is transferred to a subsequent categorization, giving rise to positive transfer. Positive transfer is found when the attention weights required for an earlier categorization are consistent with those required for the transfer categorization. We did not find strong evidence for negative transfer when the two attention weight patterns were inconsistent. Thus, one way of viewing Experiment 1 is as an extension of selective attention experiments to arbitrary dimensions.

The second extension beyond the existing literature's attentional effects in category learning has been to distinguish the importance of learning to attend to relevant dimensions and learning to ignore irrelevant dimensions. Although both effects have been observed (see the Introduction, or the extended review by Hall, 1991), their relative influences with controlled stimuli has not been studied. The current results indicated equally strong positive transfer effects when initial and transfer categorizations shared either relevant or irrelevant dimensions. This result is somewhat counterintuitive. When asked to describe how they learned to perform well on the categorization tasks, our participants generally mentioned features that they attended (e.g. "Faces in Club A were happier") rather than features that they ignored (e.g. "I learned that the apparent age of a face was irrelevant"). However, based on the equivalent amounts of transfer in the "acquired equivalence" and

"acquired distinctiveness" conditions, participants learn to both attend and ignore dimensions. Future research may well indicate that there is a dissociation between the importance of relevant and irrelevant dimensions as measured by their role in transfer and verbalizations about them. If so, then borrowing from Schooler's verbal overshadowing procedure (Brandimonte, Schooler, & Gabbino, 1997; Schooler & Engsler-Schooler, 1990), we would predict that asking participants to verbally describe the dimensions that they use to categorize faces should particularly interfere with positive transfer based on shared irrelevant dimensions. Irrelevant dimensions apparently exert more influence on categorization than is suggested by participants' own accounts of their performance.

A third extension to research on attention in category learning has been to explore learned selective attention to hard-to-describe, arbitrary dimensions. Most of the work on selective attention in human category learning has focused on separable, easily delineated dimensions (Nosofsky, 1986). The reason for this is clear; it is for such dimensions that selective attention is most efficient (Garner, 1976, 1978). However, the current work finds that selective attention is applied even when two dimensions are integral in Garner's sense that categorizing on the basis one dimension is hindered by the presence of irrelevant variation on the second dimension. We are not claiming that the arbitrary dimensions have no correlates to dimensions possessed by participants prior to the experiment. When creating dimensions by blending two faces in different proportions, it is certainly possible that participants can identify dimension values by detecting a local cue. For example, if one face seems to be slightly happier than the other, then the value along the morphed dimension may be identified by apparent happiness. Still, each face is composed out of two dimensions, and the same stimulus regions that specify the value along one dimension are also specifying the value along the other

dimension. That is, the information that is relevant for one dimension is distributed over the entire stimulus, and overlaps completely with the information relevant for the other dimension. This is true even more obviously with the saturation and brightness dimensions of Experiment 4. This is a different situation than occurs when the two dimensions can be identified by examining different parts of an object, as with color and shape for example. The type of selective attention that appears to be operating in our stimuli is like that required when two events are completely superimposed on one another (DeSchepper & Treisman, 1996; Neisser & Becklen, 1975). Furthermore, the observed selective attention was not limited to the trained stimuli and dimensions, but transferred to new stimuli when the transferred dimension was paired with new dimensions. The negative transfer conditions from Experiment 1 provide evidence that transfer was not based on similarity between stimuli used in the initial and transfer categorizations, but was instead due to the compatibility of learned selective attention to dimensions in the two categorizations.

Dimension Differentiation

Although our results strongly indicate a role for selective attention to arbitrary dimensions, selective attention alone is not sufficient to account for the pattern of transfer between categorizations. We propose that our participants also learned to organize faces into dimensions that combine to create the faces. Organizing faces into dimensions involves isolating independent sources of information in the faces, and is apparently informed by the category assignments given to the faces. When Dimension A was relevant for a categorization and Dimension B was irrelevant, participants apparently learned to isolate these two dimensions, so as to more effectively attend to Dimension A. Once Dimensions A and B are isolated, subsequent categorizations are facilitated if they make use of the same dimensional organization, and are impeded if

they make use of an incompatible organization. The pieces of evidence for this dimension differentiation process are: (1) surprisingly good performance when the irrelevant dimension becomes relevant <u>and</u> the relevant dimension becomes irrelevant, compared with situations in which only one of these changes is made (Experiment 1); (2) better transfer for categorization rules related by 90 than 45 degrees (Experiments 2, 3, and 4); (3) the advantage of categorization rules related by 90 rather than 45 degrees is only found when dimensions are overlapping and highly integral (Experiment 2); and 4) the 90 degree advantage is replicated when the stimuli offer no privileged dimensional axes (Experiment 3) and when classic examples of integral dimensions are used (Experiment 4). Categorization rules related by 45 degrees overlap partially in their selective attention requirements, but are completely incompatible in the dimensions that they encourage to be extracted. Conversely, categorization rules related by 90 degrees require the same dimensions to be isolated from each other, but are completely incompatible in their selective attention requirements. Thus, the relative superiority of transfer between rules related by 90 degrees indicates that the advantage due to consistent dimensional organization is sometimes stronger than the advantage due to consistent selective attention. This superiority is found only when highly overlapping dimensions are to be extracted, as is expected if the superiority is caused by dimension differentiation.

The results were generally consistent in suggesting that the transfer advantage for consistent dimensional organizations over inconsistent ones is largely due to difficulties caused by incompatible organizations. Particularly poor transfer performance was observed when the dimensions that needed to be isolated for a categorization were partially correlated with previously extracted dimensions. With these partially correlated dimensions, participants continue to try to interpret objects in terms of previously helpful dimensions despite their inadequacy. Although our dimensions are hard to describe, an intuition for this negative transfer can be generated with the rectangle stimuli in Figure 5. It is easy to see how an understanding of the rectangles in terms of height or width might interfere with a task that later required coding them in terms of area or shape. The former dimensions are incompatible with the latter dimensions in that they form cross-cutting organization of the stimulus space. An impressive outcome of this negative transfer is that category learning on the second categorization is often slower than initial category learning.

A critic might argue that dimension differentiation is a logical impossibility. By this argument, if two dimensions are fused together at some point in perceptual processing, then they can never be later split apart. By analogy, once red ink has been blended with blue ink, there is no simple procedure for later isolating only the blue ink. Fortunately, several computational models have recently been proposed that explain how a dimension differentiation mechanism might operate. Competitive learning networks differentiate inputs into categories by specializing detectors to respond to classes of inputs. Random detectors that are slightly more similar to an input than other detectors will learn to adapt themselves toward the input and will inhibit other detectors from doing so (Rumelhart & Zipser, 1985). The end result is that originally similar detectors that respond almost equally to all inputs become increasingly specialized and differentiated over training. Detectors develop that respond selectively to particular classes of input patterns or dimensions within the input. Smith, Gasser, and Sandhofer (1997) present a neural network simulation of the development of separated dimensions in children. In the network, dimensions become separated by detectors developing strong connections to specific dimensions while weakening their connections to all other dimensions. The model

captures the empirical phenomenon that dimension differentiation is greatly facilitated by providing comparisons of the sort "this red square and this red triangle have the same color."

The neural network approach for developing diagnostic dimensions which is perhaps most relevant to our experiments is the Expectation Maximization algorithm for factorial learning (Dempster, Laird, & Rubin, 1977; Ghahramani, 1995; Hinton, Dayan, Frey, & Neal, 1995; Tenenbaum, 1996). When presented with a set of inputs, this approach finds an underlying set of independent components that, when combined in different arrangements, reproduce the set of inputs. For example, imagine a set of 64 patterns that are generated by combining a horizontal line in any one of 8 positions with a vertical line in any one of 8 positions. From these patterns, an EM algorithm could generate the set of 16 horizontal and vertical lines that suffice for generating the 64 patterns (Ghahramani, 1995). It does so by finding the weightings of different hidden dimensions that would be most likely to have produced each of the 64 patterns. Impressively, the algorithm is able to discover both the hidden dimensions and their weightings by iterating between two steps: 1) computing the expected hidden dimensions given the current weights, and 2) maximizing the likelihood of the weights given these expected dimensions. Such an algorithm could uncover dimensions underlying our set of bald faces. However, as the EM algorithm is unsupervised, it would have to be extended so that the dimensions that it extracted would be influenced by the category feedback supplied with the faces. That is, our results indicate that people have a mechanisms that allows them to create part descriptions that are guided by the needed categorizations. In short, recent advances in neural networks can potentially supply us with an answer to our imaginary critic. The separate contributions of red and blue ink can be isolated if one has not only a single sample of mixed ink, but several samples with different proportions of red and blue ink.

Although we have argued that dimension differentiation is a different phenomenon than selective attention, similar mechanisms may underlie them. In particular, associative learning accounts of selective attention (e.g. Kruschke, 1992, 1996) are candidate accounts to explain dimension differentiation, as long as they operate at different levels. That is, the way a system learns to selectively attend to a detector for a categorization may be similar to the way that the detector learns to select particular stimulus elements. Imagining a system with connections from the external world to hypothetical detectors and connections from the detectors to categories is a helpful way of seeing the similarities and differences between selective attention and dimension differentiation. Learning the former connections involves dimension differentiation whereas learning the latter connections involves selective attention. Selectively attending to a particular property for categorization is only possible if a detector has already isolated that property. Having a detector develop a specialized response to a single property may be a rather slow process, but once it has become specialized, selective attention to that property may be rather fast. Goldstone et al. (2000) present more details on how a single network can develop specialized detectors at the same time that associations between the detectors and categories are acquired.

Mechanisms of Flexibility in Category Learning

We have hypothesized two distinctive mechanisms that flexibly adapt object descriptions to the requirements imposed by category learning. The difference between selective attention and dimension differentiation may be difficult to understand because both involve flexibly focusing on a specific source of information. If one observes that a person's categorization based on one dimension becomes less sensitive to variation along another dimension, it may be due to either changes in

selective attention or dimension differentiation, or both. The difference, in essence, is between learning appropriate weights for dimensions, and learning how to learn appropriate weights for dimensions. People have difficulty in learning to appropriately attend to some dimensions, such as brightness and saturation, because they cannot even isolate these dimensions from each other. However, integral dimensions are not necessarily doomed to remain inextricably fused. Training can help people isolate the two dimensions (Burns & Shepp, 1988; Goldstone, 1994-a), and once isolated, selective attention can operate with greater efficiency. The current experiments have described a new source of evidence for dimension differentiation that does not suffer from some of the problems of previous methods. Perceptual dimensions that are not originally privileged for interpreting objects can become

privileged when they reliably predict important categories.

The current work fits with recent efforts to describe how learned categories affect subsequent cognitive processes (Ross, 1999). In particular, perceptual and attentional processes are modified by category learning. Arbitrary dimensions can be selectively attended, and can be isolated from other dimensions. Attentional and perceptual processes provide object descriptions that serve as the foundational basis for our visual concepts. However, these are foundational processes that adapt to the concepts that they support (Goldstone & Barsalou, 1998; Goldstone et al., 2000; Schyns et al., 1998). To be foundational does not mean to be static and rigid. Rather, like a good pair of shoes that provides support by flexibly conforming to the foot, the processes that produce object descriptions support our concepts by conforming to these concepts.

References

Brandimonte, M. A., Schooler, J. W, & Gabbino, P. (1997). Attenuating verbal overshadowing through color retrieval cues. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23, 915-931.

Blough, D. S., & Blough, P. M. (1997). Form perception and attention in pigeons. Animal Learning and Behavior, 25, 1-20.

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.

Cheng, P. W., & Pachella, R. G. (1984). A psychophysical approach to dimensional separability. Cognitive Psychology, 16, 279-304.

Clapper, J. P., & Bower, G. H. (1994). Category invention in unsupervised learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 20, 443-460.

Cook, G., & Stephens, J. T. (1995). The priority of separable perception in stimulus classifications of children with mental retardation. Child Development, 66, 1057-1071.

Dempster, A., Laird, N. & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society Series B, 39, 1-38.

DeSchepper, B., & Treisman, A. (1996). Visual memory for novel shapes: Implicit coding without attention. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 27-47.

Edelman, S. (1999). Representation and recognition in vision. Cambridge, MA: Bradford Books, MIT Press.

Ennis, D. M. (1988). Confusable and discriminable stimuli: Comment on Nosofsky (1986) and Shepard (1986). Journal of Experimental Psychology: General, 117, 408-411.

Estes, W. K. (1994). Classification and cognition. New York: Oxford University Press.

Farah, M. J. (1992). Is an object an object an object? Cognitive and neuropsychological investigations of domain-specificity in visual object recognition. Current Directions in Psychological Science, 1, 164-169.

Feldman, J., & Richards, W. (1998). Mapping the mental space of rectangles. Perception, 27, 1191-1202. Perception, 27, 1191-1202.

Foard, C. F., & Kemler, D. G. (1984). Holistic and analytic modes of processing: The multiple determinants of perceptual analysis. Journal of Experimental Psychology: General, 113, 94-111.

Fox, E. (1995). Negative priming from ignored distracters in visual selection: A review. Psychonomic Bulletin and Review, 2, 145-173.

Garner, W. R. (1976). Interaction of stimulus dimensions in concept and choice processes. Cognitive Psychology, 8, 98-123.

Garner, W. R. (1978). Selective attention to attributes and to stimuli. Journal of Experimental Psychology: General, 107, 287-308.

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

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. American Psychologist, 52, 45-56.

Gentner, D. & Namy, L.L. (in press). Comparison in the development of categories, Cognitive Development.

Gershkoff-Stowe, L., Thal, D. J., Smith, L. B., & Namy, L. L. (1997). Categorization and its developmental relation to early language, Child Development, 68, 843-859.

Ghahramani, Z. (1995). Factorial learning and the EM algorithm. In G. Tesauro, D. S. Touretzky, & T. K. Leen (Eds.), Advances in Neural Information Processing Systems 7. Cambridge, MA: MIT Press, 617-624.

Goldstone, R. L. (2000). Unitization during category learning. Journal of Experimental Psychology: Human Perception and Performance, 26, 86-112.

Goldstone, R. L. (1994-a). influences of categorization on perceptual discrimination. Journal of Experimental Psychology: General, 123, 178-200.

Goldstone, R. L. (1994-b). An efficient method for obtaining similarity data. Behavior Research Methods, Instruments, & Computers, 26, 381-386.

Goldstone, R. L., & Barsalou, L. (1998). Reuniting perception and conception. Cognition, 65, 231-262.

Goldstone, R. L., Steyvers, M., Spencer-Smith, J., & Kersten, A. (2000). Interactions between perceptual and conceptual learning. in E. Diettrich & A. B. Markman (Eds.) Cognitive Dynamics: Conceptual Change in Humans and Machines. (pp. 191-228). Lawrence Erlbaum and Associates.

Grau, J. W., & Nelson, D. K. (1988). The distinction between integral and separable dimensions: Evidence for the integrality of pitch and loudness. Journal of Experimental Psychology: General, 117, 347-370.

Haider, H., & Frensch, P. A. (1996). The role of information reduction in skill acquisition. Cognitive Psychology, 30, 304-337.Hall, G. (1991). Perceptual and Associative Learning. Oxford: Clarendon Press.

Hall, G. (1991). Perceptual and Associative Learning. Oxford: Clarendon Press.

Handel, S., & Imai, S. (1972). The free classification of analyzable and unanalyzable stimuli. Perception & Psychophysics, 12, 108-116.

Herrnstein, R. J. (1990).   Levels of stimulus control: A functional approach. Special Issue: Animal cognition. Cognition, 37, 133-166.

Hinton, G. E., Dayan, P., Frey, B. J., & Neal, R. M. (1995). The "wake-sleep" algorithm for unsupervised neural networks. Science, 268, 1158-1161.

Hock, H. S., Tromley, C., & Polmann, L. (1988). Perceptual units in the acquisition of visual categories. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 75-84.

Honey, R. C., & Hall, G. (1989). The acquired equivalence and distinctiveness of cues. Journal of Experimental Psychology: Animal Behavior Processes, 15, 338-346.

Kayser, A. (1997). Heads. New York: Abbeville Press.

Kemler-Nelson, D. G., (1993). Processing integral dimensions: The whole view. Journal of Experimental Psychology: Human Perception and Performance, 19, 1105-1113.

Kemler, D. G., & Smith, L. B. (1978). Is there a developmental trend from integrality to separability in perception? Journal of Experimental Child Psychology, 26, 498-507.

Kersten, A. W., Goldstone, R. L., & Schaffert, A. (1998). Two competing attentional mechanisms in category learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24, 1437-1458.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. Child Development, 67, 2797-2822.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. Psychological Review, 99, 22-44.

Kruschke, J. K. (1993). Human category learning: Implications for backpropagation models. Connection Science, 5, 3-36

Kruschke, J. K. (1996). Dimensional relevance shifts in category learning. Connection Science, 8, 225-247.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 6, 293-323.

Lawrence, D.H. (1949). Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. Journal of Experimental Psychology, 39, 770-784.

Lubow, R. E., & Kaplan, O. (1997). Visual search as a function of type of prior experience with target and distracter, Journal of Experimental Psychology: Human Perception and Performance, 23, 14-24.

Maddox, W. T. (1992). Perceptual and decisional separability. In F. G. Ashby (Ed.), Multidimensional models of perception and cognition (pp. 147-180). Hillsdale, NJ: Erlbaum.

Markman, A. B., & Gentner, D. (1996). Commonalities and differences in similarity comparisons. Cognitive Psychology, 25, 431-467.

Markman, A. B., & Makin, V. S. (1998). Referential communication and category acquisition. Journal of Experimental Psychology: General, 127, 331-354.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. Psychological Review, 100, 254-278.

Medin, D. L., & Shaeffer, M. M. (1978). A context theory of classification learning. Psychological Review, 85, 207-238.

Melara, R. D. (1989). Similarity relations among synesthetic stimuli and their attributes. Journal of Experimental Psychology: Human Perception and Performance, 115, 212-231.

Melara, R. D., Marks, L. E., & Potts, B. C. (1993). Primacy of dimensions in color perception. Journal of Experimental Psychology: Human Perception and Performance, 19, 1082-1104.

Melcher, J. M., & Schooler, J. W. (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. Journal of Memory and Language, 35, 231-245.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. Cognitive Psychology, 7, 480-494.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. Journal of Experimental Psychology: General, 115, 39-57.

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 87-108.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. Journal of Experimental Psychology: Human Perception and Performance, 17, 3-27.

Nosofsky, R. M., & Palmeri, T. J. (1996). Learning to classify integral-dimension stimuli. Psychonomic Bulletin & Review, 3, 222-226.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. Psychological Review, 101, 53-79.

Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. Psychological Review, 94, 61-73.

Pevtzow, R., & Goldstone, R. L. (1994). Categorization and the parsing of objects.   Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society. (pp. 717-722). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. Journal of Experimental Psychology, 77, 353-363.

Regehr, G., & Brooks, L. R. (1995). Category organization in free classification: The organizing effect of an array of stimuli. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 347-363.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black and W. F. Prokasy (Eds.) Classical conditioning II: Current research and theory (pp. 64-99). Appleton-Century-Crofts: New York.

Ross, B. H. (1999). Postclassification category use: The effects of learning to use categories after learning to classify. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25, 743-757.

Rumelhart, D. E., & Zipser, D. (1985). Feature discovery by competitive learning. Cognitive Science, 9, 75-112.

Schoenemann, P. H. (1977). Similarity of rectangles. Journal of Mathematical Psychology, 16, 161-165.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. Cognitive Psychology, 22, 36-71.

Schyns, P. G., Goldstone, R. L, & Thibaut, J. (1998). Development of features in object concepts. Behavioral and Brain Sciences, 21, 1-54.

Schyns, P. G., & Murphy, G. L. (1994). The ontogeny of part representation in object concepts. In D. L. Medin (Ed.). The Psychology of Learning and Motivation, 31, 305-354. Academic Press: San Diego, CA.

Schyns, P. G., & Rodet (1997). Categorization creates functional features. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23, 681-696.

Shepard, R. N., Hovland, C. L., & Jenkins, H. M. (1961). Learning and memorization of classification. Psychological Monographs, 75 (13), Whole No. 517.

Shepp, B. E., Burns, B., & McDonough, D. (1980). The relation of stimulus structure to perceptual and cognitive development: Further tests of a separability hypothesis. In F. Wilkening, J. Becker, & T. Trabasso (Eds.), Information integration by children. (pp. 113-146). Hillsdale, NJ: Erlbaum.

Shiffrin, R. M., & Lightfoot, N. (1997). Perceptual learning of alphanumeric-like characters. In R. L. Goldstone, P. G. Schyns, & D. L. Medin (Eds.) The Psychology of Learning and Motivation, Volume 36. San Diego: Academic Press. (pp. 45-82).

Shiffrin, R. M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual Learning, automatic attending and a general theory. Psychological Review, 84, 127-190.

Smith, L. B. (1989a). From global similarity to kinds of similarity: The construction of dimensions in development. In S. Vosniadou and A. Ortony (Eds.), <u>Similarity and analogical reasoning</u> (pp. 146 -178). Cambridge: Cambridge University Press.

Smith, L. B. (1989b). A model of perceptual classification in children and adults. <u>Psychological Review</u>, <u>96</u>, 125-144.

Smith, L. B., & Evans. P. (1989). Similarity, identity, and dimensions: Perceptual classification in children and adults. In B. E. Shepp & S. Ballesteros (Eds.), <u>Objects perception: Structure and process</u>. Hillsdale, NJ: Erlbaum

Smith, L. B., Gasser, M., & Sandhofer, C. (1997). Learning to talk about the properties of objects: A network model of the development of dimensions. (pp. 220-256). In R. L. Goldstone, P. G. Schyns, & D. L. Medin (Eds.) <u>Psychology of Learning and Motivation</u>, Vol. 36. San Diego, CA: Academic Press.

Smith, L. B., & Kemler, D. G. (1978). Levels of experienced dimensionality in children and adults. <u>Cognitive Psychology</u>, <u>10</u>, 502-532.

Sutherland, N. S., & Mackintosh, N. J. (1971). <u>Mechanisms of animal discrimination learning.</u> New York: Academic Press.

Steyvers, M. (1999). Morphing techniques for generating and manipulating face images. Behavior Research Methods, Instruments, & Computers, 31, 359-369.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. <u>Quarterly Journal of Experimental Psychology</u>, <u>46A</u>, 225-245.

Tenenbaum, J. B. (1996). Learning the structure of similarity. In G. Tesauro, D. S. Touretzky, & T. K. Leen (Eds.), <u>Advances in Neural Information Processing Systems 8</u>. Cambridge, MA: MIT Press, 4-9.

Thorndike, E. L. (1903). <u>Education psychology</u>. New York: Lemke & Buechner.

Tighe, T. J., & Tighe, L. S. (1969). Facilitation of transposition and reversal learning in children by prior perceptual training. <u>Journal of Experimental Child Psychology</u>, <u>8</u>, 366-374.

Tipper, S. P. (1992). Selection for action: The role of inhibitory mechanisms. <u>Current Directions in Psychological Science</u>, <u>1</u>, 105-109.

Ward, T. B., & Vela, E. (1986). Classifying color materials: Children are less holistic than adults. <u>Journal of Experimental Child Psychology</u>, <u>42</u>, 273-302.

Author Notes

Table 1. Initial and Transfer Categorizations of Experiment 1

Transfer Condition	Initial Phase		Transfer Phase	
	Relevant	Irrelevant	Relevant	Irrelevant
Identity (A|B)	A	B	A	B
Acquired Distinctiveness (A|C)	A	C	A	B
Acquired Equivalence (C|B)	C	B	A	B
Negative Priming (C|A)	C	A	A	B
Attentional Capture (B|C)	B	C	A	B
90 Degree Rotation (B|A)	B	A	A	B
Neutral Control (C|D)	C	D	A	B

Table 2. Summary of Results from Experiments 2-4.

Experiment	Condition	1st	Phase 2nd Perserved Assignments 25%	50%	75%	Item Assignment Different	Same
2a	45-degrees	65.4	59.7		63.2	62.1	61.2
	90-degrees	64.7		65.7		60.0	70.7
2b	45-degrees	67.9	65.6		64.7	67.5	63.1
	90 -degrees	64.3		65.0		65.9	64.1
3	45-degreees	59.4	54.1		53.2	53.6	53.7
	90-degrees	58.6		58.9		58.2	60.4
4	45-degrees	79.1	74.1		74.8	75.7	73.0
	90-degreees	80.2		82.1		81.3	82.9

Note: Preserved assignments refers to the percentage of category assignments that were preserved between the first and second phases of the experiment across the 8 (Experiments 2A and 2B) or 16 (Experiments 3 and 4) stimuli. Item assignment refers to whether a particular stimulus received the same or different assignment between the two phases.

Table 3. Commission Internationale de L'Eclairage (CIE) color coordinates for the 16 stimuli used in Experiment 4.

Stimulus	Luminance	Purity
1	39	52
2	38	62
3	34	71
4	31	77
5	28	80
6	22	77
7	17	71
8	15	62
9	14	52
10	15	40
11	17	31
12	22	27
13	28	24
14	31	27
15	34	31
16	38	40

Note: Luminance values are expressed in cd/m^2; purity values are percentages
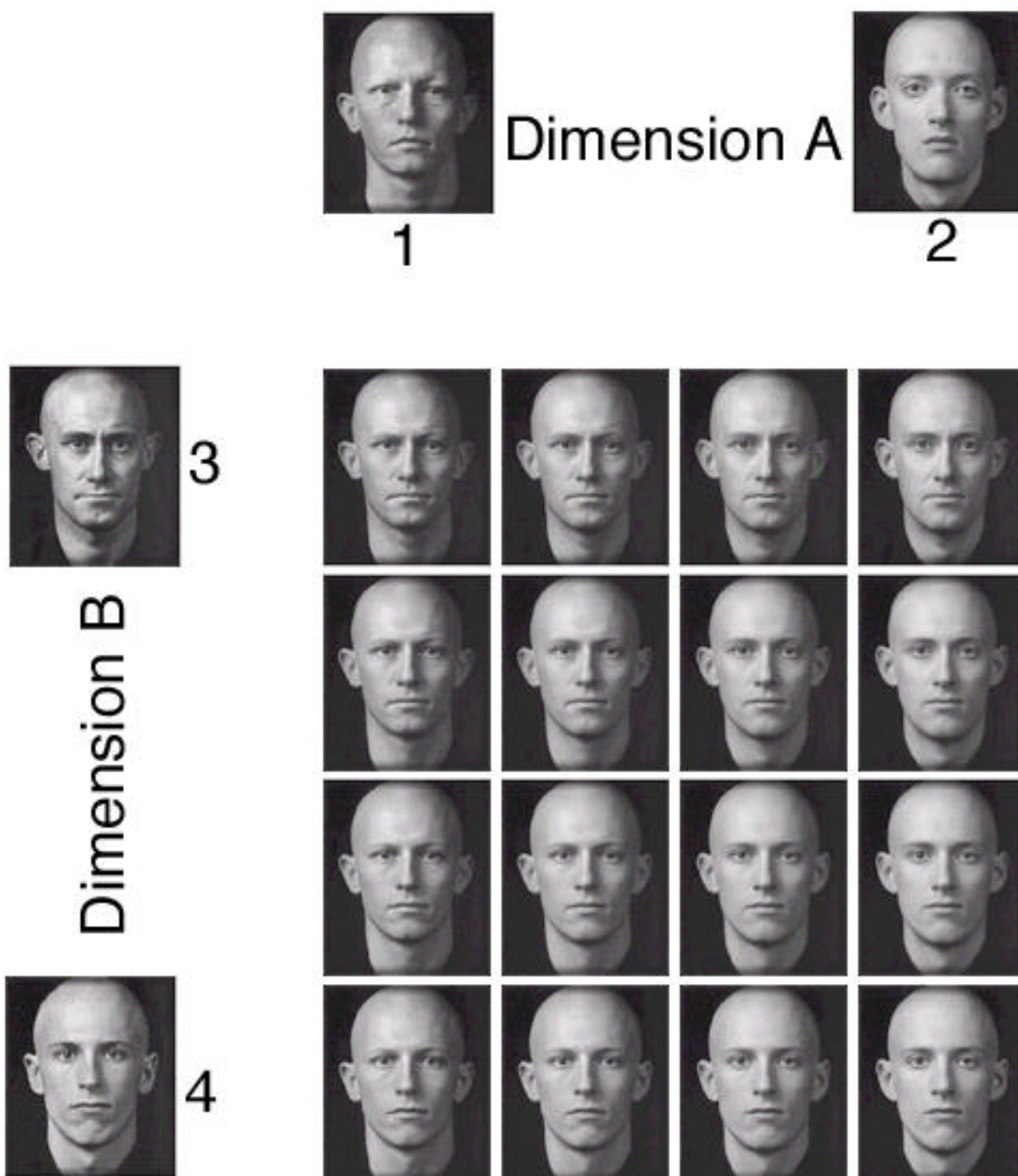
Figure 1. The four Faces 1, 2, 3 and 4 are blended in different proportions to create a 4 X 4 matrix of faces. The proportions of Faces 1 and 2 are negatively correlated such that the more of Face 1 present in one of the 16 center faces, the less of Face 2 there will be. This negative correlation establishes Dimension A, and a similar negative correlation between Faces 3 and 4 establishes Dimension B. Each of the 16 center faces is defined half by its value on Dimension A and half by its value on Dimension B.
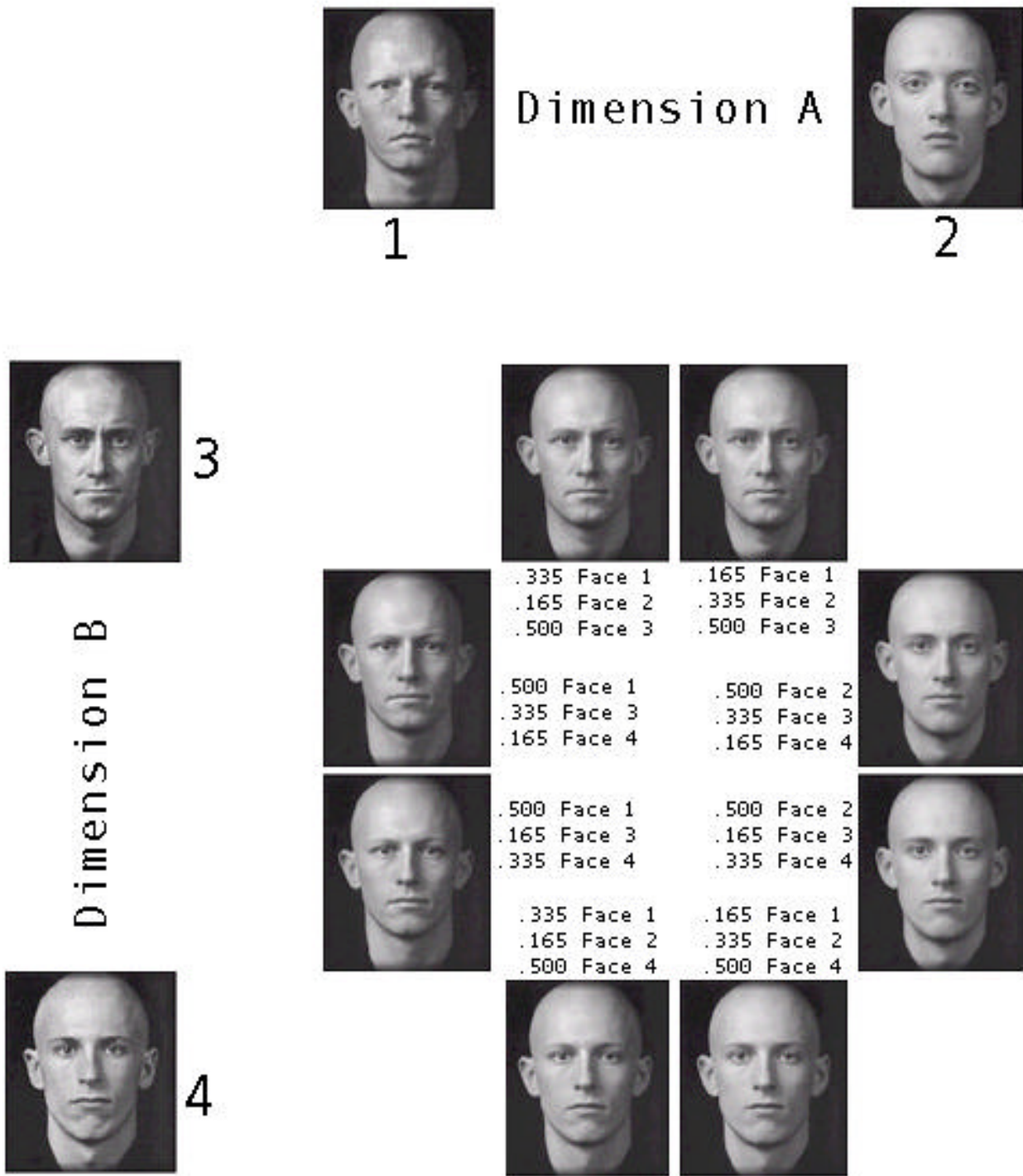
Figure 2. A subset of 8 of the faces from the 4 X 4 matrix of Figure 1 was chosen as the stimuli for Experiment 1. Beside each face are the proportions of Faces 1, 2, 3, and 4 that were used to generate it.
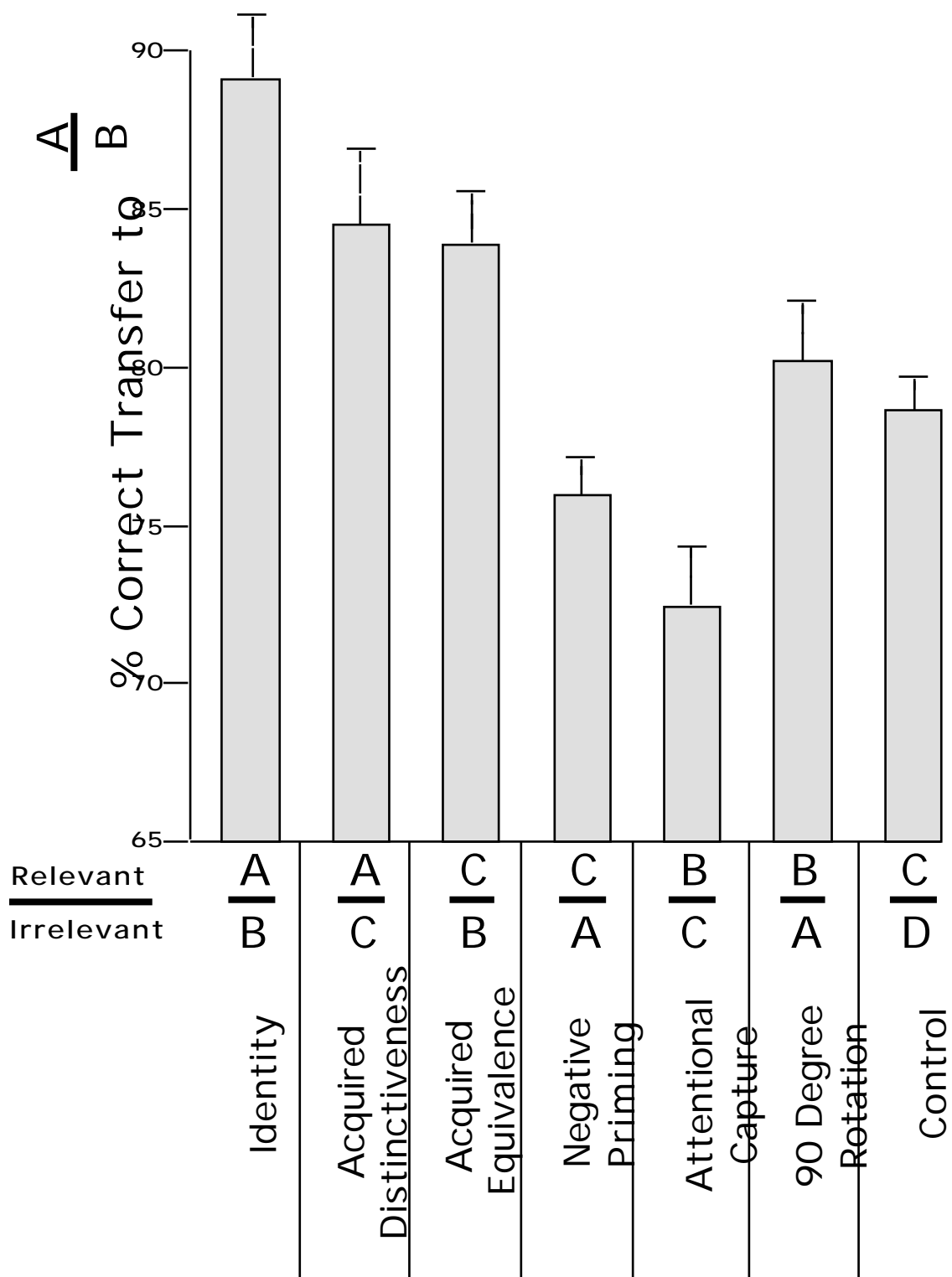
Figure 3. Results from Experiment 1, showing average categorization accuracy on the transfer categorization in which Dimension A was relevant and Dimension B was irrelevant. The conditions can be compared to the control condition on the far right.
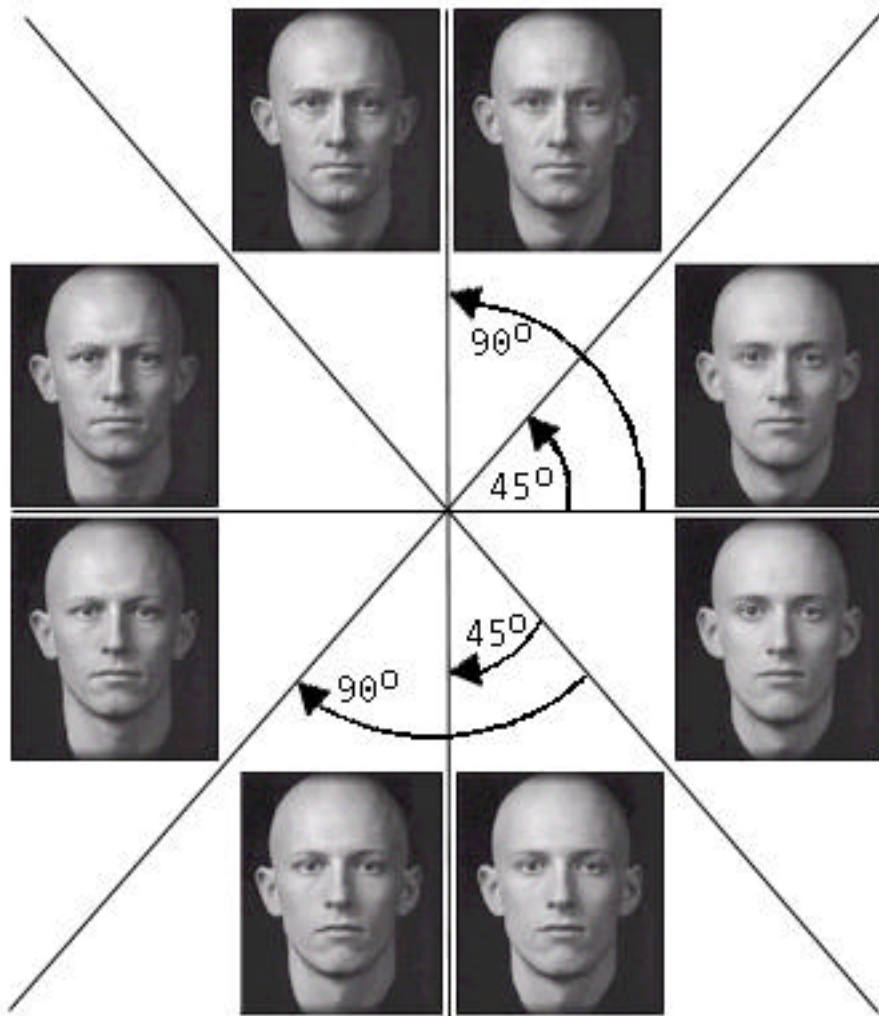
Figure 4. Sample stimuli from Experiment 2, showing two of the possible initial categorization boundaries. For each initial categorization boundary, sample transfer boundaries are shown for 45 and 90 degree rotations. Whether a transfer boundary is a 45 or 90 degree rotation is independent of its absolute orientation.

| | Compatible | | Incompatible | | Compatible |
	Height	Width	Shape	Area
☐	1	1	1:1	1
▭	1	2	2:1	2
▢	2	2	1:1	4
▯	2	1	1:2	2

Figure 5. An illustration of how stimuli can be understood in terms of competing dimensional organizations. Rectangles can be understood in terms of height and width, or in terms of their area and shape. An analysis in terms of width is compatible with an analysis into height, but is incompatible with an analysis in terms of shape because width and shape are not independent sources of variation.
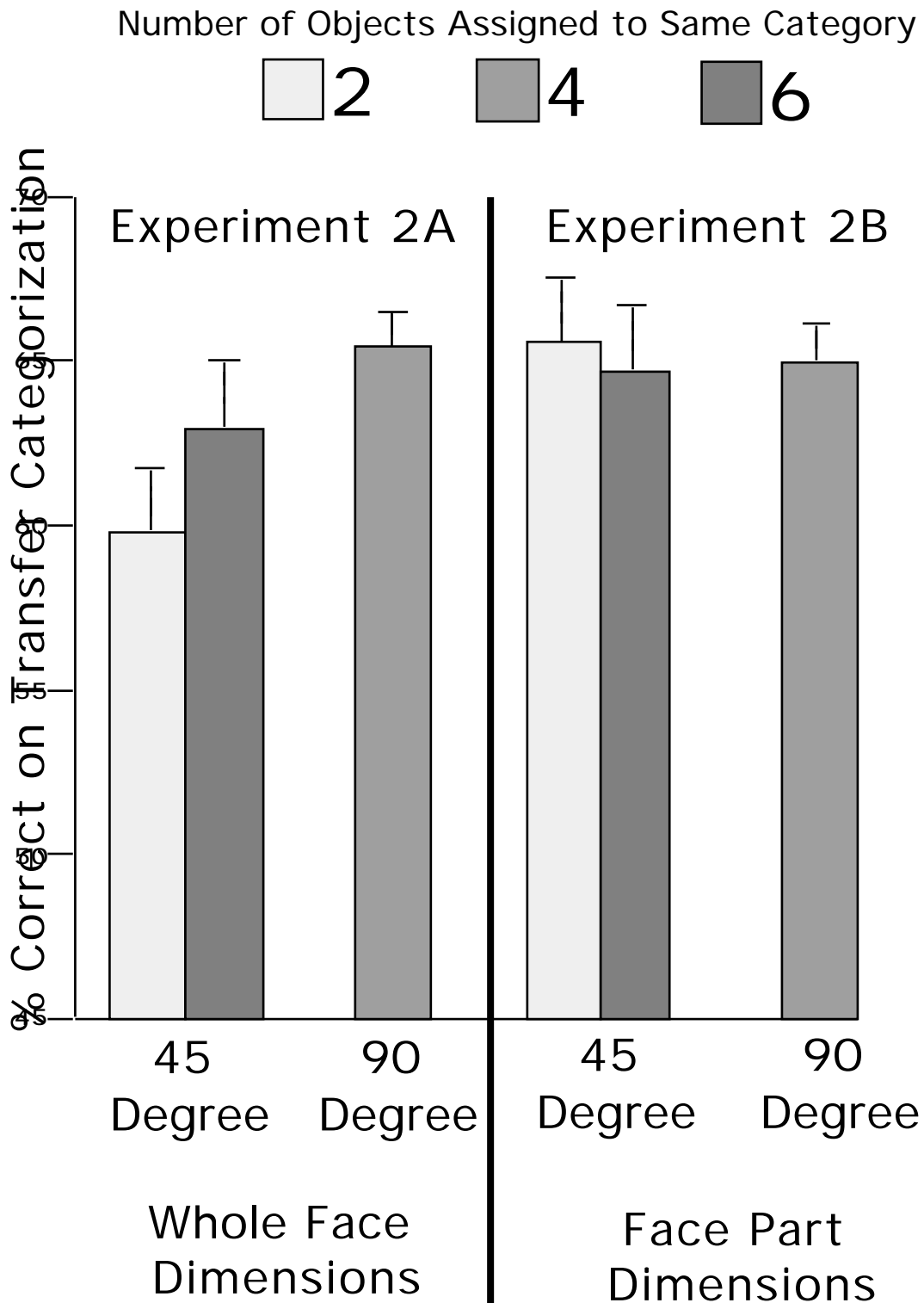
Figure 6. Results from Experiment 2A and 2B. The 90 degree rotation condition produced better transfer on the final categorization than the 45 degree condition, but only for overlapping dimensions (Experiment 2A) and not separated dimensions (Experiment 2B).
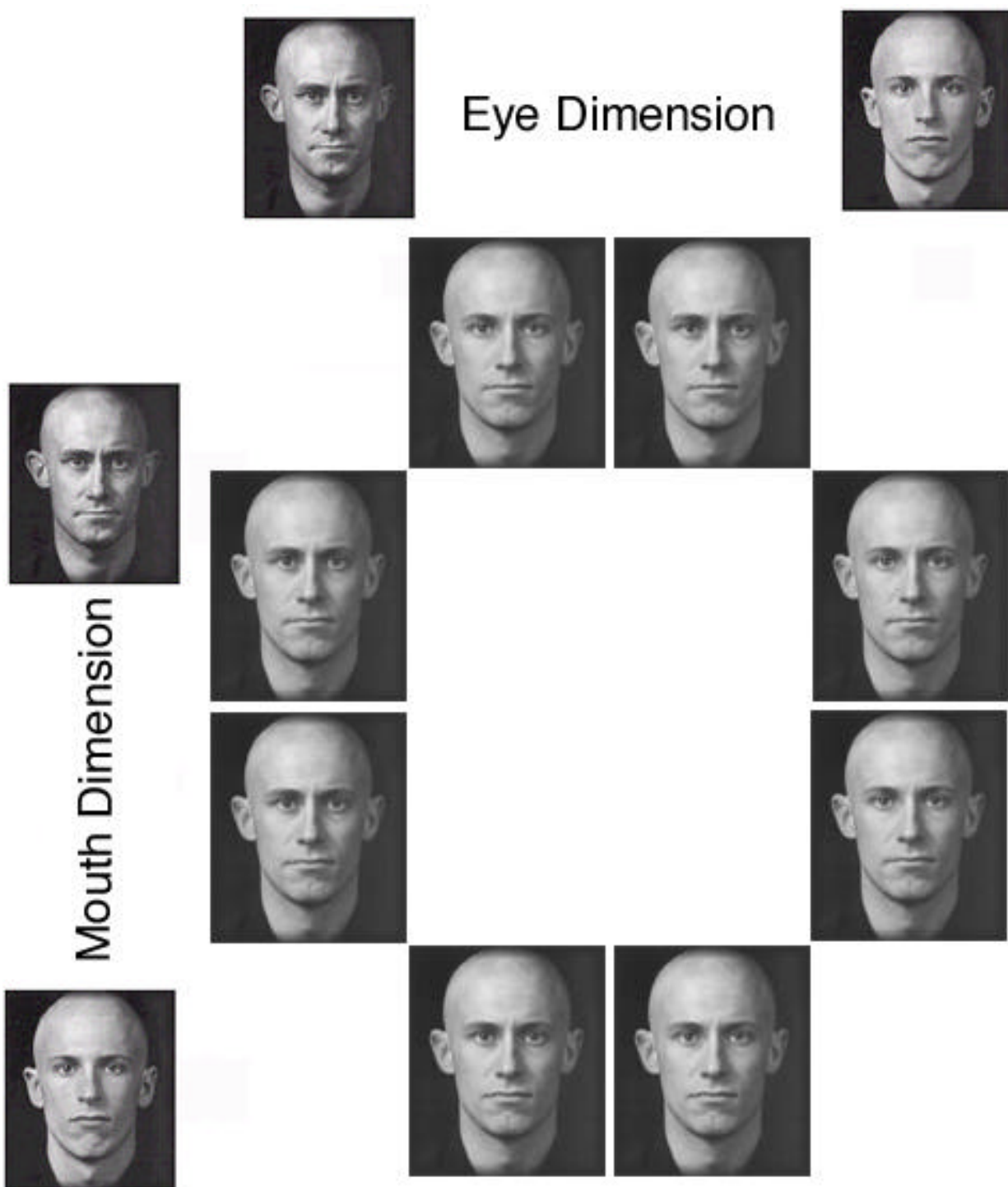
Figure 7. Stimuli from Experiment 2B. The dimensions were defined by the relative contributions from two faces to particular regions of a morphed face. The two dimensions were defined by the eyes and mouths of the faces. The contributing faces are shown on the ends of the dimension axes, and the presented stimuli are the eight faces in the middle.
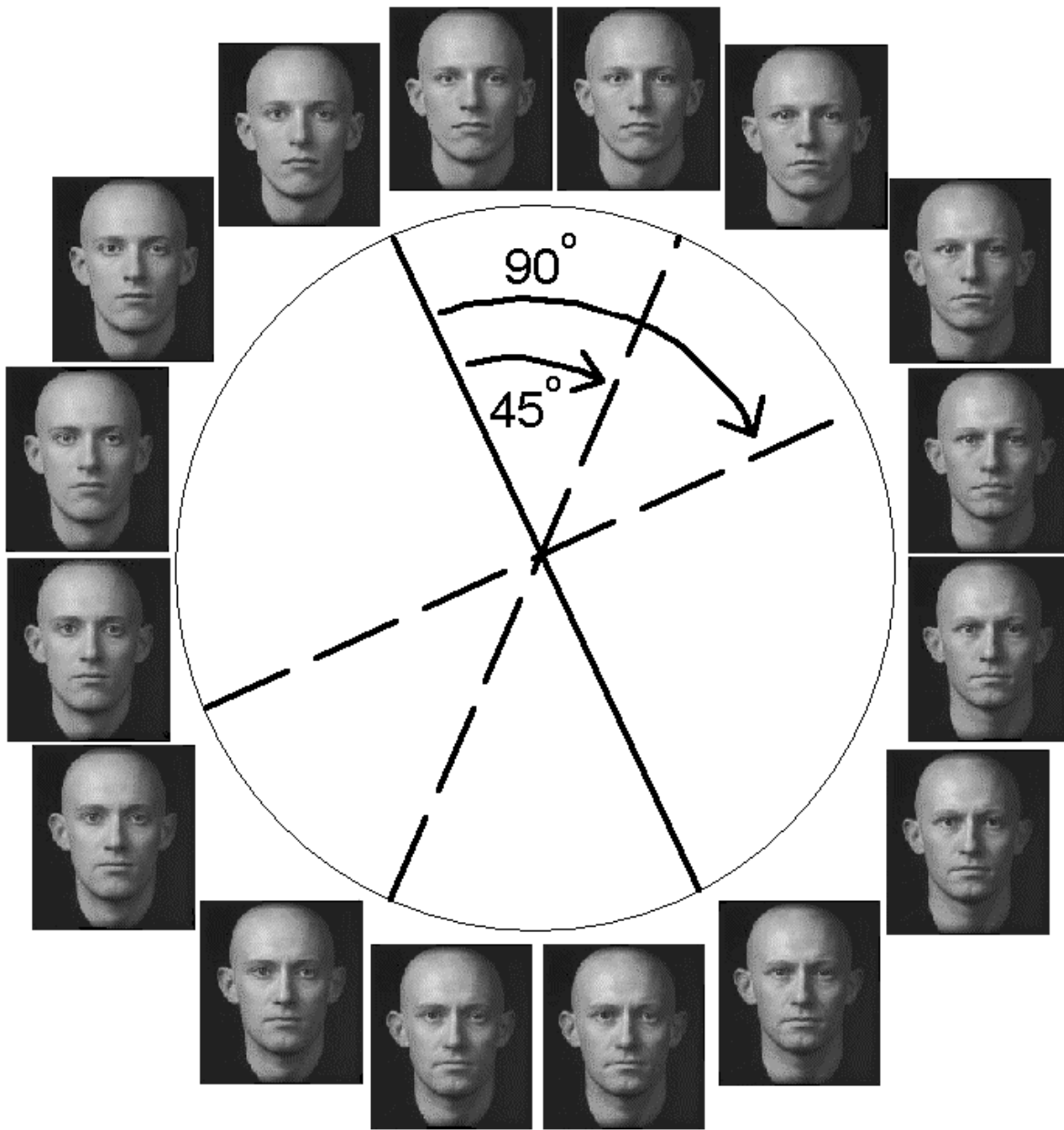
Figure 8. Sample stimuli from Experiment 3. The circularly arranged set of faces do not imply a preferred orientation for the dimensional axes.
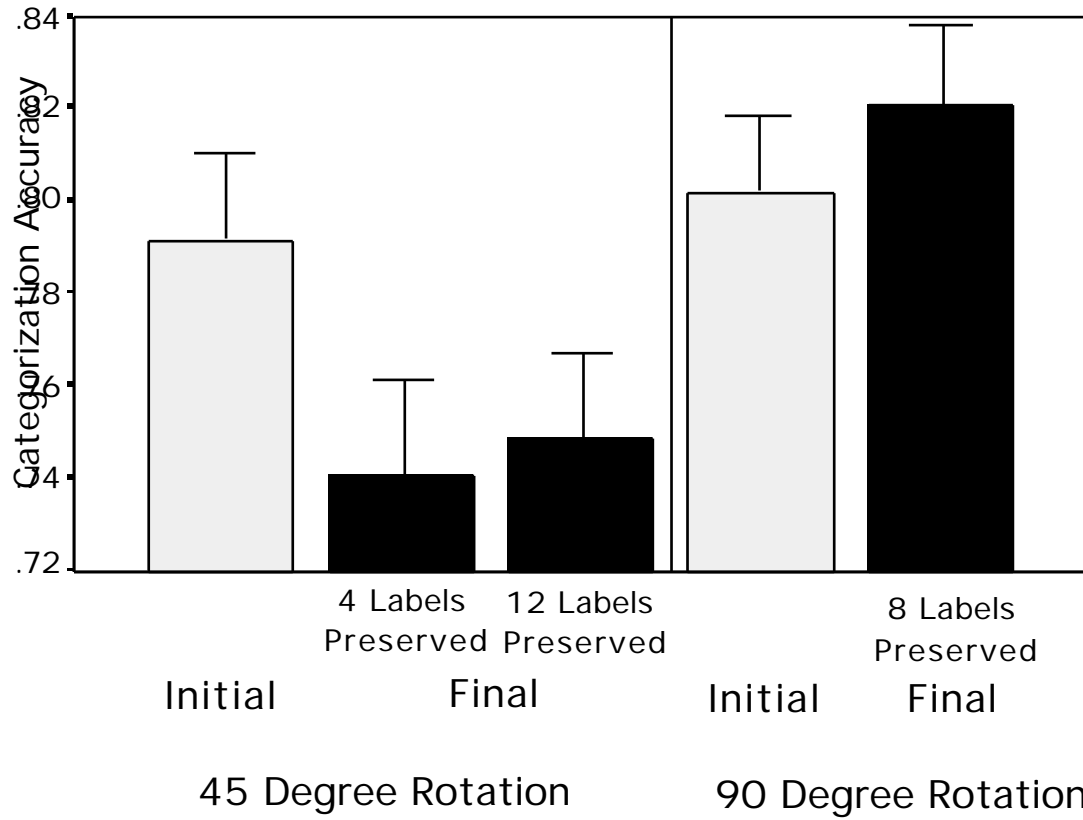
Figure 9. Results from Experiment 4, involving saturation and brightness as component dimensions of color.