

# Conceptual interrelatedness and caricatures

ROBERT L. GOLDSTONE

*Indiana University, Bloomington, Indiana*

MARK STEYVERS

*University of California, Irvine, California*

and

BRIAN J. ROGOSKY

*Indiana University, Bloomington, Indiana*

Concepts are interrelated to the extent that the characterization of each concept is influenced by the other concepts, and are isolated to the extent that the characterization of one concept is independent of other concepts. The relative categorization accuracy of the prototype and caricature of a concept can be used as a measure of concept interrelatedness. The prototype is the central tendency of a concept, whereas a caricature deviates from the concept's central tendency in the direction opposite the central tendency of other acquired concepts. The prototype is predicted to be relatively well categorized when a concept is relatively independent of other concepts, but the caricature is predicted to be relatively well categorized when a concept is highly related to other concepts. Support for these predictions comes from manipulations of the labels given to simultaneously acquired concepts (Experiment 1) and of the order of categories during learning (Experiment 2).

Concepts seem to be connected simultaneously to each other and to the external world. On one hand, concepts seem to gain their meaning through the role that they play within a network of concepts (Collins & Quillian, 1969; Field, 1977). The notion of a “conceptual web” through which concepts all mutually define one another has been highly influential in all of the major fields that comprise cognitive science, including linguistics (Saussure, 1915/1959), computer science (Lenat & Feigenbaum, 1991), psychology (Landauer & Dumais, 1997), and philosophy (Block, 1999). However, there is also dissatisfaction in some quarters with the circularity of this conceptual web account. Researchers have argued that concepts must be grounded in the external world rather than merely related to each other (Harnad, 1990). The British empiricists argued that our conceptual ideas originate in recombinations of sensory impressions (Hume, 1740/1973). More recently, Barsalou (1999; Goldstone & Barsalou, 1998; Solomon & Barsalou, 2001) has argued that concepts are not amodal, completely abstracted symbols, but rather are intrinsically perceptually based.

In an attempt to reconcile the argument for a conceptual web and that for externally grounded concepts, Gold-

stone (1996) described a continuum between purely isolated and purely interrelated concepts, arguing that a concept is interrelated with other concepts to the extent that its characterization is influenced by them. For example, a Caucasian in a white-majority culture may characterize black people in terms of their relation to white people (Linville & Jones, 1980). This Caucasian's concept of *black person* would be dependent on and influenced by his or her *white-person* concept, with the possible consequence that his or her representation of black people may be distorted to accentuate perceived differences between the two races. Goldstone's empirical basis for the continuum was the convergence of a set of experimental manipulations and measures of conceptual interrelatedness. A set of manipulations was designed to influence the degree of interrelatedness between simultaneously acquired concepts, and the influence of these manipulations was gauged by a set of measures of interrelatedness. These experiments gave support to the hypothesis that small experimental manipulations were capable of changing how influential one concept was on another concept's representation and processing. The goal of the present experiments is to further test the claim for a continuum between isolated and interrelated concepts, using rich and naturalistic stimuli, and new manipulations and measures of interrelatedness.

## Isolated and Interrelated Concepts

In evaluating the claim for a continuum between isolated and interrelated concepts, it is helpful to consider theories at the two poles of the continuum. We will consider representative models of isolated and interrelated concepts, leaving a fuller description to Goldstone (1996).

---

This research was funded by NIH Grant MH56871 and NSF Grant 0125287. The authors wish to express their thanks to Evan Heit, John Kruschke, Elinor McKone, Robert Nosofsky, Sarah Queller, and Jim Sherman for helpful suggestions on this work. Correspondence concerning this article should be addressed to R. Goldstone, Psychology Department, Indiana University, Bloomington, IN 47405 (e-mail: rgoldsto@indiana.edu). Further information about the laboratory can be found at <http://cognitntrn.psych.indiana.edu>.

Conceptual interrelatedness is a component of many linguistic treatments of concepts. Ferdinand de Saussure (1915/1959) argued that all concepts are completely “negatively defined”—that is, defined solely in terms of other concepts. He contended that “language is a system of interdependent terms in which the value of each term results solely from the simultaneous presence of the others” (p. 114) and that “concepts are purely differential and defined not in terms of their positive content but negatively by their relations with other terms in the system” (p. 117). This notion has evolved into the modern treatment of semantic networks (Collins & Quillian, 1969; Quillian, 1967). In these networks, concepts are represented by nodes in a network and gain their functionality by their links to other concept nodes. Often, these links are labeled, in which case different links refer to different kinds of relations between nodes. *Dog* would be connected to *animal* by an *is-a* link, to *bone* by an *eats* link, and to *paw* by a *has-a* link. Lenat and Feigenbaum (1991) argued that interconceptual linkages are sufficient to establish conceptual meanings even without any external grounding of the concepts. A computational approach to word meaning that has received considerable attention recently has been to base word meanings solely on the patterns of co-occurrence between a large number of words in an extremely large text corpus (Burgess, Livesay, & Lund, 1998; Burgess & Lund, 2000; Landauer & Dumais, 1997). Mathematical techniques are used to create vector encodings of words that efficiently capture their co-occurrences. If two words, such as “cocoon” and “butterfly,” frequently co-occur in an encyclopedia or enter into similar patterns of co-occurrence with other words, then their vector representations will be highly similar. The meaning of a word—that is, its vector in a high dimensional space—is completely based on the word’s contextual similarity to other words. Finally, researchers have argued that concepts are frequently characterized by their associative relations to other concepts. Barr and Caplan (1987), by having subjects list features associated with words, provided evidence that many concepts are characterized by what they called “extrinsic features,” or features that are “represented as the relationship between two or more entities” (p. 398).

From the theories outlined above, one might conclude that concepts cannot stand alone and that there could not be such a thing as a system with only one concept (Stich, 1983). However, if one looks at the field of pattern recognition rather than at theories inspired by linguistics, then examples of isolated concepts become apparent. One way to represent an isolated concept is as a feature detector. A feature detector can become active when an input with a particular perceptual feature is present. A concept can also be represented as a template in a physical or more abstract space (Edelman, 1999). In this view, concepts are applied to external objects in order to determine category membership based on the perceived input patterns of the objects. If patterns are categorized by how they compare with stored templates for categories, the representation of one category does not depend on those of other categories.

A category’s representation is simply the image that best matches its members. It is possible to have a feature detector or template for a concept without having any other concepts in the system. Categorizing an object may require comparing the relative degrees of the object’s match to the representations of the candidate categories (Nosofsky, 1986), but if the categories themselves are represented by templates or feature detectors, then each can exist independently of the other categories.

A comparison of some representative examples of interrelated and isolated concepts suggests a useful heuristic for assessing degree of interrelatedness. A Concept X is dependent on Concept Y to the extent that Concept X cannot exist without Concept Y. If the concept *vermicelli* is represented as “thinner pasta than spaghetti” then no system could possess *vermicelli* without also possessing *spaghetti*. However, if the concept *vermicelli* is represented by “a long pasta with a typical diameter of 6 mm,” then possession of the concept of *vermicelli* does not depend on possession of the concept of *spaghetti*. Partial degrees of dependency owe to the multifaceted nature of conceptual representations. One’s concept of *vermicelli* may incorporate both characterizations given above, and the relative importance of these characterizations determines how much *vermicelli*’s representation is affected by the presence or absence of a *spaghetti* concept.

### Prototypes and Caricatures

Consider the example of the two categories shown in Figure 1. Categories A and B each have six members, and each member has a unique combination of values on Dimensions 1 and 2. We will define the prototype of a category as the central tendency of the category along each of its dimensions (Posner & Keele, 1968; Reed, 1972). The prototype of Category A has a value close to 2 on Dimension 1, whereas Category B’s prototype has a value close

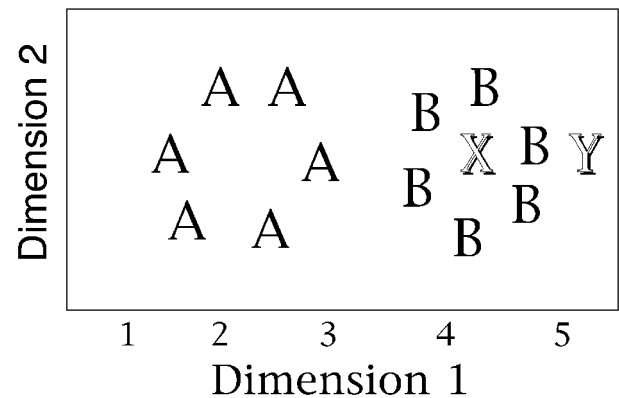


Figure 1. Category A and Category B members have average values of about 2 and about 4 respectively, on Dimension 1. Accordingly, the prototypical value for Category B on Dimension 1 is close to 4 (shown by the “X”), but the caricature is displaced away from the prototype in the direction opposite Category A (shown by the “Y”).

to 4 on Dimension 1. In this usage, a prototype for a category would be the physical object that assumes average values along the dimensions that comprise the category's members, rather than the internal representation of the category (Lakoff, 1987). In the experiments that follow, we use uniform distributions of dimension values in constructing categories, and, consequently, our description of a category prototype remains the same if we define "central tendency" as the average or median of that category.

We will define a caricature of a category as an object that assumes dimension values that depart from the central tendency of the category in the opposite direction of the central tendency of other, simultaneously acquired concepts. In Figure 1, "X" represents the prototype of Category B, and "Y" represents a caricature of Category B. Y has a value close to 5 on Dimension 1, and this value is a distortion of Category B's central tendency away from Category A. This definition of "caricature" captures the intuitive notion of a caricature as an exaggeration. If Category B has a large value on Dimension 1 (relative to Category A), then the caricature of Category B exaggerates this large value, in the same way that a newspaper caricature exaggerates distinctive facial features of a politician. The term "caricature" is often reserved for exaggerations of a member of a category from the category's prototype. A caricature of a face typically distorts the face away from a prototypical face (Rhodes, Brennan, & Carey, 1987). Although our use of the term "caricature" is consistent with this traditional use, in that both refer to distortions of a representation away from those of contrasting objects, our use extends the traditional use to cover situations in which the contrasting objects comprise an alternative, rather than a superordinate, category. If the prototype of a superordinate category such as *face* is established, then it can be used as the standard away from which all caricatures of specific faces are distorted. In the present research, we pursue the alternative strategy of creating two categories based on novel faces and treating the caricature of one category as the distortion of one face away from the other, simultaneously acquired category, rather than as a distortion away from a prototypical face.

In the present experiments, we investigate the conditions under which the prototype or caricature of a category is most easily categorized. On one hand, one might predict better categorization for the prototype because it is, by definition, the item that is most similar to the members of its category (Posner & Keele, 1968; Rosch, 1975). On the other hand, one might predict better categorization accuracy for the caricature because it emphasizes a distinctive value for a category. In fact, several experiments have found a categorization advantage for caricatures relative to prototypes (Goldstone, 1996; Nosofsky, 1991; Palmeri & Nosofsky, 2001; Rhodes et al., 1987).

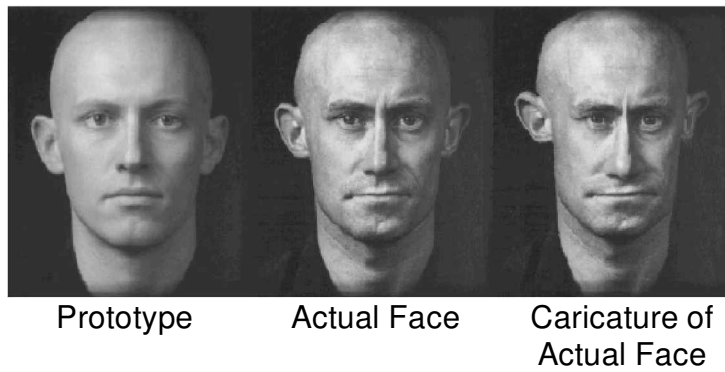
Our present goal is not to argue that caricatures are better categorized than prototypes or vice versa, but to identify experimental factors that modulate the benefit of one over the other. Specifically, experimental factors that pro-

mote relatively isolated concepts should tend to promote an advantage for prototypes over caricatures. In the absence of interconceptual influences, the representation that best exemplifies a concept will be its central tendency. If we try to represent Category B and do not know anything about Category A, then our best representation of Category B will be the point marked "X" in Figure 1. However, if a concept is characterized relative to other, simultaneously acquired concepts, then characterizations of the form, "Concept B members have relatively large Dimension 1 values in comparison with Concept A" will be formed. Caricatures fit these relational descriptions better than do prototypes. Nosofsky (1991) argued that classification of an object into a category depends on both its absolute similarity to members of the category and its relative similarity to members of all of the candidate categories. The present experiments explore factors that affect the relative importance of these absolute and relative determinants of categorization.

We instantiate caricatures and prototypes by face stimuli that are formed using automatic morphing software developed by Steyvers (1999). An example of caricaturization using this software is shown in Figure 2. A prototypical bald head was generated by combining 62 bald heads taken from Kayser (1997). To create this prototype without the blurred quality typical of superimposed face photographs (Busey, 1998; Galton, 1878), 127 defining points were found for each of the 62 heads, and the average location for each point was assigned to create the average face. The gray-scale values of corresponding pixels across the 62 heads were blended to create the gray-scale values for the average face. The caricature (shown on the right in Figure 2) of a particular face (shown in the middle of Figure 2) was generated by distorting the positions of each of the 127 defining points on the face by 20% away from the defining points on the average face. In the experiments to be reported, the caricatures are distortions away from an alternative face category rather than away from the prototype for the superordinate category of all of the faces shown in Figure 2.

### Manipulations of Conceptual Interrelatedness

The relative ease of categorizing prototypes and caricatures will be used as the measure of concept interrelatedness. The other main task is to develop experimental manipulations that are predicted to affect the interrelatedness of concepts. There already exists a body of literature suggesting such manipulations. Goldstone (1996) found that interrelated categories were generated when (1) participants were encouraged to look for features that discriminated between the learned categories, (2) categories were alternated frequently, (3) participants were practiced categorizers, and (4) stimuli were relatively undistorted versions of the categories' prototypes. In contrast, isolated categories were generated when (1) participants were encouraged to form images of the categories, (2) categories were alternated rarely, (3) participants were relatively un-



**Figure 2.** Using the method described by Steyvers (1999), the face labeled “prototype” is obtained by blending together faces of 62 bald men. The caricature of an actual face is obtained by shifting the appearance of the face away from the prototype.

practiced, and (4) stimuli were highly distorted versions of the categories’ prototypes. Niedenthal and Beike (1997) explored whether people’s self-concepts were relatively independent of other people or relationally defined. They found that self-concepts were relatively interrelated when participants were asked to consider their distinctive attributes, and when participants were younger siblings comparing themselves to older siblings rather than vice versa. McKenzie (1998, 1999) has explored the conditions under which finding out information about one hypothesis affects mutually exclusive alternative hypotheses. Even in situations in which participants are told that a patient has one and only one of two candidate diseases, evidence that increases participants’ confidence that the patient has one disease does not always decrease their confidence that the patient has the other disease. Concurrent rather than successive presentation of the two diseases makes the diagnoses of the diseases more (negatively) dependent on one another, as does mentioning both diseases when participants make confidence judgments about each disease.

The present research explores two experimental manipulations that might be expected to affect concept interrelatedness. In Experiment 1, the labels given to categories being acquired were manipulated. One category was given a positive label (“Club A”) and the other category was given a negation label (“not Club A”). Although this manipulation was used by Goldstone (1996), it has never been used with the caricature-versus-prototype measure of interrelatedness, and never with naturalistic stimuli. In Experiment 2, the order of learning the categories was manipulated by a test of the hypothesis that the first learned category would be relatively isolated, whereas the second category would tend to be characterized relative to the first category.

### Faces as Categories

One of the goals of the experiments was to explore interrelated and isolated categories using materials that are natural and familiar to participants. A particularly apt domain for exploring the categorization of caricatured representa-

tions is faces, from which the notion of caricatures originally developed. An interesting empirical question is whether the prior results indicating caricature-based representations for interrelated categories will generalize to categories defined by faces. One reason for questioning the applicability of the previous results to faces is that faces are often thought to be processed by special-purpose mechanisms that are functionally and physiologically distinct from those involved in recognizing other objects (Bentin, Sagiv, Mecklinger, Friederici, & von Cramon, 2002; Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997). Researchers have argued that faces are typically processed holistically in terms of entire face patterns, rather than analyzed into separate components (Maurer, Le Grand, & Mondloch, 2002; Tanaka & Farah, 1993). If faces are processed holistically, one might predict that a face category would be represented by its prototypical image rather than by a caricature that selectively exaggerates certain components of the faces in the category. However, previous research has indicated better categorization for caricatures than for prototypes of faces (McKone, Martini, & Nakayama, 2001; Rhodes et al., 1987), in agreement with previous results with unfamiliar objects.

In the choice of faces as materials, several levels of categorization are possible. The categories to be learned can be defined by particular individuals; by groups of individuals based on gender, race, emotion, age, or another category; or by the superordinate category of faces (as opposed to nonfaces). The framework of isolated and interrelated concepts is general enough to be applicable at all of these levels of face categorization. The present experiments focus on categories defined by particular individuals, and the variability within a category is obtained by a gradual transformation of the identity of a face using morphing software (Steyvers, 1999). Several other researchers have created morph-based continua between faces and tested people’s ability to discriminate or categorize faces (Beale & Keil, 1995; Levin & Angelone, 2002; Levin & Beale, 2000; McKone et al., 2001). A result commonly obtained from these studies is that categorization accuracy

for a face increases monotonically as a function of the distance of the face from the category boundary (see also Ashby & Maddox, 1993). However, in these studies, the endpoints of the continua are presented as the representatives of the categories, and this, rather than these endpoints' status as caricatures, may explain why the endpoints are well categorized. In the present experiments, we explore whether a caricature advantage for categorization is still found even when prototypes are presented as often as caricatures are during category learning. We also explore factors that modulate the degree of the caricature advantage.

## EXPERIMENT 1

Numerous studies have shown that the labels given to categories of objects influence the characterization of those categories (Harnad, 1987; Malt, Sloman, Gennari, Shi, & Wang, 1999; Waxman, 1990; Wisniewski & Medin, 1994). One example that is particularly related to concept interrelatedness is the mutual exclusivity hypothesis (Markman, 1990; Waxman, Chambers, Yntema, & Gelman, 1989). According to this hypothesis, children determine the referent of a noun by assuming that nouns are mutually exclusive, and, consequently, if a new term is applied to one of two objects and one object already has a name, children will tend to assume that the term refers to the other object. Similar to Saussure's (1915/1959) notion of competition between concepts, the mutual exclusivity hypothesis assumes that as one concept gains control of a conceptual region, its competitor concepts lose control of the region. This is the same competition that is predicted to make interrelated concepts increasingly characterized by their caricatures rather than by their prototypes. In both cases, one category is displaced away from another category.

Labeling may make pairs of concepts asymmetrically dependent on one another. One concept can be labeled "Category A," whereas another concept is labeled "not Category A." In this case, the concept labeled "not Category A" is predicted to be more influenced by "Category A" than vice versa (Clark, 1990). The concept that has a label that refers to another concept is predicted to be highly influenced by the referenced concept. Even though the category structures are symmetric and the labels are randomly assigned to the two categories, the "not-Category A" concept is predicted to be characterized more in terms of a caricature than the "Category A" concept is. More specifically, there should be a tendency to associate the "not-Category A" concept with a stimulus that is more of a caricature than the stimulus associated with the "Category A" concept. There may be a bias to associate both categories with caricatures rather than with prototypes (Goldstone, 1996; Palmeri & Nosofsky, 2001; Rhodes et al., 1987), but the extent of caricaturization is predicted to be greater for the "not-Category A" concept. If we compare the relative categorization accuracies for prototypes and caricatures, we predict the advantage of caricatures over prototypes to be larger for the concept labeled "not Category A" than for the "Category A" concept.

The basis for this prediction comes from a combination of two assumptions. First, the more dependent Concept B is on a mutually exclusive Concept A, the more Concept B's characterization will be caricatured away from Concept A. This assumption is supported by previous research showing that mutually exclusive concepts that are highly interrelated tend to be characterized by features that distinguish each from the other (Goldstone, 1996; Kruschke, 1996; Niedenthal & Beike, 1997). Exaggerating distinctive properties of a concept through caricature allows a concept to be differentiated from its close conceptual neighbors. This exaggeration will be most extensive along attributes that distinguish a concept from others that are most likely to be brought to mind by the concept—that is, by other concepts that most influence the concept in question. The second assumption is that explicitly labeling Concept B as not being Concept A makes Concept B dependent on Concept A. This assumption is supported by previous work showing that, when a concept's label explicitly makes reference to another concept, then this referenced concept exerts more influence over the referencing concept's use and representation than vice versa (Goldstone, 1996; Van Wallendaal & Hastie, 1990). Putting these two assumptions together, we predict that the extent of the categorization advantage for caricatures over prototypes will be greater for a category labeled "not club members" than for a category labeled "club members."

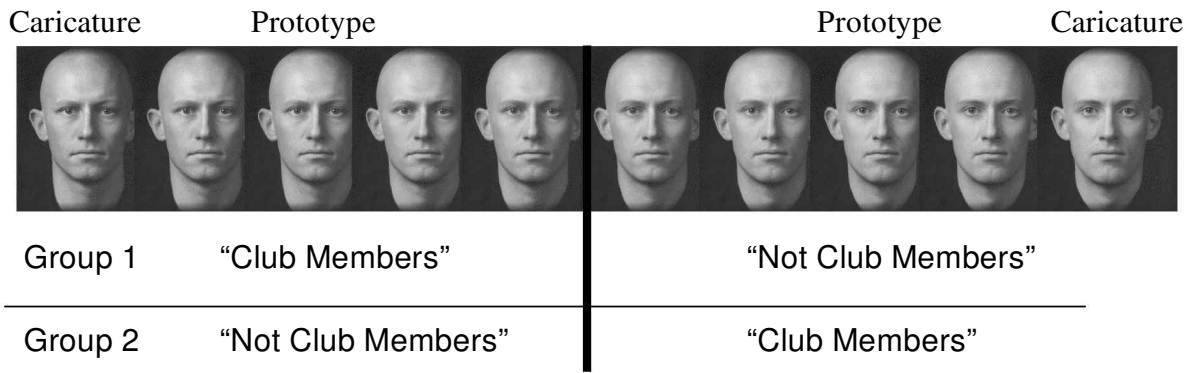
## Method

**Participants.** Sixty-two undergraduate students from Indiana University served as participants in order to fulfill a course requirement. The students were split evenly into the two labeling conditions.

**Materials.** The stimuli were faces that were generated by morphing between photographs of two bald heads selected from Kayser (1997). Previous research has suggested that morphs generated from the two selected faces did not introduce conspicuous nonlinearities between physical and psychological scalings (Goldstone & Steyvers, 2001). The morph sequence of the 10 faces used is shown in Figure 3. Each of the morphs was automatically generated using a morphing technique described by Steyvers (1999). Through 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 2 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. The faces at the left and right ends of Figure 3 are actual faces, and the 8 intermediate faces are blends of the 2 actual faces, with the proportions shifting away from those of the left face and toward those of the right face in equal 11.11% increments.

The prototype for a category was defined as the center face within the category's set of five faces. This is the face that is the most similar, on average, to the other faces within its category. The caricature of a category is defined as the face that is least like the faces from the other category. The face between the caricature and the prototype is also a face that has been systematically caricatured relative to the prototype, but to a lesser extent.

Each face was displayed in gray scale with 256 possible brightness values per pixel (1 pixel = 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 Apple Imac computer screen. The average viewing distance was 46 cm.



**Figure 3.** In Experiment 1, a morph sequence of 10 faces was divided into two categories. For half of the participants, the left five faces belonged to a category of “club members” and the remaining faces were labeled “not club members.” For the other half of the participants, these labels were reversed.

**Procedure.** On each trial, the participants saw a face and categorized it by pressing either “Y” or “N” on the keyboard, with feedback from the computer on each trial indicating with a check or an “X” whether the participant was correct or incorrect, respectively, and also indicating the correct category assignment for the face. The participants were instructed: “You will see faces appear on the screen. Half of them belong to a certain club, while the remaining half do not. If you think that a face belongs to the club, press the ‘Y’ key for ‘Yes.’ If you think that it does not belong to the club, press the ‘N’ key for No.” The dividing line between club members and non-club members is shown by the vertical line in Figure 3. For the half of the participants in Group 1, the first five faces were club members, and the last five faces were not club members. For Group 2, the first five faces were not club members, and the last five faces were club members.

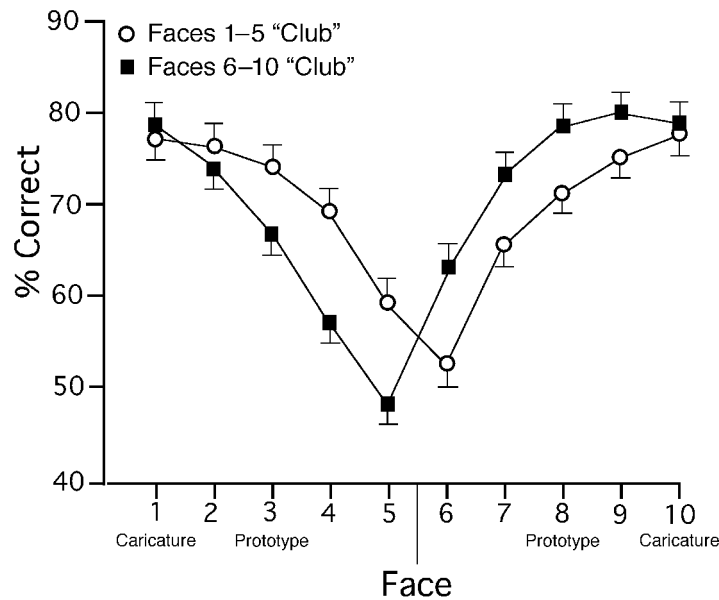
The experiment consisted of 60 repetitions of the 10 faces shown in Figure 3, for a total of 600 trials. The order of the 600 trials was randomized. The placement of a face’s center was also randomized within a  $6 \times 6$  cm square in the center of the screen. Each trial began with a face appearing on the screen. The face remained on the screen until the participant pressed the “Y” or “N” key. Immediately after one of the keys was pressed, feedback was given to the participant, and after 1.5 sec the screen was erased. Written feedback was in the form of “Yes, this face is a club member,” “No, this face is not a club member,” “Yes, this face is not a club member,” or “No, this face is a club member.” The blank interval between trials was 1 sec. The participants were given breaks every 100 trials. During these breaks, the participants were informed of their accuracy and speed during the preceding block.

## Results

The primary data of interest related to the participants’ accuracy at categorizing particular faces into the two categories. Accuracy averaged over the 600 trials was variable enough that it was a sensitive dependent measure, and was less noisy than response times. However, response times closely mirrored the accuracy data. The categorization accuracies for each face and each group of participants are shown in Figure 4. The data reveal a bias to respond “club” rather than “not club.” This is shown by the horizontal offset between the two lines. Overall, the participants responded “club” and “not club” on 54% and 46% of trials, respectively [paired  $t$  test:  $t(61) = 6.24, p < .01$ ].

The accuracy results were submitted to a repeated measures analysis of variance (ANOVA) with within-subjects factors for the stimulus presented (10 levels) and the block (2 levels, for the first and second halves of the experiment), and a between-subjects factor for the labeling condition (Faces 1–5, “Club A,” and Faces 1–5, “not Club A”). There was a main effect of stimulus presented—namely, that the stimuli closer to the category boundary were harder to categorize than the stimuli near the endpoints [ $F(9,540) = 34.59, MS_e = 4.20, p < .001$ ]. There was also a main effect of block [ $F(1,60) = 23.49, MS_e = 3.86, p < .001$ ], but not of labeling. There were interactions between stimulus presented and labeling [ $F(9,540) = 3.45, MS_e = 3.55, p < .01$ ] and between stimulus presented and block [ $F(9,540) = 2.38, MS_e = 3.18, p < .05$ ], but not between labeling and block.

To directly test hypotheses that compare categorization accuracy of caricatures and prototypes, we conducted a restricted analysis in which only Faces 1, 3, 8, and 10 were considered. Faces 1 and 10 were coded as caricatures, and Faces 3 and 8 were coded as prototypes. The labeling variable was recoded to take into account the interaction between labeling and stimulus presented. Given that the results for the 10 faces were highly symmetric across the categorization boundary, we created a new labeling variable that combined responses from Faces 1–5 with those from Faces 6–10. The labels were recoded as “negation labels” when Faces 1 and 3 were presented and Faces 1–5 were “not Club A” faces, and when Faces 8 and 10 were presented and Faces 1–5 were “Club A” faces. The labels were recoded as “standard labels” when Faces 1 and 3 were presented and Faces 1–5 were “Club A” faces, and when Faces 8 and 10 were presented and Faces 1–5 were “not Club A” faces. With this recoding, the computed ANOVA was now a repeated measures ANOVA with three within-subject factors of stimulus (prototype or caricature), labeling (standard or negation), and block (first or second block). This restricted analysis yielded a significant main effect of stimulus presented, with caricatures



**Figure 4.** Percent correct categorizations for the 10 faces of Figure 3 in Experiment 1. The categorization advantage of caricatures (Faces 1 and 10) over prototypes (Faces 3 and 8) was greater for the “not club member” category than for the “club member” category. The error bars show one standard error of the mean.

being more accurately categorized than prototypes, with respective accuracies of 78.3% and 72.6% [ $F(1,61) = 8.85, MS_e = 3.13, p < .01$ ]. There was no main effect of labeling, but there was a significant interaction between stimulus presented (prototype or caricature) and labeling [ $F(1,61) = 8.25, MS_e = 3.18, p < .01$ ]. For negation-labeled faces, caricatures and prototypes were categorized with respective accuracies of 78.6% and 68.6%. For standard-labeled faces, these respective accuracies are much more similar, at 78.0% and 76.6%. The faces between the caricature and the prototype (Faces 2 and 9) produced results intermediate to those of the caricature and the prototype for both labeling conditions.

Still with the ANOVA restricted to Faces 1, 3, 8, and 10, exploratory analyses revealed a main effect of blocks, with accuracy lower for the first 300 trials (70.7%) than for the final 300 trials (80.3%) [ $F(1,61) = 13.76, MS_e = 2.92, p < .001$ ]. However, there was also an interaction between stimulus presented and block [ $F(1,61) = 9.14, MS_e = 4.14, p < .01$ ]. For the early trials, caricatures and prototypes were categorized with respective accuracies of 72.0% and 69.4%, and for the late trials, these percentages increased to 84.6% and 76.0%. This interaction indicates that the categorization advantage for caricatures over prototypes increases with practice. This is consistent with earlier results (Goldstone, 1996) and is predicted by the hypothesis that in a categorization task in which two categories are frequently alternated, the categories will become increasingly interrelated. As categories become more interrelated, the degree of caricaturization of the category representations is expected to increase.

## Discussion

Category labeling influences the relative advantage of caricatures over prototypes even when the structure of the categories remains constant. A category that is labeled as a negation of a standard category tends to be represented as a caricature, systematically distorted away from the standard category. This is revealed by the significantly greater categorization accuracy for the negation category’s caricature than for its prototype. The standard category itself shows approximately equal categorization accuracies for prototypes and caricatures. This set of results is predicted by the distinction between interrelated and isolated concepts. One way of making Concept B dependent on Concept A is simply to make Concept B’s label depend on Concept A. This produces asymmetrically dependent concepts because B will depend on A more than A depends on B. The expected result of this is that Concept B’s representation will be caricatured away from Concept A’s representation, and that the degree of this caricaturization should be greater than the caricature of A away from B.

## EXPERIMENT 2

In Experiment 1, asymmetrically dependent concepts were created through asymmetric category labeling relations. A second way of creating asymmetric concepts is by presenting the categories sequentially. The first presented category should develop a relatively isolated category representation, because there is no other category to serve as a standard of comparison. In contrast, the second presented category can be compared with the earlier cat-

egory, and thus the second category should be more dependent on the first category than vice versa. Empirical evidence for asymmetric dependencies between sequentially learned categories comes from Kruschke (1996). He found that the second of a pair of learned categories tends to be characterized in terms of features that are diagnostic for distinguishing the second category from the first. The first learned category does not show this bias to as large an extent.

If the features that are particularly diagnostic for distinguishing a second category from a first category are emphasized in the second category's representation, then one would also expect a larger caricature-over-prototype advantage for the second than for the first category. A caricature is defined as a stimulus that emphasizes features of a category that distinguish it from other learned categories. Accordingly, we predict that the asymmetric dependency of the second category on the first category should cause the second category to be represented in a more caricatured form than the first category.

## Method

**Participants.** Sixty-six undergraduate students from Indiana University served as participants in order to fulfill a course requirement. The students were split evenly into the two category-order conditions.

**Materials and Procedure.** The stimuli were identical to those used in Experiment 1. The participants were told that they would eventually be asked to categorize faces into one of two groups. They were told that they would first see examples from one group, then examples from a second group, and, finally, they would be asked to categorize faces that they had seen before into one of these two groups. In agreement with these instructions, the experiment was broken down into three phases. During the first phase, half of the participants were presented with the left five faces from Figure 3, and the other half were presented with the right five faces. Each face

was presented a total of 15 times, and the presentation order was randomized. On each trial, the face remained on the screen for 3 sec with the label "Category A" below it, followed by a blank screen for 1 sec. In the second phase of the experiment, the faces not shown during the first phase were presented to participants with the label "Category B" below them, using the same procedure as that used in the first phase.

During the categorization phase of the experiment, the participants were shown the 10 faces from Figure 3 a total of 15 times each, in a random order. The participants were instructed to decide whether each face belonged to Category A or to Category B. The participants indicated their categorizations by pressing either the "A" or the "B" key. The participants did not receive feedback during the categorization phase.

## Results

The primary data of interest related to the participants' accuracy at categorizing particular faces into the two categories averaged over 150 trials. The categorization accuracies for each face and each group of participants are shown in Figure 5. The data reveal a small bias to categorize a face into the second, more recently seen category (Category B). Overall, the participants responded "Category A" and "Category B" on 48.7% and 51.3% of trials, respectively [paired  $t$  test:  $t(65) = 2.12, p < .05$ ].

The accuracy results were submitted to a repeated measures ANOVA with within-subjects factors for stimulus presented (10 levels) and block (2 levels, for the first and second halves of the experiment), and a between-subjects factor for order of categorizations (Faces 1–5 presented first and Faces 6–10 presented second). There was a main effect of stimulus presented [ $F(9,576) = 29.21, MS_e = 3.53, p < .001$ ]. There was also a main effect of block [ $F(1,64) = 25.06, MS_e = 3.23, p < .001$ ], but not of order of categorization. There were interactions between stimu-

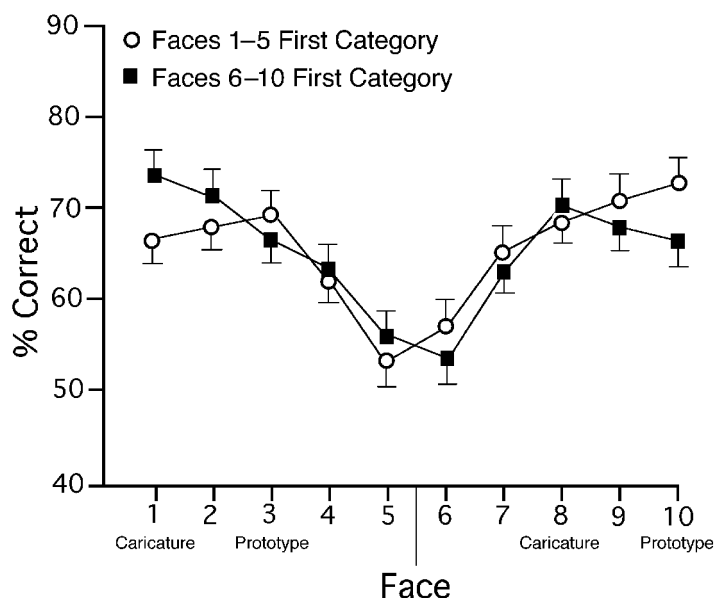


Figure 5. Percent correct categorizations in Experiment 2. The error bars show one standard error of the mean.



lus presented and order [ $F(9,576) = 4.05$ ,  $MS_e = 3.55$ ,  $p < .01$ ] and between stimulus presented and block [ $F(9,576) = 2.50$ ,  $MS_e = 4.18$ ,  $p < .05$ ], but not between order and block.

Unlike in Experiment 1, there was no main effect such that the caricatures were categorized more accurately than the prototypes. Overall, the extreme caricatures (Faces 1 and 10) and prototypes (Faces 3 and 8) were categorized with accuracies of 70.1% and 69.1%, respectively [paired  $F(1,64) = 1.48$ ,  $MS_e = 4.04$ ,  $p > .1$ ]. However, there was an interaction between face type and category order when we restricted our analysis to just the caricatures and prototypes (see Figure 4). As in Experiment 1, a recoded order variable was created that combined responses to Faces 1–5 with those to Faces 6–10. Faces 1 and 3 for the group that received Faces 1–5 first, and Faces 8 and 10 for the group that received Faces 1–5 second, were categorized together as “first category” faces. Faces 1 and 3 for the group that received Faces 1–5 second, and Faces 8 and 10 from the group that received Faces 1–5 first, were categorized together as “second category” faces. With this recoding, the computed ANOVA was now a repeated measures ANOVA with three within-subjects factors of stimulus (prototype and caricature), order (first and second categories), and block (first and second blocks). For faces from the first category, presented earlier, the caricatures and prototypes were categorized with respective accuracies of 67.1% and 70.0%. For faces from the second category, presented later, these respective accuracies reverse their rank order, with accuracies of 73.3% and 68.1% for caricatures and prototypes, respectively. These four comparisons represent a significant face type (prototype or caricature)  $\times$  order (first or second category) interaction [ $F(1,64) = 10.24$ ,  $MS_e = 3.87$ ,  $p < .01$ ]. The faces between the caricature and the prototype (Faces 2 and 9) produced results intermediate to those of the caricature and the prototype for both labeling conditions.

## Discussion

According to the results from Experiment 2, whether a caricature advantage or a prototype advantage is found during categorization depends on whether the category is first or second in a pair of sequentially presented categories. For the first category presented, the prototype is better categorized than the caricature. The opposite effect is found for the second category. This is consistent with the hypothesis that the second category is more likely to be characterized relative to the first category than vice versa. Isolated categories are expected to be represented in terms of their prototypes, whereas a category that depends on a second, mutually exclusive category is expected to be represented in terms of a caricature that is systematically distorted away from the second category.

In Experiment 1, a general advantage for caricatures over prototypes was found, whereas in Experiment 2, there was no significant main effect of face type. Given procedural differences between the experiments, this is naturally accounted for by the framework of isolated and interrelated concepts. In Experiment 1, items taken from the two cat-

egories were randomly intermixed, whereas in Experiment 2, the two categories were presented separately. Categories are predicted to be more interrelated when they are presented simultaneously, because there is more opportunity for interactions between developing category representations. This prediction is supported by previous evidence that, when simultaneously acquired categories are alternated frequently, the category representations are more interrelated than when categories are alternated rarely (Goldstone, 1996). Dividing categories into completely separate blocks, as was done in Experiment 2, can be interpreted as an extreme way of alternating categories rarely.

In Experiment 1, there was a bias to respond with the standard-labeled rather than the negation-labeled category, and the negation-labeled category was more caricatured than the standard-labeled category. In Experiment 2, there was a bias to respond with the label of the category learned second rather than with that of the the first category, but the second category was more caricatured than the first. Accordingly, across the two experiments, response bias is dissociable from caricaturization. This is useful because it shows that the amount of caricaturization of a category is not simply a function of how often the category is given as a response to a stimulus.

## GENERAL DISCUSSION

The effects of the two manipulations of Experiments 1 and 2 are well integrated by the proposal that concepts differ in how interrelated they are. In particular, two concepts may be asymmetrically dependent on one another, such that Concept B depends more on Concept A than vice versa. From this perspective, manipulations that might not otherwise be seen as similar are seen as cohering together. Labeling a category “not Category A” and presenting examples of the category after presenting Category A exemplars are two ways of making the category highly dependent on Category A. Given this dependence and the mutual exclusivity between the categories (items belong to one and only one of the categories), the representation of the category is expected to be caricatured away from the prototype of Category A. Category A, by contrast, is expected to be more independent in these experimental conditions, and its representation is expected to be more prototypical and less caricatured.

The present experiments extend theories of isolated and interrelated concepts to face stimuli. Previous experiments have explored interrelated concepts using abstract and unfamiliar shape categories (Goldstone, 1996) or verbal stimuli (Niedenthal & Beike, 1997). Although there may be specialized perceptual processes for faces (McCarthy et al., 1997; but also see Gauthier & Tarr, 2002), categorization of faces also shows similarities with other domains. In general, a categorization advantage for caricatures over prototypes is frequently observed for faces and other objects; however, the extent of this advantage can be manipulated by experimental variables. The previous demonstrations of malleable categorization are apparently not simply artifacts of the arbitrary and unfamil-

iar stimuli used. In addition, the present results suggest that previous arguments for the superior categorization of caricatures over prototypes (Rhodes et al., 1987) may need to be qualified by experimental conditions. The caricature advantage can be accentuated by giving a category a label that refers to another category, and can be eliminated by presenting the category members before other alternative categories to be learned. Rather than trying to obtain definitive evidence as to whether faces are represented by caricatures, prototypes, or images close to category boundaries, a more fruitful empirical mission is to determine the experimental and observer characteristics that affect the choice of face representation.

### Caricaturization and the Perception of Others

The present results are novel in that the experimental manipulations of category interrelatedness had not been tested with rich, naturalistic stimuli. Goldstone (1996) predicted that category presentation order should affect concept interrelatedness, but this notion had not been directly tested until now. However, there are several results in social psychology that are related to the present demonstration. The basis for this link is that an individual is likely to think of him- or herself in a relatively isolated, independent manner, and to think of others in relation to him- or herself (Niedenthal & Beike, 1997). This principle can also be applied at the group level, with the consequence that a person is likely to think of his or her own group in a relatively isolated manner, and to think of other groups in relation to his or her own (Linville & Jones, 1980). In combination with the present framework, this suggests that people should perceive other people and other groups as more caricatured than they really are, but that they should perceive themselves and their own groups in a more veridical fashion.

In agreement with this hypothesis, Linville and Jones (1980) found that members of other groups are more likely to be appraised in an extreme manner than members of one's own group. For example, white participants judged a black applicant more favorably than they did a white applicant when both applicants had relatively strong credentials, but judged the black applicant more negatively when both applicants had weak credentials. A mechanism that may underlie this effect is that distinctive attributes of a relatively unfamiliar or rare group have an exaggerated effect on people's judgments (McConnell, Sherman, & Hamilton, 1994). Caricatures are exactly the kind of representation that exaggerates features of one group relative to those of another group. The relatively unfamiliar outgroup is particularly prone to being represented in a caricatured manner.

Social cognition researchers have also found a general bias by which people accentuate the differences between categories (Corneille & Judd, 1999; Krueger & Rothbart, 1990). As in McConnell et al. (1994), the mechanism for this contrast effect appears to be that the importance of properties that are particularly diagnostic for distinguishing between categories is exaggerated. The present results extend these findings by showing that these exaggerations can be asymmetric. An independent concept is less sus-

ceptible to exaggeration than a relationally defined concept. Under the assumption that white participants living in white-majority environments view the category *whites* as relatively isolated and view blacks in terms of their differences from whites, the prediction is that these participants would tend to represent *whites* by the prototype of white people, but *blacks* by a caricature distorted away from the prototype for black people in the direction opposite of *whites*. Although this kind of asymmetry has not been directly tested to our knowledge, greater exaggeration of group differences has been found when participants assessed individuals from minority rather than majority groups (see Hewstone, Rubin, & Willis, 2002, for a review).

### Models of Isolated and Interrelated Concepts

Existing models of categorization differ in their predictions of the relative accuracy for categorizing prototypes and caricatures. Decision boundary models of categorization (Ashby & Maddox, 1993) predict that categorizations become easier as stimuli move away from the boundaries between categories. A natural consequence of these models is that caricatures should be easier to categorize than prototypes in a two-category situation because caricatures are designed to be displaced away from the boundary between the categories. Exemplar models represent categories in terms of the members belonging to each category, and categorize new items by measuring their summed similarity to the existing members of the candidate categories (Medin & Schaffer, 1978). In one influential exemplar model—the generalized context model (GCM; Nosofsky, 1986, 1991)—the likelihood of correctly categorizing a caricature and a prototype of a category depends on whether similarity is assumed to be based on an exponential or on a Gaussian function of stimulus distance (Nosofsky, 1991, p. 135). Under an exponential relation between subjective similarity and distance, and given only one single-dimension item in each of two categories, caricatures and prototypes are predicted to be exactly equally well categorized. Under a Gaussian relation, categorization accuracy increases monotonically as the extremity of caricature increases. Thus, neither of these functions allows GCM to predict either a categorization advantage for a prototype relative to a caricature, or greatest categorization accuracy for an intermediate degree of caricaturization. By contrast, other models have predicted that intermediate stimuli between prototypes and extreme caricatures will be most accurately categorized. These models are useful in accommodating the “peak shift” phenomenon, according to which the maximal response is given to a stimulus that is displaced away from a reinforced stimulus by a specific amount in the direction opposite an unreinforced stimulus (McLaren, Bennett, Guttman-Nahir, & Kim, 1995; Spence, 1936).

The present data challenge any model of categorization that predicts a constant advantage of caricatures over prototypes or vice versa. We found that whether a caricature-over-prototype advantage was found depends on whether the concepts involved were relatively independent or inter-

related. This kind of interaction is accommodated by Goldstone's (1996) RECON model. This connectionist model is a two-layer recurrent network. One layer of units represents the input dimensions, and the other represents the learned categories. All units in RECON are connected to each other by weighted links. Unlike standard feed forward networks, RECON has recurrent connections between category units. These connections provide a mechanism through which categories influence each other. By varying a single parameter—the degree of influence of category units on each other—varying degrees of concept interrelatedness are obtained. Simulations have shown that as the magnitude of category-to-category weights increases, the categorization advantage of caricatures over prototypes increases. Strong lateral inhibition between categories leads to the categories' emphasizing distinctive features that uniquely characterize just one category.

RECON can also accommodate asymmetrically interrelated categories by removing the constraint that the weight from Category A to Category B must equal the weight from B to A. If the weight from A to B is greater than the weight from B to A, then the representation of Category B will be caricatured (away from A) relative to Category A's representation. In this manner, the present experiments can be accommodated by RECON, under the assumption that labeling B so as to refer to A, or presenting B after A, will make the weight from Category A to Category B greater than that from B to A.

More generally, the present results are also consistent with models that assume that the representation of a concept is based not only on its connections to the external, perceptual world (Barsalou, 1999; Goldstone & Barsalou, 1998), but also on its connections to other concepts (Goldstone & Rogosky, 2002). Concepts do not typically act like independent detectors polling the world. Concepts are also influenced by the simultaneous presence of other concepts. Understanding the interaction between external and internal determinants of conceptual meaning is an important, but difficult, project. The present experiments take one step toward this goal by developing empirical tools for manipulating and measuring the extent to which concepts depend on each other and on perceptual information. If there are indeed individual differences among concepts in terms of how independent or interrelated they are, then studying these individual differences can shed light on how internal and external sources of meaning are represented.

## REFERENCES

- ASHBY, F. G., & MADDOX, W. T. (1993). Relations among prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*, **38**, 423-466.
- BARR, R. A., & CAPLAN, L. J. (1987). Category representations and their implications for category structure. *Memory & Cognition*, **15**, 397-418.
- BARSALOU, L. W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences*, **22**, 577-660.
- BEALE, J. M., & KEIL, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, **57**, 217-239.
- BENTIN, S., SAGIV, N., MECKLINGER, A., FRIEDERICI, A., & VON CRAMON, Y. D. (2002). Priming visual face-processing mechanisms: Electrophysiological evidence. *Psychological Science*, **13**, 190-193.
- BLOCK, N. (1999). Functional role semantics. In R. A. Wilson & F. C. Keil (Eds.), *MIT encyclopedia of the cognitive sciences* (pp. 331-332). Cambridge, MA: MIT Press.
- BURGESS, C., LIVESAY, K., & LUND, K. (1998). Explorations in context space: Words, sentences, and discourse. *Discourse Processes*, **25**, 211-257.
- BURGESS, C., & LUND, K. (2000). The dynamics of meaning in memory. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics: Conceptual change in humans and machines* (pp. 117-156). Mahwah, NJ: Erlbaum.
- BUSEY, T. A. (1998). Physical and psychological representations of faces: Evidence from morphing. *Psychological Science*, **9**, 476-483.
- CLARK, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language*, **17**, 417-431.
- COLLINS, A. M., & QUILLIAN, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior*, **8**, 240-247.
- CORNEILLE, O., & JUDD, C. M. (1999). Accentuation and sensitization effects in the categorization of multifaceted stimuli. *Journal of Personality & Social Psychology*, **77**, 927-941.
- EDELMAN, S. (1999). *Representation and recognition in vision*. Cambridge, MA: MIT Press.
- FIELD, H. (1977). Logic, meaning, and conceptual role. *Journal of Philosophy*, **74**, 379-409.
- GALTON, F. J. (1878). Composite portraits. *Nature*, **18**, 97-100.
- GAUTHIER, I., & TARR, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 431-446.
- GOLDSTONE, R. L. (1996). Isolated and interrelated concepts. *Memory & Cognition*, **24**, 608-628.
- GOLDSTONE, R. L., & BARSALOU, L. (1998). Reuniting perception and conception. *Cognition*, **65**, 231-262.
- GOLDSTONE, R. L., & ROGOSKY, B. J. (2002). Using relations within conceptual systems to translate across conceptual systems. *Cognition*, **84**, 295-320.
- GOLDSTONE, R. L., & STEVYERS, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General*, **130**, 116-139.
- HARNAD, S. (Ed.) (1987). *Categorical perception*. Cambridge: Cambridge University Press.
- HARNAD, S. (1990). The symbol grounding problem. *Physica D*, **42**, 335-346.
- HEWSTONE, M., RUBIN, M., & WILLIS, H. (2002). Intergroup bias. *Annual Review of Psychology*, **53**, 575-604.
- HUME, D. (1973). *A treatise of human nature: Being an attempt to introduce the experimental method of reasoning into moral subjects*. London: Thomas Longman. (Original work published 1740)
- KANWISHER, N., MCDERMOTT, J., & CHUN, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, **17**, 4302-4311.
- KAYSER, A. (1997). *Heads*. New York: Abbeville Press.
- KRUEGER, J., & ROTHBART, M. (1990). Contrast and accentuation effects in category learning. *Journal of Personality & Social Psychology*, **59**, 651-663.
- KRUSCHKE, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 3-26.
- LAKOFF, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- LANDAUER, T. K., & DUMAIS, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, **104**, 211-240.
- LENAT, D. B., & FEIGENBAUM, E. A. (1991). On the thresholds of knowledge. *Artificial Intelligence*, **47**, 185-250.
- LEVIN, D. T., & ANGELONE, B. L. (2002). Categorical perception of race. *Perception*, **31**, 567-578.
- LEVIN, D. T., & BEALE, J. M. (2000). Categorical perception occurs in newly learned faces, other-race faces, and inverted faces. *Perception & Psychophysics*, **62**, 386-401.
- LINVILLE, P. W., & JONES, E. E. (1980). Polarized appraisals of out-group members. *Journal of Personality & Social Psychology*, **38**, 689-703.

- MALT, B. C., SLOMAN, S. A., GENNARI, S., SHI, M., & WANG, Y. (1999). Knowing versus naming: Similarity and the linguistic classification of artifacts. *Journal of Memory & Language*, **40**, 230-262.
- MARKMAN, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, **14**, 57-77.
- MAURER, D., LE GRAND, R., & MONDLOCH, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, **6**, 255-260.
- MCCARTHY, G., PUCE, A., GORE, J. C., & ALLISON, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, **9**, 605-610.
- MCCONNELL, A. R., SHERMAN, S. J., & HAMILTON, D. L. (1994). Illusory correlation in the perception of groups: An extension of the distinctiveness-based account. *Journal of Personality & Social Psychology*, **67**, 414-429.
- McKENZIE, C. R. M. (1998). Taking into account the strength of an alternative hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 771-792.
- McKENZIE, C. R. M. (1999). (Non)Complementary updating of belief in two hypotheses. *Memory & Cognition*, **27**, 152-165.
- McKONE, E., MARTINI, P., & NAKAYAMA, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 573-599.
- MCLAREN, I. P. L., BENNETT, C. H., GUTTMAN-NAHIR, T., & KIM, K. (1995). Prototype effects and peak shift in categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 662-673.
- MEDIN, D. L., & SCHAFFER, M. M. (1978). A context theory of classification learning. *Psychological Review*, **85**, 207-238.
- NIEDENTHAL, P. M., & BEIKE, D. R. (1997). Interrelated and isolated self-concepts. *Personality & Social Psychology Review*, **1**, 106-128.
- NOSOFSKY, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, **115**, 39-57.
- NOSOFSKY, R. M. (1991). Typicality in logically defined categories: Exemplar-similarity versus rule instantiation. *Memory & Cognition*, **19**, 131-150.
- PALMERI, T. J., & NOSOFSKY, R. M. (2001). Central tendencies, extreme points, and prototype enhancement effects in ill-defined perceptual categorization. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **54A**, 197-235.
- POSNER, M. I., & KEELE, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, **77**, 353-363.
- QUILLIAN, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science*, **12**, 410-430.
- REED, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, **3**, 382-407.
- RHODES, G., BRENNAN, S., & CAREY, S. (1987). Identification and ratings of caricatures: Implications for mental representations of faces. *Cognitive Psychology*, **19**, 473-497.
- ROSCH, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: Human Perception & Performance*, **1**, 303-322.
- SAUSSURE, F. DE (1959). *Course in general linguistics*. New York: McGraw-Hill. (Original work published 1915)
- SOLOMON, K. O., & BARSALOU, L. W. (2001). Representing properties locally. *Cognitive Psychology*, **43**, 129-169.
- SPENCE, K. W. (1936). The nature of discrimination learning in animals. *Psychological Review*, **43**, 427-429.
- STEYVERS, M. (1999). Morphing techniques for manipulating face images. *Behavior Research Methods, Instruments, & Computers*, **31**, 359-369.
- STICH, S. P. (1983). *From folk psychology to cognitive science: The case against belief*. Cambridge, MA: MIT Press.
- TANAKA, J. W., & FARAH, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, **46A**, 225-245.
- VAN WALLENDIAEL, L. R., & HASTIE, R. (1990). Tracing the footsteps of Sherlock Holmes: Cognitive representations of hypothesis testing. *Memory & Cognition*, **18**, 240-250.
- WAXMAN, S. R. (1990). Linguistic biases and the establishment of conceptual hierarchies: Evidence from preschool children. *Cognitive Development*, **5**, 123-150.
- WAXMAN, S. R., CHAMBERS, D. W., YNTEMA, D. B., & GELMAN, R. (1989). Complementary versus contrastive classification in preschool children. *Journal of Experimental Child Psychology*, **48**, 410-422.
- WISNIEWSKI, E. J., & MEDIN, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science*, **18**, 221-281.

(Manuscript received March 20, 2002;  
revision accepted for publication September 5, 2002.)