COGNITION

Cognition 78 (2001) 27–43

www.elsevier.com/locate/cognit

Altering object representations through category learning

Robert L. Goldstone[a,*], Yvonne Lippa[a,b], Richard M. Shiffrin[a]

[a] *Indiana University, Bloomington, IN 47405, USA*
[b] *Max Planck Institute for Psychological Research, Amalienstraße 33, D-80799 Munich, Germany*

Received 1 June 1999; received in revised form 12 October 1999; accepted 16 June 2000

Abstract

Previous research has shown that objects that are grouped together in the same category become more similar to each other and that objects that are grouped in different categories become increasingly dissimilar, as measured by similarity ratings and psychophysical discriminations. These findings are consistent with two theories of the influence of concept learning on similarity. By a Strategic Judgment Bias account, the categories associated with objects are explicitly used as cues for determining similarity, and objects that are categorized together are judged to be more similar because similarity is not only a function of the objects themselves, but also the objects' category labels. By a Changed Object Description account, category learning alters the description of the objects themselves, emphasizing properties that are relevant for categorization. A new method for distinguishing between these accounts is introduced which measures the difference between the similarity ratings of categorized objects to a neutral object. The results indicate both strategic biases based on category labels and genuine representational change, with the strategic bias affecting mostly objects belonging to different categories and the representational change affecting mostly objects belonging to the same category. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Concept learning; Categorical perception; Perceptual learning; Top-down perception

1. Introduction

It is relatively uncontroversial that the concepts that we humans learn depend upon our perceptual representations. Our 'Cat' concept depends upon encoding

* Corresponding author. Psychology Department, Indiana University, Bloomington, IN 47405, USA. Fax: +1-812-855-4853.

E-mail address: rgoldsto@indiana.edu (R.L. Goldstone).

0010-0277/01/$ - see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S0010-0277(00)00099-8

features such as 'four legs', 'meows', and 'whiskers'. Whether the members of concepts are represented by rules (Bruner, Goodnow, & Austin, 1956), features (Rosch, 1978), or dimensional coordinates (Nosofsky, 1986), these representations are assumed to be based at least partially on perceptual descriptions of the objects. Theorists have also explored the possibility that the descriptions given to concept members do not only influence, but are also influenced *by*, the learned concepts. Several empirical results and theoretical treatments lead to the suggestion that the relation between perceptual descriptions and concept representations is not uni-directional, but rather is bi-directional and mutually supporting (Goldstone, 2000; Goldstone, Steyvers, Spencer-Smith, & Kersten, 2000; Lin & Murphy, 1997; Schyns, Goldstone, & Thibaut, 1998; Schyns & Murphy, 1994; Schyns & Rodet, 1997; Wisniewski & Medin, 1994).

One of the largest sources of evidence that concepts influence perceptual descriptions comes from the field of categorical perception (Harnad, 1987). By this phenomenon, people are better able to distinguish between physically different stimuli when the stimuli come from different categories than when they come from the same category. For example, Liberman, Harris, Hoffman, and Griffith (1957) generated a set of 14 vowel–consonant syllables going from /be/ to /de/ to /ge/ by varying the onset frequency of the second-formant transition of the initial consonant. The 14 speech sounds were created by making equal physical spacings between neighboring sounds. Observers listened to three sounds – A followed by B followed by X – and indicated whether X was identical to A or B. X was chosen to always be identical to either A or B. Participants performed this task more accurately when syllables A and B belonged to different phonetic categories than when they were physical variants of the same phoneme, even when the physical difference between A and B was equated. Liberman et al. (1957) concluded that the phonemic categories possessed by an adult speaker of English influence the perceptual discriminations that can be made.

Some researchers have argued that the phonemic categories that yield categorical perception effects are innate, or at least present in 4-month-old infants (Eimas, Siqueland, Jusczyk, & Vigorito, 1971). However, a number of studies have suggested that learned rather than innate categories can also produce categorical perception effects. A sound difference that crosses the boundary between phonemes in a language will be more discriminable to speakers of that language than to speakers of a language in which the sound difference does not cross a phonemic boundary (Repp & Liberman, 1987). Laboratory training on the sound categories of a language can produce categorical perception among speakers of a language that does not have these categories (Pisoni, Aslin, Perey, & Hennessy, 1982).

There have also been a number of studies showing categorical perception effects for learned visual categories involving arbitrary categories. For example, in a study by Katz (1963), children were required to associate four similar geometrical shapes with nonsense syllables. One group learned to assign two syllables to the four stimuli, such that two shapes had the same name. Another group assigned four different syllables to the four stimuli. The results of a subsequent same/different judgment task showed that children who learned only two syllables judged those shapes that were named identically as being the same more often than children who learned to label each shape with a different name. This result is consistent with others (for an overview see Arnoult, 1957; Cantor, 1965) in showing that objects that are associated with the same label or outcome subsequently are treated as being more similar to each other than they were prior to training. More generally, researchers have found that the particular category labels that are provided when a set of objects are learned influences how the categories are structured (Landau & Shipley, 1996; Pothos & Chater, 1997).

In a more recent exploration of this notion, Goldstone (1994a) first trained subjects on one of several categorization conditions in which one physical dimension was relevant and another was irrelevant. Subjects were then transferred to same/different judgments ('Are these two squares physically identical or not?'). Ability to discriminate between squares in the same/different judgment task, measured by Signal Detection Theory's d', was greater when the squares varied along dimensions that were relevant during categorization training. In one case, experience with categorizing objects actually decreased people's ability to spot subtle perceptual differences between the objects, if the objects belonged to the same category. These acquired categorical perception effects were observed for the easily separated dimensions of brightness and size as well as the psychologically fused dimensions of brightness and saturation. Beale and Keil (1995, 1996) found that participants show categorical perception for continua between familiar faces. Levin and Beale (2000) found heightened discriminability at category boundaries even for unfamiliar and inverted faces, and conclude that such effects are based on the rapid learning of perceptual categories. Particularly strong categorical perception effects have been found when researchers have used similarity ratings rather than psychophysical measures of discriminability. For example, Livingston, Andrews, and Harnad (1998) found that when participants learned to classify objects (animal body parts or artificial cells) into a single category, then these objects received far higher similarity ratings than objects that belonged to different categories, or than objects that were not categorized at all (see also Kurtz, 1996). Their finding of greater similarity associated with objects grouped together in a category is consistent with results that very young infants (2 months old) show sensitivity to differences between speech sounds that they lose by the age of 10 months (Werker & Tees, 1984). This desensitization only occurs if the different sounds come from the same phonetic category in the children's native language.

2. How does categorization influence judgments?

Despite the wealth of evidence that learned categories can produce categorical perception effects, there are a number of unanswered questions from this literature. One of the most troubling is: do learned categories merely influence strategic judgments (the Strategic Judgment Bias account) made about objects, or do they affect the psychologically encoded descriptions of the objects (the Altered Object Description account)? This issue is clearly apparent with the results of Livingston et al.

(1998). Imagine participating in their experiment. You have just learned a categorization in which objects A and B belong to the same category. You are now asked to rate their similarity. You may well give A and B a fairly high similarity rating, reasoning that 'The experimenter just told me that they are in the same category, so I suppose I should give them a high similarity rating.' This is a 'task demand' account in which participants may be responding to the expectation that items that receive the same label should be judged as similar. Similarity judgments are particularly prone to strategic cognitive processing of this sort (Goldstone, 1994b). People may have an explicit strategy of increasing the similarity rating between two objects by a certain amount if the objects were previously grouped in the same category together. They may also have a strategy of decreasing the similarity between objects that are assigned to different categories, or a combination of these two strategies.

An alternative to explicitly basing similarity judgments on prior category assignments is that category learning actually changes the description of the categorized objects. Object properties that the different members of a category have in common may become selectively emphasized because of their relevance for categorization. After category learning, these relevant properties continue to be particularly important for representing objects, and thus in a similarity rating task, objects that share these important properties will tend to seem more similar. By definition, the objects that share these important properties will be the objects that were categorized together. The difference between the former Strategic Judgment Bias account and this Altered Object Description account is that in the former account, the object descriptions themselves are not influenced by categorization; only the similarity judgments, acting on unchanged object representations, are influenced by categorization. The difference between these positions is important because several researchers have argued that learned categorical perception shows that object representations are influenced by concept learning (Goldstone and Steyvers, in press; Goldstone et al., 2000; Livingston et al., 1998). However, we would only have confidence in this position if the Altered Object Description account can be unambiguously supported.

Although similarity ratings are particularly prone to task demands and explicit strategies, a version of the Strategic Judgment Bias account can also be developed for psychophysical measures. For example, the finding by Goldstone (1994a) that people have greater sensitivity at distinguishing objects that belong to different categories rather than the same category may be because participants adopt the strategy of labeling the objects they see with their assigned categories, and respond 'same' or 'different' depending on whether the objects receive the same label. This strategy of relying on coded labels rather than perceptual traces is particularly common if delays and therefore memory demands are introduced (Pisoni, 1973). Goldstone argues against this labeling interpretation because objects that receive the same label but differ on a category-relevant dimension become increasingly discriminable with category training. For example, if a category boundary exists on the size dimension between 3 and 4 cm, then Goldstone found that training on the size categorization promotes discrimination between 2 and 3 cm objects. This result does argue against a deterministic labeling process, but not a stochastic labeling process in which a 3 cm object is sometimes labeled incorrectly because it is close to the category boundary (Liberman et al., 1957). The discrimination between 2 and 3 cm objects may be improved by making use of category labels even if they are incorrectly assigned, under the assumption that the 3 cm object will receive an incorrect label more often than will the 2 cm object.

The primary purpose of the present experiment is to obtain evidence that can distinguish between a Strategic Judgment Bias and an Altered Object Description account of why category learning affects categorized objects' similarity ratings and psychophysical measures of similarity. In this study we use similarity ratings rather than same/different judgments, because similarity ratings are typically more sensitive to conceptual processing (Goldstone, 1994b; Goldstone & Barsalou, 1998). Although similarity ratings are more sensitive to task demands, if object descriptions are altered by category learning, then similarity ratings would probably also be more sensitive to this change. To find evidence for the Altered Object Description account that could not be explained by the Strategic Judgment Bias account, we examined the difference between similarity ratings given to a neutral, uncategorized object. Specifically, we determined the absolute difference, across participants, between the similarity ratings for categorized objects to a neutral object. Consider a situation where objects A and B belong to one category, objects C and D belong to another category, and object E has not been associated to any category. If category learning alters objects' descriptions, then we predict that the similarity ratings between objects A and E should become more similar to the similarity ratings between objects B and E after category learning. That is, if categorizing objects A and B together makes their encoded descriptions more similar to each other, then they should enter into similar similarity relations with other objects. In the limiting case, if objects A and B developed identical representations because of their category membership and if there were no noise in similarity ratings, then their similarity ratings to other objects would be identical. The Altered Object Description account might also predict that objects that are placed in different categories should have relatively dissociated similarity ratings to other neutral objects. If categorization accentuates stimulus aspects that differentiate between learned categories, then similarity ratings involving objects from different categories might be expected to become less related. To the extent that objects have more related similarity ratings when categorized together, we would conclude that properties common to a category are perceptually emphasized. To the extent that objects have more dissociated similarity when placed in separate categories, we would conclude that properties that differentiate between categories are emphasized.

In contrast to these predictions of an Altered Object Description account, by a Strategic Judgment Bias account, category learning should not systematically affect the similarity ratings involving object E because it has not been previously categorized. The strategy of basing similarity ratings on whether or not two stimuli have received the same category label is not applicable when one stimulus has not been assigned to a category. Objects A, B, C, D, and E are chosen so as to be approximately equally similar to each other. As such, the Strategic Judgment Bias account should predict that objects A, B, C, and D should be rated equally similar relative to

E, before as well as after categorization learning. Assuming proper counterbalancing, any bias to increase the subjective similarity between objects that share a label would not predict greater concordance of similarity ratings between objects A and E and objects B and E.

In addition to collecting data on similarity ratings of two objects relative to a neutral object, we also collect direct similarity ratings between two objects that each receive a category label. However, these direct similarity ratings provide less diagnostic evidence with respect to the two accounts of categorization's influence on similarity. By the Altered Object Description account, objects that are placed in the same category would be expected to have their similarity increased if features shared by members of the same category are emphasized. A second, non-mutually exclusive possibility by this account is that objects that are placed in different categories have their similarity decreased if features that distinguish between categories are emphasized. Features shared by the members of a category need not always be distinctive, and vice versa. For example, when categorizing birds as either cardinals or blue jays, the feature 'bright red' is diagnostic of cardinals but is not a common feature possessed by all cardinals because only males are bright red. The feature 'has wings' is a feature common to all cardinals, but is not diagnostic because blue jays also possess this feature. These two possible predictions of the Altered Object Description account are also consistent by a Strategic Judgment Bias account.

Although the direct similarity ratings between categorized objects are not useful for distinguishing Altered Object Description and Strategic Judgment Bias accounts, they are useful for diagnosing whether category learning has as its primary impact increased similarity between objects in the same category (within-category *compression*, also known as acquired equivalence), decreased similarity between objects in different categories (between-category *expansion*, also known as acquired distinctiveness), or both. Whereas some researchers have found the primary impact of category learning to be decreasing people's confusions when making discriminations involving perceptual dimensions relevant to the categorization (Goldstone, 1994a; Levin & Beale, 2000), other researchers have found category learning to primarily increase similarity ratings between objects belonging to the same category (Kurtz, 1996; Livingston et al., 1998). One possible reason for this may be that psychophysical measures of discriminability seem to promote expansion whereas similarity ratings promote compression (Levin & Beale, 2000). Another proposed resolution is that if the categorized objects are extremely similar, then expansion is found; otherwise, compression is found (Livingston et al., 1998). In the current experiment, we systematically altered the between- and within-category similarities, such that objects belonging to different categories were either similar or dissimilar, as were objects belonging to the same category. One hypothesis is that if two categories are very similar to each other but must be treated differently, then there is pressure to find aspects that discriminate between the categories, yielding expansion effects (Goldstone, 1996). Similarly, if the objects within one category are highly different from each other, then there is pressure to find or emphasize aspects that make them more similar, yielding compression effects. This framework would predict particularly large compression effects for categorizations with low rather than high within-category similarity, and particularly large expansion effects for categorizations with high rather than low between-category similarity.

3. Method

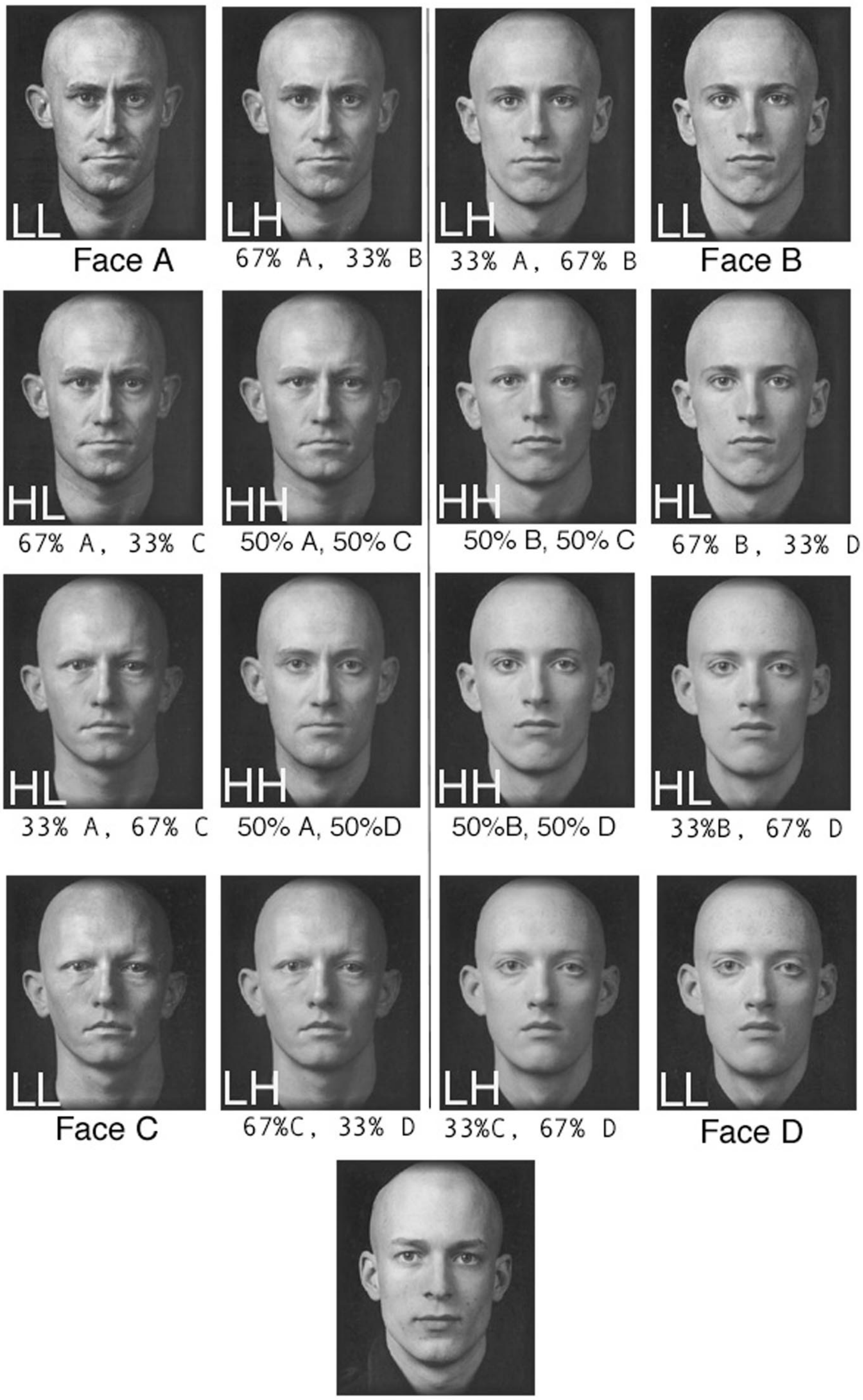
3.1. Participants

One hundred ninety-three undergraduate students from Indiana University served as participants in order to fulfill a course requirement and were run in parallel sessions. Four participants did not meet the learning criterion of 80% correct in the last 72 trials of the categorization task and were excluded from further analyses. Of the remaining 189 subjects, 50 were randomly assigned to the HH group, 45 to the HL group, 48 to the LH group, and 46 to the LL group.

3.2. 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 morphs are shown on the corners of Fig. 1 and are labeled as Faces A, B, C, and D. An additional Face E, also shown in Fig. 1, was used as a neutral comparison face. A set of 16 faces was created for the categorization portion of the experiment. Four of these faces were undistorted faces, and the other faces blended in different proportions two of the faces from the set {A, B, C, D}. Using morphing routines developed by Steyvers (1999), 60 control points were located on each of the original faces at salient face positions such as corners of eyes, pupils, cheekbones, and the midway point of the lower lip. The face labeled '66% Face A and 33% Face B' was created by moving each of the control points to a position located 66% of the distance from B to A starting with Face B, and assigning a gray-scale value to each pixel that was a weighted average of A and B values, with three times more weight for A than B values. The other faces in Fig. 1 were created in a similar fashion.

The five original Faces A, B, C, D, and E were selected from a larger database of 62 bald faces. The subjective similarity between each pair of these 62 faces was obtained by a method described by Goldstone (1994a). The five faces were selected because each possible pair from this set of faces received average subjective similarities that were within 20% of any other pair. Face E is as similar to each of the Faces A–D as they are to each other. As such, E did not systematically have intermediate values on the dimensions shown in Fig. 1. Each face was displayed with 256 gray-scale brightness values per pixel (one pixel is 0.034 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.


LL
Face A
LH
67% A, 33% B
LH
33% A, 67% B
LL
Face B
HL
67% A, 33% C
HH
50% A, 50% C
HH
50% B, 50% C
HL
67% B, 33% D
HL
33% A, 67% C
HH
50% A, 50%D
HH
50%B, 50% D
HL
33%B, 67% D
LL
Face C
LH
67%C, 33% D
LH
33%C, 67% D
LL
Face D

3.3. Procedure

Participants were presented with three tasks in the 55 min experiment: a pre-categorization similarity rating task, a category learning task, and a post-categorization similarity rating task identical to the first task. Each of the similarity rating tasks required approximately 15 min, and the category learning task required 25 min.

Each participant was randomly assigned to one of four different groups: HH, HL, LH, LL, where the first letter of each group refers to whether the within-category similarity of faces was high ('H') or low ('L'), and the second letter refers to whether the between-category similarity of faces was high or low. The faces were categorized by the vertical boundary line shown in Fig. 1, such that the eight faces on the left belonged to one category and the eight faces on the right belonged to a second category. Each participant was shown a subset of four of the 16 of Fig. 1 during their similarity rating and category learning tasks. The particular four faces shown to each group of participants are depicted in Fig. 1. The LL faces are characterized by having low within- and between-category similarity; the four faces are completely unrelated to each other. The HH faces are characterized by high within- and between-category similarity; each face shares half of its identity with one face, and the other half of its identity with another face. Half of its identity is shared with a face in its same category, and the other half is shared with a face in the opposite category. The HL faces have high within-category similarity and low between-category similarity, and as such, the categories should be particularly easy to learn. As shown in Fig. 1, faces from the same category of the HL set are based on the same faces, differing only in the percentages of these two faces, whereas faces from different categories of the HL set are not based on any of the same faces at all. Conversely, faces from the LH set have low within-category similarity and high-between category similarity. Each face has a similar face in the opposite category, and has no original face element in common with the other face in its category.

During each trial of the categorization task, participants saw one of the four faces from their set (never seeing Face E), and categorized it by pressing either 'A' or 'B' on the keyboard, with feedback on each trial from the computer showing an 'X' for incorrect responses or a check for correct responses, and also indicating the correct category assignment for the face. Participants saw each of the four faces 54 times, yielding 216 categorization trials. The order in which the faces were presented was randomized. Participants received short breaks every 54 trials. During these breaks, the computer displayed the participants' accuracy and average response time on the previous block of trials.

Fig. 1. The four Faces A, B, C and D are blended in different proportions to create the stimuli for four categorization conditions, and Face E is used as a neutral, unlabeled face. The thin vertical line indicates the category boundary for all four conditions. The faces labeled LH belong to the categorization condition with low within-category similarity and high between-category similarity. The HL faces involve high within-category similarity and low between-category similarity. The LL faces have low within- and between-category similarity. The HH faces have high within- and between-category similarity.

During the two similarity rating tasks, participants saw two faces on the computer screen, selected from the four faces in their set shown in Fig. 1 and including another unrelated Face E. The faces appeared side by side, separated by 5 cm. The first face appeared on the left side of the screen for 2500 ms, followed by a blank screen lasting 250 ms, followed by the second face that also appeared for 2500 ms. After the second face was removed, similarity ratings were collected by having participants move a cursor on the screen with a mouse. A horizontal line was drawn along the bottom of the screen. The left and right edges of the line were labeled ‘Not Very Similar’ and ‘Highly Similar’, respectively. The length of the horizontal line was 20 cm, and was divided into 500 units. Participants were instructed to press a button on the mouse when the cursor was positioned along the line at their subjective similarity estimate. Each of the five faces was paired with every other face, yielding ten comparisons, and these ten comparisons were each repeated seven times, yielding 70 ratings in all. The ordering of the ten comparisons was randomized, as was the left/right ordering of the two compared faces.

4. Results

4.1. Categorization performance

Analyses of variance on percent correct and response time (RT) were conducted, using learning group (HH, HL, LH, and LL) as a between-subject variable and block of trials (first, second, and third) as a within-subject variable. A highly significant main effect of block indicated faster ($F(2, 370) = 96.55$, $P < 0.001$) and more accurate responses ($F(2, 370) = 130.43$, $P < 0.001$) with increasing practice. A highly reliable main effect of learning group for RT ($F(3, 185) = 19.32$, $P < 0.001$) and accuracy ($F(3, 185) = 23.46$, $P < 0.001$) showed that both learning groups with high within-category similarity showed the best and comparable performance (HH, 627 ms and 96.6% correct; HL, 588 ms and 97.7% correct), followed by the learning group with low within-category and low between-category similarity (LL, 696 ms and 95.7% correct), which significantly differed only from the HL group (Scheffe’s post-hoc test, $P < 0.05$). Performance in the learning group with low within-category and high between-category similarity was significantly worse compared to the performance of the other three groups (LH, 836 ms and 92.3% correct, Scheffe’s post-hoc tests, $P < 0.05$). A reliable interaction between block and learning group showed that learning group differences decreased considerably across blocks for both RT ($F(6, 370) = 5.52$, $P < 0.001$) and accuracy ($F(6, 370) = 27.90$, $P < 0.001$).

4.2. Similarity ratings

A first analysis tested whether the categorization task changed the judged similarity between the faces that were categorized. For this analysis, an ANOVA was conducted on the similarity ratings involving only categorized faces, using learning group (HH, HL, LH, and LL) as a between-subject variable and condition (within-category versus between-category comparison) and time (before versus after category learning) as within-subject variables. There was a significant main effect of learning group ($F(3, 185) = 33.63$, $P < 0.001$) showing that both learning groups with high within-category similarity judged the faces to be more similar (ratings of 330 and 318 for HH and HL, respectively) than the two learning groups with low within-category similarity (245 and 258 for LL and LH, respectively). The main effect of condition was also highly significant ($F(1, 185) = 181.84$, $P < 0.001$). With a rating difference of 76, within-category pairs were judged more similar than between-category pairs. A reliable interaction between learning group and condition revealed that this difference was true for the HL group (192) ($F(1, 185) = 273.81$, $P < 0.001$), for the HH group (151) ($F(1, 185) = 187.29$, $P < 0.001$) and for the LH group (29) ($F(1, 185) = 34.22$, $P < 0.001$), whereas in the LL condition between-category pairs were judged more similar than within-category pairs (-65) ($F(1, 185) = 6.26$, $P < 0.01$). The interaction between condition, time, and learning group did not approach significance ($P > 0.80$). Instead, there was a condition × time interaction ($F(1, 185) = 3.89$, $P = 0.050$). As depicted in the top panel of Fig. 2, participants judged between-category pairs to be less similar after the categorization task than before ($F(1, 185) = 5.49$, $P < 0.05$), whereas the judgments for within-category pairs remained unaffected by category learning ($F < 1$). Hence, for ratings of the similarity between the categorized faces, we found an overall effect of expansion for faces previously assigned to different categories, and no influence of categorization for faces that were previously assigned to the same category.

In a second analysis we tested whether the categorization task changed the similarity of the categorized faces relative to the neutral, non-categorized Face E. We calculated the absolute difference for each participant between their average ratings given to pairs of comparisons that involved a categorized face and the neutral face and averaged separately across those pairs where the two categorized faces belonged to the same category (a difference of within-category comparisons) and those pairs where the two categorized faces belonged to different categories (a difference of between-category comparisons). An ANOVA was conducted on these differences, using learning group (HH, HL, LH, and LL) as a between-subject variable and condition (within-category versus between-category comparison) and time (before versus after categorization) as within-subject variables. The main effect of learning group was significant ($F(3, 185) = 15.83$, $P < 0.001$), indicating that the two learning groups of low between-category similarity rated stimulus pairs involving the neutral face more differently (differences of 121 and 97 for LL and HL, respectively) than the two groups of high between-category similarity (differences of 66 and 79 for HH and LH, respectively). A reliable main effect of condition ($F(1, 185) = 88.56, P < 0.001$) showed that the overall difference for within-category comparisons was smaller than the difference for between-category comparisons (69 and 111, respectively). A learning group × condition interaction revealed that this effect of a smaller difference for within-category comparisons occurred in the HH group ($F(1, 185) = 50.09$, $P < 0.001$), HL group ($F(1, 185) = 85.86$, $P < 0.001$), and LL group ($F(1, 185) = 12.14$, $P < 0.01$), but not in the LH group ($F(1, 185) = 1.16, P = 0.283$). Of special interest is the significant

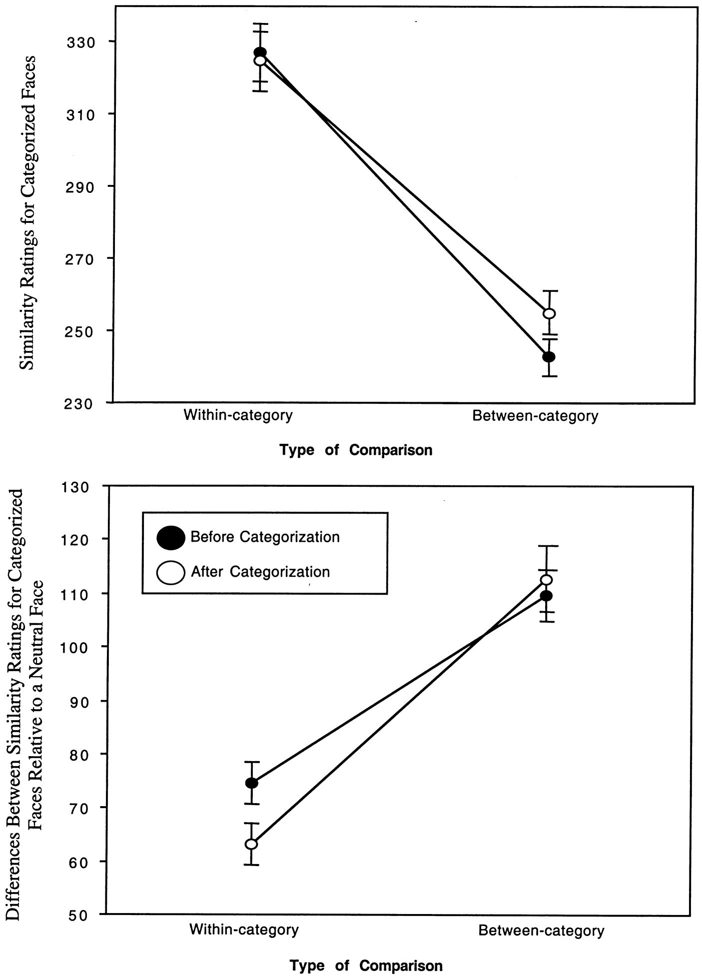


Fig. 2. The top panel shows the similarity ratings between faces that belong to the same category (Within) or different categories (Between), before and after category learning. The lower panel shows the absolute difference between pairs of similarity ratings involving a neutral face that was not shown during category learning. These ratings could either involve faces that belonged to the same category or to different categories.

condition × time interaction ($F(1, 185) = 5.35, P < 0.05$). As depicted in the bottom panel of Fig. 2, the difference of between-category comparisons did not change over time ($F < 1$), whereas the difference between within-category comparisons became smaller after the categorization task ($F(1, 185) = 10.51, P < 0.01$). This effect was not modulated by learning group ($P > 0.25$). Hence, similarity ratings relative to a neutral face became more similar for faces that were assigned to the same category (compression effect), whereas they remained unaffected for categorized faces that were assigned to different categories.

5. Discussion

In addition to finding an influence of category learning on similarity ratings for the categorized faces, category learning also affected the differences between similarity ratings for categorized faces when compared to a neutral face. The former dependent measurement may be caused by a tendency to base similarity ratings on category label similarity specifically. Thus, it might be that participants rated two categorized faces as similar or different to the extent that they belonged to the same category or to different categories, respectively. However, the impact of the categorization task on the differences between similarity ratings for categorized faces relative to a neutral face cannot be explained by such a label-based bias, because the neutral face was not assigned to a category. Our finding that after category learning the similarity ratings of same-category faces relative to a neutral face became more similar therefore suggests that category learning has changed the representations of the objects themselves. When two objects are placed in the same category, the objects' common features are apparently emphasized, and hence their object descriptions become more similar. As the objects become more similar, there is an increased positive dependency between their judged similarities to other objects.

The two significant influences of category learning on measures of similarity were generally replicated across all of the four conditions of within- and between-category similarity. There was no significant interaction between the learned categorization and changes in perceived similarity. This null result was apparently not due to an insufficiently strong manipulation of category similarities. Our learning groups differed widely in their ease. For example, the learning groups with high within-category similarity learned their categories much more quickly and judged their faces as more similar than did the learning groups with low within-category similarity. Although future research could prove otherwise, our current conclusion is that different category structures, including those with either widely separated or highly overlapping categories, should all influence the representation and similarity of the categorized objects to equivalent extents.

The two dependent measures of object similarity suggested different impacts of category learning on perceived similarity. While the similarity ratings for categorized faces revealed only a decrease of similarity between faces of different categories (hence, an expansion effect), the differences between similarity ratings for categorized faces relative to a neutral face showed only an increase in similarity between faces of the same category (hence, a compression effect). This difference supports our conjecture that the two dependent measures are contaminated by strategic use of category labels to a different extent. In the case of categorized objects, judgment performance reflects the prominence of category labels. The presence of only an expansion, but not compression, effect for similarity ratings between categorized faces was not predicted, but a post-hoc account is available. When two objects are compared that have been assigned to different categories, then participants give relatively low similarity ratings because of the obvious discrepancy between their category labels. However, objects assigned to the same category may not have their similarity increased very much because labels may only be prominent when they show variation (Garner, 1962; Goldstone, Medin, & Halberstadt, 1997). For example, Goldstone et al. (1997) showed that when people are given pairs of objects to rate for similarity, they are unlikely to code objects in terms of the thickness of their lines unless the compared objects differ in their line thickness. Otherwise, line thickness is a 'backgrounded' dimension that is ignored. Similarly, if explicitly labeled objects share a common label, then this commonality appears to be backgrounded, given that it does not increase similarity. However, if the compared objects have different labels, then this labeling dimension is promoted to the foreground because of its variation, and consequently the objects' similarity is decreased.

A different mechanism is required to account for the differences between similarity ratings for categorized faces relative to a neutral face. By this measure, faces that are categorized together became more similar, whereas no effect was observed for faces from different categories. The above explanation of backgrounded dimensions becoming foregrounded when variation exists along them should no longer be relevant given that the neutral face has no category label at all. Instead, judgments must rather be based on the representations of the objects themselves. The form of this representational change could have been either within-category compression or between-category expansion. For this measure, emphasizing features shared by category members may be more powerful than emphasizing features that distinguish between categories because participants are likely to form complete images for each of the two categories. Previous work has suggested that when image-like concept representations are encouraged, concepts tend to be represented in terms of their within-category commonalities rather than the features that distinguish them from other categories (Goldstone, 1996). Once properties shared by the members of a category have been emphasized, these properties will bring representations for members of the same category closer together.

Thus, the current data pattern suggests that the traditional use of similarity ratings to measure similarity does not exclusively measure the representational similarity of the objects per se, but also measures the similarity of their associated labels and categories. By contrast, measuring the similarity of two objects indirectly by measuring their similarity to other objects may provide a less contaminated gauge of their representational similarity (for another use of indirect similarities, see Landauer & Dumais, 1997). At a minimum, given that one measure of similarity shows expansion while the other shows compression, we can be fairly confident that implicit labeling of objects is not contaminating our indirect measure of learned similarity. Other researchers have found compression effects with traditional, direct similarity ratings (Livingston et al., 1998), and so more research will be needed to identify the factors that determine whether compression or expansion effects are observed. However, given that we used the same stimulus materials and rating task for the direct and indirect measures, we can be confident that one such determining factor is whether the compared objects are explicitly labeled or not.

In summary, our findings contribute to our understanding of what is the psychological impact of learning new categories. By showing that after category learning two members of the same category agree more in their judged similarity to a third non-categorized object, we have provided stronger evidence than before that grouping two objects together changes their internal descriptions. The elements that the objects share, elements that by definition specify their category, become more important parts of the objects' descriptions. This result may even have some relevance to the classic debate between Gibson and Gibson (1955a,b) and Postman (1955). The dispute as characterized by the Gibsons concerns the question "Is learning a matter of enriching previously meager sensations or is it a matter of differentiating previously vague impressions?" (p. 34). According to the enrichment view, perceptions change as sensory information becomes associated with and enriched by accompanying information such as labels, outcomes, or contexts. According to the Gibsons' differentiation view, perceptions change not by becoming connected to learned associations, but by becoming more connected to the external world and its properties. Thus, it is assumed that learning involves responding to previously ignored sensory information. An enrichment view would suggest that objects become associated with their accompanying category labels, and that these labels are later triggered when the objects are presented. This can account for the decreased similarity found for objects associated with different category labels. However, it does not provide a natural account for how similarities involving a neutral, unlabeled object change systematically with learning. There appears to be more happening during category learning than simply assigning objects with labels that enter into subsequent judgments. Category learning also changes what object properties are emphasized. Similarities between the objects are directly influenced by the similarity of their categories, but also the categories indirectly influence the objects' similarities by causing the objects to appear differently.

Acknowledgements

Many useful comments and suggestions were provided by Stevan Harnad, Ken Livingston, and Robert Nosofsky. The authors wish to thank Casey Haines, Matt Licht, Erica Martin and Yasuaki Sakamoto for assistance in running participants. This research was funded by NIH Grant R01 MH56871-01A2 and a Gill fellowship to the first author, and NIMH Grant 12717 to the third author. More information about the laboratory can be found at http://cognitrn.psych.indiana.edu/

References

Arnoult, M. D. (1957). Stimulus predifferentiation: some generalizations and hypotheses. *Psychological Bulletin*, *54*, 339–350.

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, *57*, 217–239.

Beale, J. M., & Keil, F. C. (1996). Categorical perception as an acquired phenomenon: what are the implications? In L. Smith, & P. Hancock (Eds.), *Neural computation and psychology: workshops in computing series* (pp. 176–187). Berlin: Springer-Verlag.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.

Cantor, J. H. (1965). Transfer to stimulus pretraining to motor paired-associate and discrimination learning tasks. In L. P. Lipsitt, & C. C. Spiker (Eds.), *Advances in child development and behavior* (Vol. 2, pp. 19–58). New York: Academic Press.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*, 303–306.

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.

Gibson, J. J., & Gibson, E. J. (1955a). Perceptual learning: differentiation or enrichment? *Psychological Review*, *62*, 32–41.

Gibson, J. J., & Gibson, E. J. (1955b). What is learned in perceptual learning? A reply to Professor Postman. *Psychological Review*, *62*, 447–450.

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

Goldstone, R. L. (1994b). The role of similarity in categorization: providing a groundwork. *Cognition*, *52*, 125–157.

Goldstone, R. L. (1996). Isolated and interrelated concepts. *Memory and Cognition*, *24*, 608–628.

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

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

Goldstone, R. L., Medin, D. L., & Halberstadt, J. (1997). Similarity in context. *Memory and Cognition*, *25*, 237–255.

Goldstone, R. L., & Steyvers, M. The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General*.

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). Mahwah, NJ: Lawrence Erlbaum Associates.

Harnad, S. (1987). *Categorical perception*. Cambridge University Press: Cambridge.

Katz, P. A. (1963). Effects of labels on children's perception and discrimination learning. *Journal of Experimental Psychology*, *66*, 423–428.

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

Kurtz, K. J. (1996). Category-based similarity. In G. W. Cottrell (Ed.), *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*, 290.

Landau, B., & Shipley, E. (1996). Object naming & category boundaries. In A. Stringfellow (Ed.), *Proceedings of the Boston University Conference on Language Development*. Brookline, MA: Cascadillo Press.

Landauer, T., & Dumais, S. T. (1997). A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Levin, D. T., & Beale, J. (2000). Categorical perception occurs in newly learned faces, other-race faces, and inverted faces. *Perception and Psychophysics*, *23*, 1153–1169.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*, 358–368.

Lin, E. L., & Murphy, G. L. (1997). Effects of background knowledge on object categorization and part detection. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1153–1169.

Livingston, K. R., Andrews, J. K., & Harnad, S. (1998). Categorical perception effects induced by category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 732–753.

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

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, *13*, 253–260.

Pisoni, D. B., Aslin, R. N., Perey, A. J., & Hennessy, B. L. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 297–314.

Postman, L. (1955). Association theory and perceptual learning. *Psychological Review*, *62*, 438–446.

Pothos, E. M., & Chater, N. (1997). Classification and prior assumptions about category "shape": new evidence concerning prototype and exemplar theories of categorization. *Proceedings of the Nineteenth*

Annual Conference of the Cognitive Science Society (pp. 620–625). Mahwah, NJ: Lawrence Erlbaum Associates.

Repp, B. H., & Liberman, A. M. (1987). Phonetic category boundaries are flexible. In S. Harnad (Ed.), *Categorical perception* (pp. 89–112). Cambridge: Cambridge University Press.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd, *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

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* (Vol. 31, pp. 305–354). San Diego, CA: Academic Press.

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

Schyns, P., Goldstone, R. L., & Thibaut, J.-P. (1998). The development of features in object concepts. *Behavioural and Brain Sciences*, *21*, 1–54.

Steyvers, M. (1999). Morphing techniques for manipulating face images. *Behavioral Research Methods, Instruments, and Computers*, *31*, 359–369.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, *7*, 49–63.

Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science*, *18*, 221–281.