Central tendencies, extreme points, and prototype enhancement effects in ill-defined perceptual categorization

Thomas J. Palmeri
Vanderbilt University, Nashville, U.S.A.
Robert M. Nosofsky
Indiana University, Bloomington, U.S.A.

In three perceptual classification experiments involving ill-defined category structures, extreme prototype enhancement effects were observed in which prototypes were classified more accurately than other category instances. Such empirical findings can prove theoretically challenging to exemplar-based models of categorization if prototypes are psychological central tendencies of category instances. We found instead that category prototypes were sometimes better characterized as psychological extreme points relative to contrast categories. Extending a classic and widely cited study (Posner & Keele, 1968), participants learned categories created from distortions of dot patterns arranged in familiar shapes. Participants then made pairwise similarity judgements of the patterns. Multidimensional scaling (MDS) analyses of the similarity data revealed the prototypes to be psychological extreme points, not central tendencies. Evidence for extreme point representations was also found for novel prototype patterns displaying a symmetry structure and for prototypes of grid patterns used in recent studies by McLaren and colleagues (McLaren, Bennet, Guttman-Nahir, Kim, & Mackintosh, 1995). When used in combination with the derived MDS solutions, an exemplar-based model of categorization, the Generalized Context Model (Nosofsky, 1986), provided good fits to the observed categorization data in all three experiments.

One of the major research paradigms used for investigating perceptual categorization has been the dot-pattern prototype-distortion task introduced by Posner and Keele (1968, 1970). In this task, prototype dot patterns are created, and participants are trained to

Requests for reprints should be sent to Thomas J. Palmeri, Department of Psychology, 301 Wilson Hall, Vanderbilt University, Nashville, TN 37240, U.S.A. Email: Thomas.J.Palmeri@Vanderbilt.edu

This work was partially supported by Grant PHS R01 MH48494 from the National Institute of Mental Health to Robert M. Nosofsky and by Vanderbilt University Research Council Direct Research Support Grants to Thomas J. Palmeri. We thank Carolyn Cave, I.P.L. McLaren, Timothy McNamara, and an anonymous reviewer for their comments on earlier versions of this work. Thanks to Jason Bechler, Debra Coomes, Andrew Domescik, Heather Fahey, Brenda Graves, Kristen Hicks, Stephanie Jurek, Nicole Naylor, Thomas Sharpe, and Stephanie Tang for testing participants.

http://www.tandf.co.uk/journals/pp/02724987.html DOI:10.1080/02724980042000084

categorize statistical distortions of these prototypes. During transfer, old distortions, new distortions, and the prototypes are presented to be classified without feedback. As discussed, for example, by Homa (1984), a major advantage of these experiments is that the dot patterns are essentially infinitely variable and have a highly complex dimensional structure, so that the properties of the artificial categories that are created may mimic those of many natural categories.

The dot-pattern paradigm has been used to systematically investigate the effects of numerous fundamental variables on category learning and transfer, including effects of category size (e.g., Breen & Schvaneveldt, 1986; Homa & Vosburgh, 1976; Posner & Keele, 1968; Shin & Nosofsky, 1992), category variability (e.g., Barresi, Robbins, & Shain, 1975; Homa, 1978; Homa & Vosburgh, 1976; Posner & Keele, 1968), instance frequency (e.g., Homa, Dunbar, & Nohre, 1991; Shin & Nosofsky, 1992), number of categories learned (e.g., Homa & Chambliss, 1975), amount of training (e.g., Homa et al., 1991; Homa, Goldhardt, Burruel-Homa, & Smith, 1993), and delay between training and transfer (e.g., Homa, Cross, Cornell, Goldman, & Schwartz, 1973; Posner & Keele, 1970; Strange, Kenney, Kessel, & Jenkins, 1970). In addition, the paradigm has been used to investigate the relationship between categorization and old–new recognition memory (e.g., Homa et al., 1993; Metcalfe & Fisher, 1986; Onohundro, 1981; Shin & Nosofsky, 1992; Vandierendonck, 1984) and has also served as a fundamental testing ground for investigating neuropsychological aspects of categorization (e.g., Knowlton & Squire, 1993; Kolodny, 1994; Nosofsky & Zaki, 1998; Palmeri & Flanery, 1999; Reber & Squire, 1997; Reber, Stark, & Squire, 1998; Squire & Knowlton, 1995). Indeed, the dot-pattern paradigm has provided bedrock data for evaluating numerous theories of categorization and memory including prototype models (e.g., Busemeyer, Dewey, & Medin, 1984; Homa, Sterling, & Trepel, 1981), distributed memory models (e.g., Knapp & Anderson, 1984; Metcalfe, 1982), connectionist models (e.g., McClelland & Rumelhart, 1985), and exemplar models (Hintzman, 1986; Nosofsky, 1988; Shin & Nosofsky, 1992).

One of the most salient aspects of dot-pattern studies is a well-known effect called *prototype enhancement*. On average, category prototypes that are not experienced during training are typically classified during transfer as well as, and sometimes somewhat better than, the old category instances, and better than new category instances. Although such prototype enhancement effects were originally believed to provide solid evidence for the existence of prototype abstraction processes, theoretical work has shown that pure exemplar retrieval models can account for this phenomenon as well (e.g., Busemeyer et al., 1984; Hintzman, 1986; Hintzman & Ludlam, 1980; Nosofsky, 1988, 1992; Shin & Nosofsky, 1992). It may seem paradoxical that models that assume that categories are represented solely in terms of stored exemplars can account for enhanced classification of unseen prototypes. The key intuition is that, although any given old exemplar is highly similar to itself, it may not be very similar to any other old exemplars. By contrast, prototypes are typically similar to many other exemplars stored in memory. The similarity of prototypes to numerous stored exemplars makes up for the lack of stored representations for the prototypes themselves.

Shin and Nosofsky (1992) demonstrated an approach to modelling detailed aspects of dot-pattern classification performance that combined an exemplar-based model of categorization, the Generalized Context Model (GCM; Nosofsky, 1986), with multidimensional scaling (MDS) techniques (Shepard, 1980). The stimuli they used were typical of most random-dot-pattern studies. Random prototypes were generated for each category, and statistical distortions of these prototypes were created as category instances (Posner, Goldsmith, & Welton, 1967). Some of these distortions were designated as training patterns and some were designated as transfer patterns. In their studies, all participants learned to classify the same set of patterns. Over numerous trials, participants learned to classify the training patterns with corrective feedback. During transfer, they classified the training patterns, transfer patterns, and category prototypes without feedback. Across three experiments, Shin and Nosofsky examined effects of several fundamental learning variables on categorization, including level of distortion of patterns, category size, delay of the transfer phase, and individual item frequency. Their primary goal was to assess whether a pure exemplar-based model could account for the observed classification results of whether prototype abstraction processes needed to be assumed as well.

Several previous attempts to model dot-pattern classification have used randomly generated multi-element stimulus vectors as inputs to simulation models (e.g., Hintzman, 1986; Knapp & Anderson, 1984; Metcalfe, 1982; Nosofsky, 1988). Although such representations are intended to capture some elements of the physical instantiation of the dot-pattern stimuli, they may fail to capture the true psychological relationships among these complex patterns. Moreover, these methods allow only gross-level predictions to be made, such as predictions of average classification performance for the prototypes and old and new distortions. Instead, as a more detailed test of various competing models, Shin and Nosofsky (1992) aimed to account for the classification performance of particular instances, not just average classification of particular types of stimuli. In their experiments, participants provided pairwise similarity ratings of the dot patterns. These similarity rating data were then analysed using standard MDS techniques to obtain a psychological scaling solution for the stimuli. The derived scaling solution was used in conjunction with the GCM, a prototype model, and a mixed model to account for the observed classification data. Theoretical analyses revealed little evidence for the existence of a prototype abstraction process that operated above and beyond pure exemplar-based generalization. Among other qualitative and quantitative predictions, the pure exemplar-based model could account for the prototype enhancement effects that Shin and Nosofsky observed.

It is important to note that many reported cases of prototype enhancement in experiments using the dot-pattern paradigm have compared classification accuracy for category prototypes relative to the *average* classification accuracy for other category instances; in those articles that report classification probabilities for individual stimuli, the prototype is *not* the best classified item overall. For example, the degree of prototype enhancement reported by Shin and Nosofsky (1992) was not very large, and many old category instances were classified more accurately than the category prototypes. A potential challenge for the GCM and other exemplar models is whether or not they could ever predict an *extreme prototype enhancement* effect in this paradigm, in which the prototypes are classified significantly more accurately than other category instances.

An observation of extreme prototype enhancement could provide a serious challenge to the GCM and other exemplar models. According to many theories of perceptual categorization, including the GCM, objects are represented as points in a multidimensional psychological space (e.g., Ashby, 1992; Homa, 1984; Nosofsky, 1986; Reed, 1972; Shepard & Chang, 1963). It is natural to assume that category prototypes in dot-pattern studies, as well as other experimental paradigms, are psychological as well as physical central tendencies of category instances (Homa, 1984; Homa et al., 1981; Nosofsky, 1987; Posner, 1969; Reed, 1972; Rosch, 1975b, 1978; Smith & Medin, 1981). For example, in summarizing the results of Posner and Keele's (1968, 1970) studies, Anderson (1980, p. 140, 142) wrote: "One of the most impressive demonstrations of subjects' ability to extract the central tendency of a set of instances is a series of experiments performed by Posner and Keele (1968, 1970). . . . The prototype for Posner's dot patterns would have been the average of the studied dot patterns." This assumption about a central-tendency representation for the prototypes has been largely confirmed in multidimensional scaling studies involving randomly generated prototypes (Homa, 1984; Homa, Rhoads, & Chambliss, 1979; Shin & Nosofsky, 1992). Simplifying the multidimensional representation of the patterns somewhat, the left panel of Figure 1 depicts category prototypes (cA, cB, and cC) as psychological central tendencies of the category instances. A key point is that if the category prototypes are indeed central tendencies in psychological space as well as being physical central tendencies of category instances, then the GCM *cannot* predict an extreme prototype enhancement effect. Instead, category instances lying in extreme regions of the psychological space (relative to the contrast categories) will tend to be classified more accurately than the central prototypes, as is illustrated later.

An alternative possibility, however, to be explored in the present research, is that prototypes that are physical central tendencies of category instances may sometimes reside not as psychological central tendencies, but rather as psychological extreme points relative to the category instances. The right panel of Figure 1 depicts a situation in which physical category central tendencies may be represented as extreme points (eA, eB, and eC) in the psychological space relative to the category instances. Under such conditions, the GCM does predict extreme prototype enhancement, as illustrated next.[1]

To illustrate how predictions of the GCM can vary depending on whether the prototypes are central tendencies or extreme points, we generated predictions of the model based on the idealized configuration shown in the left panel of Figure 2. We compared predicted classification accuracy for "close" exemplars, which are close to members of contrast categories, "far" exemplars, which are far from members of contrast categories, central-tendency prototypes, and extreme-point prototypes. The right panel of Figure 2 shows predicted classification accuracy for each of these types of stimulus across a range of parameter values (this theoretical analysis allowed the sensitivity parameter, c, which scales distances between instances in psychological space, to vary across a range of values; we also assumed equal

[1] Extreme prototype enhancement has often been observed in experimental paradigms involving discrete-dimension stimuli. However, in experimental paradigms involving discrete dimensions, such as those initially tested by Medin and Schaffer (1978), the interpretation of a prototype as a "central tendency" versus an "extreme point" is unclear. Consider stimuli varying along four binary-valued dimensions. Suppose the prototypes of Categories A and B are 1111 and 2222, respectively, and that category instances are generated by distorting these prototypes by varying degrees (e.g., 1112 and 1121 might be examples of Category A, and 2221 and 2212 might be examples of Category B). From one point of view, these prototypes can be viewed as "central tendencies", in the sense that they have the modal values on each dimension. However, these prototypes act as extreme points in the multidimensional category structure as well. Therefore, one needs to test experimental paradigms with complex continuous-dimension stimuli, such as dot patterns, to sharply distinguish between the roles of central tendencies and extreme point representations in classification.

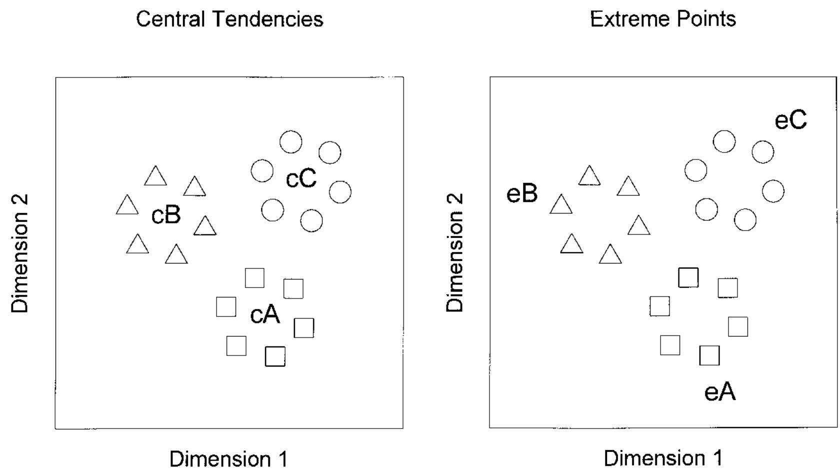

Figure 1. A schematic illustration of central tendencies and extreme points in simplified two-dimensional psychological space. In both panels, three categories with six distortions each are depicted by the squares, circles, and triangles. In the left panel, the prototypes (cA, cB, and cC) are central tendencies of the six distortions of their category. In the right panel, the prototypes (eA, eB, and eC) are extremes relative to the six distortions of their category and relative to the distortions of the other categories.

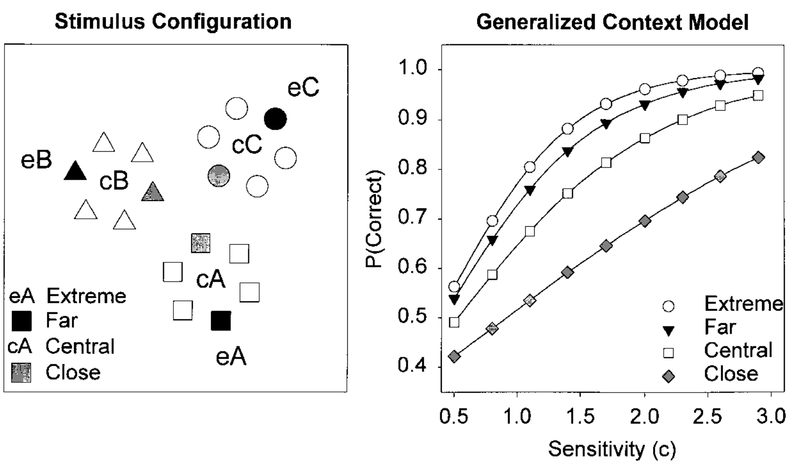

Figure 2. The left panel displays a schematic illustration of the psychological space used in the simulations reported in the text. Grey symbols indicate close exemplars, black symbols indicate far exemplars, cA indicates a central-tendency prototype, eA indicates an extreme-point prototype. The right panel displays the GCM predictions for each of these four types of stimulus as a function of the sensitivity parameter of the model, c.

response biases for the three categories and assumed equal attention to both psychological dimensions; see the discussion of the GCM following Experiment 1 for details of the model). As shown in the figure, the close exemplars were predicted to be classified with relatively low accuracy, and the far exemplars were predicted to be classified with relatively high accuracy. The GCM could not predict prototypes to be the best classified items when central-tendency representations were assumed—prototypes were always classified with intermediate accuracy. In these simulations, the failure to predict extreme prototype enhancement with central tendency representations was observed regardless of whether the prototypes were allowed to be old training items (a point that will be important when reviewing the results of the first two experiments). However, the GCM could predict the prototypes to be the best classified items when extreme-point representations were assumed. It is clear that the derived MDS solution is crucial to ascertain whether the GCM can or cannot predict an observed pattern of extreme prototype enhancement. Certain "objective" measures of pattern similarity, defined by such things as the average distances between dots in pairs of patterns (e.g., Posner, 1969), are probably insufficient—using such measures of similarity, the prototype, being a physical central tendency of category instances, would always emerge as a psychological central tendency as well.

The potential importance of psychological extremes in categorization has been noted in some other work. For example, Barsalou (1985, 1991) demonstrated the importance of ideal points in highly conceptual domains involving goal-derived categories—the best example of the category "foods to eat on a diet" is one with zero calories, not one with the average caloric content of typical diet foods. In applications of their highly successful *Fuzzy-Logical Model of Perception* (FLMP), Massaro and colleagues often tested paradigms in which the stimuli varied along two clear continuous dimensions, and the prototypes to which people compared objects were assumed to occupy extreme corners of the psychological space (Massaro, 1987; Massaro & Friedman, 1990; Oden & Massaro, 1978)—note that in their paradigms, the physically manipulated prototypes and their resulting psychological representations were *both* extreme points. Although the potential importance of psychological extremes in categorization has been suggested by the work of Barsalou, Massaro, and others (e.g., Goldstone, 1993, 1996; Rosch, 1975b), the idea that a prototype that is a central tendency in the physically defined space may emerge as an extreme point in the psychological space has not previously been suggested. Although prototypes may indeed be physical central tendencies of the distortions created from them, it does not necessarily follow that they are psychological central tendencies as well. Rather, various emergent dimensions, based on diagnostic configurations among elements of a complex physical stimulus such as a dot pattern (e.g., Hock, Tromley, & Polmann, 1988), may be formed, which cause the prototypes to be represented as psychological extremes within the context of learning particular categories.

One goal of the present research was to document that extreme prototype enhancement effects could be observed empirically. Whereas the prototypes that most recent experiments have used were random dot patterns, in one of the original Posner and Keele (1968) studies, highly recognizable dot patterns (e.g., a triangle, an M, and an F) were used as category prototypes instead. We chose to use such recognizable patterns in the first experiment, reasoning that their use as prototypes might offer an excellent chance of empirically observing an extreme prototype enhancement effect. Another goal of the present research was to examine the nature of the psychological representations of the physical category prototypes: Are they best characterized as psychological central tendencies or as psychological extreme points? To assess this issue, participants provided pairwise similarity ratings among all patterns in the set, and multidimensional scaling techniques were used to derive a psychological space for those patterns. If extreme prototype enhancement is observed and physical prototypes are represented as psychological central tendencies, then the GCM and many other exemplar models would be falsified. Finally, even if the prototypes are represented as extreme points in the MDS solution, the question still arises as to what type of decision model will provide the best account of the detailed classification performance data. We assessed this question by comparing the quantitative fits of formal exemplar, prototype, and mixed exemplar-plus-prototype models.

EXPERIMENT 1

This experiment was an extension of Posner and Keele's (1968) classic study. Following their design, the categories were based on prototype patterns formed in the shape of a triangle, a plus, and an F.[2] Category instances were distortions of those prototype patterns. During training, participants learned to classify these instances with feedback. To increase the likelihood of observing an extreme prototype enhancement effect, for one group of participants the category prototypes were also presented during training (recall from the simulations reported earlier that the inability of the GCM to account for extreme prototype enhancement with central tendency representations was not modulated by the presence or absence of prototypes during training). At transfer, participants were tested on old distortions, new distortions, and prototypes. Extending the Posner and Keele design, participants also made pairwise similarity judgements; multidimensional scaling techniques were used to analyse the similarity data to derive the psychological coordinates for the patterns (Shin & Nosofsky, 1992). The scaling solution should reveal whether the prototypes were central tendencies or extreme points in the psychological space. Exemplar, prototype, and exemplar-plus-prototype models were fitted to the observed categorization data to test whether a pure exemplar-based model needed to be supplemented by special prototype abstraction mechanisms.

Method

Participants

Participants were 280 undergraduates who received course credit. All were tested individually.

Stimuli

Patterns were composed of nine dots placed on a 50×50 grid. As shown in Figure 3, the three prototype patterns were in the shape of a triangle, a plus, and an F. These prototypes fitted within the centre 30×30 of the grid. From each prototype, nine moderate-level distortions (6 bits/dot) were created by using a standard statistical distortion algorithm (Posner et al., 1967); this algorithm moves each dot of the prototype pattern some small amount in a random direction. Six distortions were selected as old training items, and three were selected as new transfer items. Stimuli were presented on 14-in computer monitors.

[2] Unlike Posner and Keele (1968), we used a *plus* instead of an *M* so that the prototype would belong to different superordinate categories.

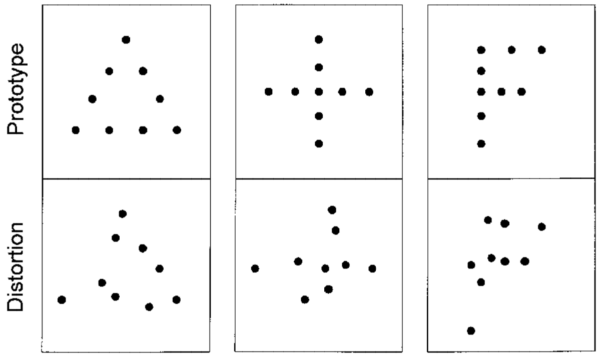

Figure 3. The top part of the figure shows the three prototype patterns used in Experiment 1. Under each prototype is an example of a moderate-level distortion created from that pattern.

Procedure

A standard category learning/transfer paradigm was used. During training, participants learned to categorize the six training patterns from each category. In the *prototype condition*, participants also saw the prototypes during training; in the *no prototype condition*, participants saw the prototypes only at transfer. Patterns were presented once per block, in random order, for eight blocks. On each trial, a pattern was presented and classified as an A, B, or C. Corrective feedback was supplied for five seconds or until the space bar was pressed. After an ITI of one second the next pattern was displayed. The assignment of category to response key was randomized for every participant. Responses were made by pressing labelled keys on a computer keyboard.

During transfer, all thirty patterns (one prototype, six old distortions, and three new distortions from each category) were presented once per block, in random order, for three blocks. No corrective feedback was provided.

Participants also rated the pairwise similarities among patterns. Because of the large number of possible pairs, each participant rated only half of them (randomly selected for each participant). On each trial, two patterns were presented side by side. Participants rated similarities by using a 10-point scale (1 = very dissimilar, 10 = very similar).

Results

Categorization data analyses

The observed category response probabilities for each individual stimulus in the no prototype and prototype conditions are reported in Table 1. A two-way repeated measures analysis of variance (ANOVA) was conducted on the categorization accuracy data (the probability of classifying the item into the correct category) with no prototype vs. prototype as a between-subjects factor and type (prototype, old, or new) as a within-subjects factor. Accuracies for prototypes, old patterns, and new patterns in the no prototype and prototype conditions are summarized in the left panel of Figure 4. Accuracy was higher in the prototype condition than the no prototype condition, $F(1, 276) = 94.12$, $MSE = 0.06$ (alpha was set at .05 for all statistical tests reported in this paper). The effect of type was significant, $F(2, 552) = 261.16$, $MSE = 0.02$; planned comparisons revealed that prototypes were

TABLE 1
Observed and predicted categorization response probabilities for the no prototype and prototype conditions in Experiment 1

	Observed						*Generalized context model*					
	No prototype			*Prototype*			*No prototype*			*Prototype*		
Stimulus	*P(T)*	*P(P)*	*P(F)*	*P(T)*	*P(P)*	*P(F)*	*P(T)*	*P(P)*	*P(F)*	*P(T)*	*P(P)*	*P(F)*
Triangles												
T_P	.798	.088	.114	.926	.045	.029	.745	.122	.134	.903	.047	.049
T_1	.774	.095	.131	.805	.093	.102	.774	.102	.124	.841	.074	.085
T_2	.829	.071	.100	.912	.031	.057	.792	.095	.113	.860	.064	.076
T_3	.429	.283	.288	.491	.264	.245	.447	.224	.329	.449	.229	.323
T_4	.738	.119	.143	.810	.107	.083	.736	.136	.128	.800	.110	.090
T_5	.862	.071	.067	.940	.048	.012	.813	.088	.099	.883	.057	.061
T_6	.614	.136	.250	.690	.107	.202	.642	.149	.209	.691	.129	.180
T_a	.602	.202	.195	.707	.155	.138	.637	.192	.170	.699	.171	.130
T_b	.521	.219	.260	.502	.252	.245	.551	.192	.257	.607	.163	.231
T_c	.462	.343	.195	.524	.343	.133	.560	.252	.187	.599	.250	.151
Pluses												
P_P	.260	.555	.186	.038	.914	.048	.211	.522	.268	.025	.934	.041
P_1	.217	.498	.286	.133	.588	.279	.224	.497	.279	.143	.639	.218
P_2	.138	.607	.255	.088	.714	.198	.142	.568	.290	.091	.674	.235
P_3	.162	.598	.241	.100	.779	.121	.179	.589	.232	.114	.703	.183
P_4	.143	.610	.248	.074	.762	.164	.120	.670	.209	.071	.788	.141
P_5	.131	.600	.269	.064	.652	.283	.151	.538	.312	.087	.639	.274
P_6	.098	.710	.193	.062	.810	.129	.108	.704	.187	.062	.814	.124
P_a	.164	.407	.429	.124	.507	.369	.220	.338	.441	.161	.415	.424
P_b	.219	.390	.391	.143	.502	.355	.219	.427	.355	.167	.504	.329
P_c	.138	.564	.298	.098	.669	.233	.144	.603	.253	.099	.709	.193
Fs												
F_P	.243	.333	.424	.012	.052	.936	.177	.280	.543	.025	.053	.922
F_1	.231	.305	.464	.162	.312	.526	.191	.280	.529	.125	.299	.576
F_2	.074	.181	.745	.076	.183	.740	.108	.274	.618	.072	.243	.685
F_3	.188	.174	.638	.129	.148	.724	.185	.196	.618	.113	.124	.763
F_4	.271	.233	.495	.183	.181	.636	.213	.214	.573	.160	.201	.639
F_5	.174	.274	.552	.152	.243	.605	.189	.214	.597	.132	.194	.674
F_6	.105	.174	.721	.057	.095	.848	.127	.185	.688	.063	.096	.842
F_a	.369	.219	.412	.291	.143	.567	.313	.240	.448	.243	.195	.561
F_b	.150	.291	.560	.124	.255	.621	.168	.331	.501	.127	.312	.562
F_c	.164	.381	.455	.148	.443	.410	.153	.471	.376	.122	.532	.346

Note: T_P = prototype of triangle category; T_1 = old instance of triangle category; T_a = new instance of triangle category. Category T = triangles; Category P = pluses; Category F = Fs.

categorized more accurately than old instances, which were categorized more accurately than new instances. A significant two-way Prototype × Type interaction was observed, $F(2, 552) = 96.86$, $MSE = 0.02$; in the prototype condition, prototypes were categorized significantly more accurately than old instances; however, in the no prototype condition, prototypes were categorized roughly as accurately as old instances. In both conditions, new

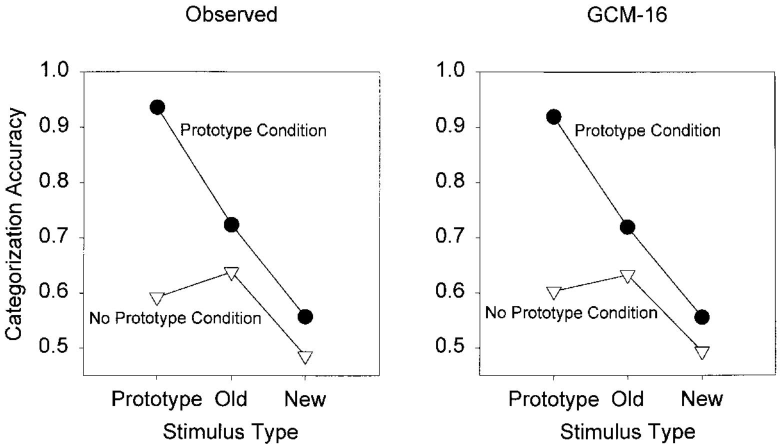

Figure 4. The left panel displays observed and the right panel displays predicted categorization accuracy, collapsed across individual stimuli, for prototypes, old items, and new items in Experiment 1. The no prototype condition is indicated by open triangles, and the prototype condition is indicated by filled circles.

instances were categorized with the lowest accuracy. Note that the prototypes in the prototype condition were the best classified of all stimuli (except for T_5 in the triangle category), a finding that will prove important in the later theoretical analyses. The following sections determine the psychological space of these stimuli and then fit the categorization probability data by using the GCM, prototype, and mixed models.

Multidimensional scaling analyses

The average similarity matrices for the no prototype and prototype conditions were used as input to the INDSCAL scaling model (Carroll & Wish, 1974; Shepard, 1980) to derive a six-dimensional MDS solution (see Appendix A). The INDSCAL procedure gives an MDS configuration that is common to both groups, together with individual dimension weights unique to each group. The overall fit of the INDSCAL-derived distances to the observed similarity ratings was quite good (stress = .069, r^2 = .954). One potential concern was that the psychological space might vary depending on whether prototypes had been viewed during training. Because the INDSCAL solution is constrained to be structurally identical for both groups (subject only to dimensional stretching or shrinking), it is necessary to verify that the MDS solution is reasonable for both groups. First, the fits of the INDSCAL solution to each individual group were also quite good: No prototype stress = .069, r^2 = .954; prototype stress = .068, r^2 = .954. Second, as indicated in Appendix A, the INDSCAL weights, reflecting the importance of each dimension, were comparable across both groups. Therefore, presence or absence of prototypes during training did not seem to appreciably change the locations of the dot patterns in the psychological space.

Figure 5 plots the first three dimensions of the MDS solution; dimensions 1–3 account for 80% of the variance in the similarity ratings (percentage of variance accounted for by a dimension equals the square of the INDSCAL dimension weight). Compare the left panel of Figure 5 with the right panel of Figure 1—the scaling solution is consistent with the idea

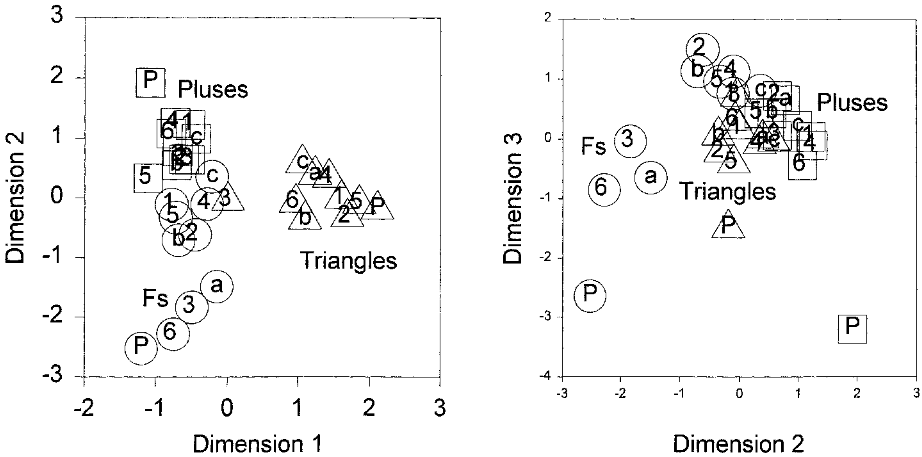

Figure 5. The left panel displays Dimensions 1 and 2, and the right panel displays Dimensions 2 and 3 of the six-dimensional MDS solution given in Appendix A for Experiment 1. The triangles are indicated by triangles, the pluses are indicated by the squares, and the Fs are indicated by circles. Numbers 1–6 indicate old exemplars, letters a–c indicate new exemplars, and P indicates the prototype of each category.

that the prototypes were psychological extremes in relation to other exemplars of their categories, not central tendencies. Although these patterns are indeed quite close to being physical central tendencies of the old instances (as determined by physically averaging the distortions), they are not psychological central tendencies. That said, we do believe that the location of objects in psychological space may be highly dependent on the context in which similarity ratings are made (e.g., Medin, Goldstone, & Gentner, 1993). We observed extreme point representations for prototypes in the context of contrast categories. It seems quite possible that if similarity ratings were collected for patterns from just a single category we could observe central tendency representations for prototypes instead.

To obtain converging evidence that the prototypes are well characterized as extreme points in the psychological space, we conducted the following analysis. First, we calculated the distance between each pair of stimuli in the MDS space as defined in the INDSCAL model. For each individual stimulus, we then computed its average distance to all members of the contrast categories. For example, if a stimulus was from the triangle category, we computed its average distance to all members of the plus and F categories. We refer to this measure as the average between-category distance (*DB*). Likewise, we calculated the average distance of each stimulus to all members within its own category. We refer to this measure as the average within-category distance (*DW*). Finally, we computed a composite measure, *DTOT*, defined as the sum of *DB* and *DW*. For each category, we then rank-ordered the stimuli according to these measures. To the extent that the prototypes occupy extreme points in the psychological space, their values of *DB*, *DW*, and *DTOT* should tend to be large. By contrast, to the extent that the prototypes are central tendencies, their values of *DW* should tend to be small, and their values of *DB* should tend to be intermediate. In the present case, the results were clear-cut: In all three categories, the prototypes were ranked first on the between-category distance measure; that is, they had the largest values of *DB*. Likewise, for the plus and F categories, the prototypes had the largest values of *DW*. (In the triangle category, the prototype was ranked second on this measure.) Finally, in all three categories, the prototypes were ranked first on the composite measure, *DTOT*. This analysis confirms the impression provided by visual inspection of Figure 4, namely, that the prototypes occupied extreme points in the psychological space in which the category exemplars were embedded.

Finally, although global measures of fit, such as stress and percentage of variance accounted for, can be useful and informative, sole reliance on them can sometimes leave undetected systematic problems with the underlying scaling solution. One possibility we thought important to investigate was whether the canonical prototype patterns were nearest neighbours of a relatively large number of category instances. As discussed by Tversky and Hutchinson (1986), spatial models of object similarity are bounded in the number of points any given point can be a nearest neighbour of. Using measures of *centrality* and *reciprocity* that Tversky and Hutchinson developed, we were able to determine that the prototypes were not nearest neighbours of a large number of points,[3] thereby strengthening the idea that the derived MDS solution provides a reasonable description to the similarity structure underlying these dot-pattern categories.

Categorization theoretical analyses

Our next step was to fit the GCM to the observed categorization data from this experiment. We begin with a brief description of the details of the model: According to the GCM, evidence favouring a given category is found by summing the similarity of a presented object and all category exemplars stored in memory. Objects are represented as points in a multidimensional psychological space, with similarity between objects i and j being a decreasing function of their distance in that space,

$$s_{ij} = \exp(-c \cdot d_{ij}) \qquad 1$$

(Shepard, 1987), where c is a sensitivity parameter that scales the psychological space. Distance, d_{ij}, is computed using a simple (weighted) Euclidean metric,

$$d_{ij} = \sqrt{\sum w_m (x_{im} - x_{jm})^2} \qquad 2$$

where x_{im} is the psychological value of object i on dimension m (i.e., the derived coordinates in the MDS solution), and w_m is the attention weight given to dimension m. The weights "stretch" the psychological space along attended dimensions and "shrink" it along unattended ones (Kruschke, 1992; Nosofsky, 1984, 1986). The probability of classifying i as a member of category J is given by

[3] Measures of *centrality* (C) and *reciprocity* (R), as defined by Tversky and Hutchinson (1986), were computed for both the observed similarity matrices and the INDSCAL derived distances. For the observed data, in the no prototype condition, $C = 1.867$ and $R = 2.267$, and in the prototype condition, $C = 1.733$ and $R = 1.967$. For the INDSCAL derived distances, in the no prototype condition, $C = 1.733$ and $R = 2.333$, and in the prototype condition, $C = 2.000$ and $R = 2.133$. Both the observed and the INDSCAL-derived centrality and reciprocity measures were small, indicating that the prototypes were not nearest neighbours of a large number of points, thereby strengthening our faith in the validity of the derived scaling solution. Such small values of C and R are similar in scale to those obtained by Tversky and Hutchinson for perceptual stimuli such as colours, sounds, odours, and simple shapes, and for more complex categories, such as birds, fruits, or weapons when superordinate terms were not included in the set.

$$P(J \mid i) = \frac{\left(b_J \sum_{j \in J} s_{ij}\right)^{\gamma}}{\sum_{K \in R} \left(b_K \sum_{k \in K} s_{ik}\right)^{\gamma}} \qquad 3$$

whereby b_J is the category J response bias, R is the set of all categories, and γ controls the level of deterministic or probabilistic responding (Maddox & Ashby, 1993; McKinley & Nosofsky, 1995; Nosofsky & Palmeri, 1997).

Our main theoretical goal was to determine whether a pure exemplar model, the GCM, could account for the categorization probabilities reported in Table 1. Recall that extreme prototype enhancement was found in the prototype condition of the experiment. As explained earlier, the extreme point representations for prototypes, which emerged from the similarity scaling solution, allow the GCM to qualitatively account for extreme prototype enhancement. Another challenge is whether or not the GCM can make reasonable quantitative predictions as well; the fits of the GCM were compared with those of a pure prototype model and a mixed exemplar-plus-prototype model.

In fitting the GCM to the observed data, for each condition (prototype and no prototype), the full version of the model has eight free parameters: an overall sensitivity parameter (c) in Equation 1; five free attention weights (w_m) in Equation 2 (the six attention weights sum to one); and two free category response biases (b_J) in Equation 3 (the three response biases sum to one); for this particular dataset, the γ parameter could be set equal to 1 without much influence on the fit of the GCM to the observed data. In the full version of the model, different parameters were assumed for the prototype and no prototype conditions, yielding a total of 16 parameters. Various restricted versions of the GCM were also investigated in which parameters were constrained to be the same across both conditions.

The GCM was fitted to the categorization probabilities given in Table 1 (120 free data points) using a maximum-likelihood measure of fit (Wickens, 1982). The predicted categorization probabilities for each individual stimulus are reported in Table 1, and maximum-likelihood parameters and summary fits are reported in Table 2 (GCM-16). The averaged predicted categorization accuracies for the three main types of stimulus (prototype, old items, and new items) are shown in the right panel of Figure 4. The GCM fitted the data quite well, accounting for 97.1% of the variance in the observed data. Averaged across items, the GCM captured all important qualitative trends, and showed very good quantitative predictions. The GCM predicted higher categorization accuracy in the prototype condition than the no prototype condition. Overall, prototypes were predicted to be categorized more accurately than old instances, which were categorized more accurately than new instances. Furthermore, the model predicted correctly that prototypes would be classified more accurately than old instances in the prototype condition, but that prototypes and old instances would be classified with roughly equal accuracy in the no prototype condition. In both conditions, new instances were categorized with the lowest accuracy.

Several restricted versions of the GCM were also tested. Although all yielded significantly worse[4] fits than the full version of the GCM, some of these restricted versions fit the

[4] Likelihood-ratio tests were used to statistically compare models (see Wickens, 1982). Let $\ln L_F$ and $\ln L_R$ denote the log-likelihoods for a full and restricted model, respectively. Assuming the restricted model is correct, the statistic $2(\ln L_F - \ln L_R)$ is distributed as a χ^2 with degrees of freedom equal to the number of constrained parameters. If the observed value of χ^2 exceeds a critical value, then the restricted version fits significantly worse than the full version.

TABLE 2
Maximum-likelihood parameters and summary fits of the categorization data for the full parameter GCM, restricted versions of the GCM, and prototype models in Experiment 1

	Parameter	*GCM-16*	*GCM-9*	*GCM-4*	*Prototype-A*	*Prototype-B*
No prototype	c	1.779	1.723	1.605	0.798	2.001
	w_1	.276	.296	.674	.607	.228
	w_2	.114	.147	.412	.153	.174
	w_3	.334	.311	.412	.000	.322
	w_4	.199	.168	.315	.225	.154
	w_5	.064	.065	.191	.000	.067
	w_6	.012	.013	.156	.015	.055
	b_T	.318	.301	.306	.393	.336
	b_P	.312	.326	.330	.274	.314
	b_F	.370	.373	.365	.333	.351
Fit	−lnL	291.44	303.87	336.01	587.57	350.72
	RMSE	0.044	0.046	0.052	0.085	0.056
	%Var	95.5	95.0	93.8	83.4	92.7
Prototype	c	2.168	2.236	2.051	1.658	1.941
	w_1	.320	.296	.714	.357	.360
	w_2	.188	.147	.474	.173	.274
	w_3	.260	.311	.262	.000	.000
	w_4	.149	.168	.281	.122	.195
	w_5	.070	.065	.206	.227	.092
	w_6	.013	.013	.170	.121	.079
	b_T	.279	.301	.306	.306	.313
	b_P	.344	.326	.330	.287	.328
	b_F	.377	.373	.365	.408	.360
Fit	−lnL	280.29	295.21	327.05	409.79	451.75
	RMSE	0.040	0.043	0.048	0.063	0.066
	%Var	98.0	97.7	97.1	94.9	94.6
Overall Fit	−lnL	571.73	599.08	663.06	997.37	802.47
	RMSE	0.042	0.045	0.050	0.075	0.061
	%Var	97.1	96.7	95.9	90.8	93.9

Note: %Var = variance accounted for (in percentages). c = general sensitivity parameter; w_m = attention weight given to dimension *m*; b_j = bias for making category *j* response; −lnL = negative value of log-likelihood; *RMSE* = root mean squared error between observed and predicted categorization probabilities. GCM-16 = full parameterized GCM; GCM-9 = GCM constrained with weights equal to INDSCAL weights and biases common between the no prototype and prototype conditions; Prototype-A = prototype model using MDS-derived "prototypes"; Prototype-B = prototype model using average of old exemplars. Underlined values are constrained parameters.

data extremely well. First, as shown in Table 2, a nine-parameter version, with attention weights and biases constrained to be the same for both conditions, fit the data very well, accounting for 96.7% of the variance; this result provides evidence that selective attention and response biases probably did not differ very much between the two conditions. Second, a four-parameter version, with attention weights set equal to the INDSCAL weights from the MDS solution and biases constrained to be the same for both conditions, also fit the data quite well, accounting for 95.9% of the variance. Although the categories were presented with equal frequency, for some reason restricted versions with equal biases across the three categories fitted the data significantly worse than the unrestricted versions.

Prototype models, which assume that categorization decisions are based on similarity to an abstracted prototype (e.g., Homa, 1984; Reed, 1972), were also formalized within the MDS framework. The prototype models were identical to the GCM except that rather than computing the summed similarity of an item to all category exemplars, one computes its similarity to the category prototype instead. Two models were tested: in Prototype-A, the derived MDS coordinates of the prototypes were used; in Prototype-B, the prototype representations were generated by spatially averaging old category instance representations (Reed, 1972; Nosofsky, 1988; Shin & Nosofsky, 1992). As shown in Table 2, both 16-parameter models fitted the data worse than the 4-parameter GCM. Predicted categorization accuracies for prototypes, old items, and new items are shown in Figure 6. Prototype-A did not capture the qualitative trends in the data from the no prototype condition, and underestimated the difference in accuracy between old and new items in both conditions. Prototype-B captured most of the qualitative relations, but systematically under- and overestimated the accuracies for prototypes and new items, respectively.

A combined exemplar-plus-prototype model was also investigated (Shin & Nosofsky, 1992). In this model, evidence for a given category was equal to the summed similarity of an item to all category exemplars plus a weighted similarity to the category prototype,

$$P(J|i) = \frac{\left[b_J\left(\sum_{j\varepsilon J} S_{ij} + y\cdot S_{ipJ}\right)\right]^{\gamma}}{\sum_{K\varepsilon R}\left[b_K\left(\sum_{j\varepsilon K} S_{ij} + y\cdot S_{ipK}\right)\right]^{\gamma}} \qquad 4$$

where y is the weight for the prototype and s_{ipJ} *is the similarity between item i* and the prototype for category J. Note that $y = 0$ yields the standard GCM. In fitting the combined model to the data, separate y terms were assumed for the no prototype and prototype conditions. When the prototype was given by the MDS-derived coordinates, the combined model did not fit significantly better than the standard GCM, $-\ln L = 571.27$, $\chi^2(2) = 0.21$. When the prototype was given by the average of the old exemplars, the combined model did fit significantly better than the standard GCM, $-\ln L = 568.19$, $\chi^2(2) = 7.08$, with y(no prototype) $= 3.408$ and y(prototype) $= 1.034$. However, note that the improvement in fit was quite small. Moreover, if anything, we expected to find greater use of prototype information in the prototype conditions; rather, the y term was greater in the no prototype condition.

Discussion

In Experiment 1, we were able to empirically document an extreme prototype enhancement effect in which category prototypes were classified more accurately than other category instances. This finding contrasts with most other studies using the dot-pattern paradigm in which the category prototypes were typically classified with intermediate accuracy relative to individual training instances of a category. The ability of exemplar-based models, such as

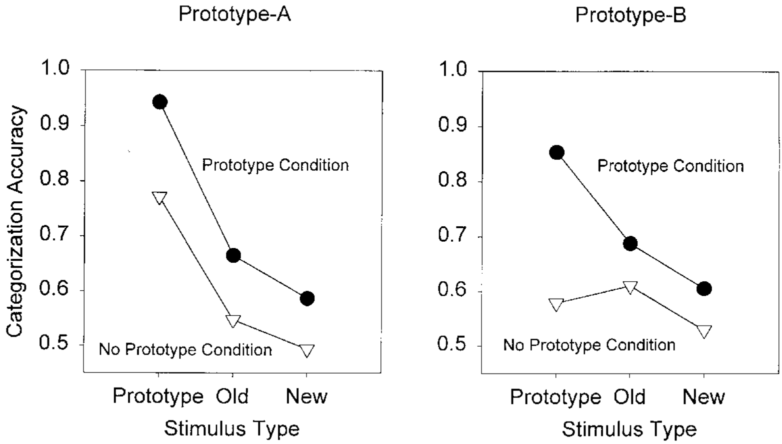

Figure 6. The left panel displays Prototype-A and the right panel displays Prototype-B predicted categorization accuracies, collapsed across individual stimuli, for prototypes, old items, and new items in Experiment 1. The no prototype condition is indicated by open triangles, and the prototype condition is indicated by filled circles.

the GCM, to account for extreme prototype enhancement hinges on the psychological representation of the category prototypes, depending on whether these physical central tendencies are represented as psychological central tendencies or as psychological extreme points.

MDS analyses of participants' similarity ratings revealed that the prototypes were represented as extreme points in the psychological space relative to the category instances. This result is informative because the typical assumption expressed in the categorization literature is that the physical manipulation of prototypes in generating category instances has a fairly direct mapping onto the psychological representations of those prototypes and category instances (e.g., McLaren, Bennet, Guttman-Nahir, Kim, & Mackintosh, 1995).

The GCM can qualitatively predict extreme prototype enhancement when prototypes are represented as extreme points in the psychological space. We were also able to demonstrate that the GCM could provide a good quantitative account of the observed categorization data without needing to include adjunct prototype abstraction processes as well. Even in conditions in which the category prototypes were familiar shapes, people still seemed to rely on exemplar information for making categorization decisions.

EXPERIMENT 2

The goal of Experiment 2 was to find additional evidence for extreme prototype enhancement and for extreme point prototype representations using novel prototypes. To "induce" extreme point representations for novel prototypes, we chose to constrain the category prototypes to have bilateral vertical symmetry (see Figure 7). Symmetry, in a variety of forms, is pervasive in the natural world (Weyl, 1952), and people are extremely sensitive to

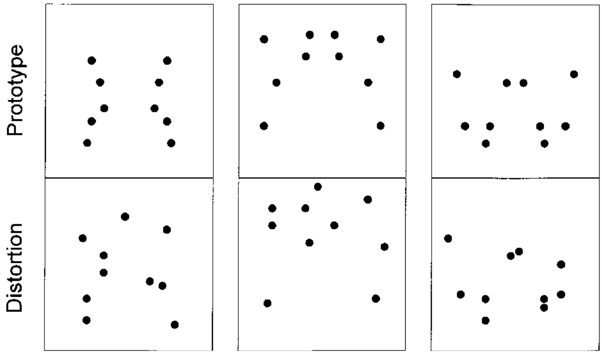

Figure 7. The top part of the figure shows the three prototype patterns for Categories A, B, and C, respectively, used in Experiment 2. Under each prototype is an example of a moderate-level distortion created from that pattern.

symmetry, especially around the vertical axis (e.g., Baylis & Driver, 1995; Bornstein & Krinsky, 1985; Wagemans, 1993; Wagemans, Van Gool, & d'Ydewalle, 1992). Moreover, canonical forms of objects such as leaves, crystals, or sea shells are often pictorially represented in nature field guides with near perfect symmetry; most people would agree that such perfect natural forms are the prototypes of those categories.

Method

Participants

Participants were 118 undergraduates who received course credit. All participants were tested individually.

Stimuli

Three prototype patterns of 10 dots each were created; each was constrained to be symmetric about the vertical axis. Five points were randomly located on the left side of the pattern and five points were mirrored on the right side of the pattern. One dot of a pair was first randomly located in the 30 × 30 grid; the second dot was located having the same vertical coordinate and the negative horizontal coordinate. For example, if the first point had a grid location of (10,12) the second point had a grid location of (−10,12). A constraint was imposed that all dots be at least three units apart. Nine moderate-level distortions (6 bits/dot) of each of these three prototypes were created; six were designated as old training patterns. Figure 7 displays the three prototype patterns along with an example distortion.

Procedure

All procedural details were identical to those used in Experiment 1, except that five training blocks were used.

Results

Categorization data analyses

The observed categorization response probabilities for each pattern in the no prototype and prototype conditions are shown in Table 3. A two-way repeated measures ANOVA was conducted on the accuracy data with no prototype vs. prototype as a between-subjects factor and type (prototype, old, or new) as a within-subjects factor. A significant main effect of type was found, $F(2, 232) = 27.12$, $MSE = 0.01$; planned comparisons revealed that old patterns were categorized more accurately than prototypes, and prototypes were categorized more accurately than new patterns. No main effect or interactions involving no prototype vs. prototype were significant. Inspection of Table 3 reveals that the effect of type depends strongly on the category. The prototypes were classified better than the old items in Categories A and B, but were classified worse than the old items in Category C; Figure 8 summarizes the observed categorization accuracy for prototypes, old items, and new items for Categories A and B (as discussed later, the MDS solution revealed a degenerate psychological configuration for the patterns belonging to Category C, so for illustrative purposes, data from Category C items were not included in the figure). The goal in the theoretical analyses will be to attempt quantitative as well as qualitative accounts of this pattern of results by the GCM, prototype, and mixed models.

Multidimensional scaling analyses

The average similarity matrices were used as input to the INDSCAL scaling model to derive a six-dimensional MDS solution (see Appendix B). The overall fit of the INDSCAL-derived distances to the observed similarity ratings was quite good (stress = .068, $r^2 = .949$). The fits of the INDSCAL solution to the no prototype and prototype groups, individually, were quite good as well (no prototype stress = .073, $r^2 = .942$; prototype stress = .063, $r^2 = .956$).[5]

Figure 9 displays the first three dimensions of the MDS solution. The Category A prototype clearly exhibits an extreme-point representation, and the Category B prototype tends more towards an extreme point than a central tendency representation. These impressions are supported by calculation of the within- and between-category distance measures that we introduced in Experiment 1. For both Categories A and B, the prototypes were ranked first in terms of the composite measure $DTOT$, and were both highly ranked on the component measures DB and DW. However, the Category C prototype exhibits neither an extreme point nor a central tendency representation. (The C prototype was ranked first on the DW measure, last on the DB measure, and intermediate on the $DTOT$ measure.) Unfortunately, by the chance nature of the prototype distortion procedure, the instances of Category C

[5] Nearest neighbour analyses (Tversky & Hutchinson, 1986) were conducted on the observed similarity matrices and the derived MDS distances. For the observed similarity matrices, in the no prototype condition, $C = 2.533$ and $R = 2.467$, and in the prototype condition, $C = 2.133$ and $R = 2.333$. For the derived MDS distances, in the no prototype condition, $C = 2.467$ and $R = 2.700$, and in the prototype condition, $C = 2.267$ and $R = 2.600$. As in Experiment 1, it does not appear that the category prototypes were nearest neighbours of many category instances.

TABLE 3
Observed and predicted categorization response probabilities for the no prototype and prototype conditions in Experiment 2

	Observed						*Generalized context model*					
	No prototype			*Prototype*			*No prototype*			*Prototype*		
Stimulus	*P(A)*	*P(B)*	*P(C)*	*P(A)*	*P(B)*	*P(C)*	*P(A)*	*P(B)*	*P(C)*	*P(A)*	*P(B)*	*P(C)*
Category A												
A_P	.915	.045	.040	.912	.037	.051	.857	.074	.069	.881	.062	.058
A_1	.921	.028	.051	.910	.025	.065	.812	.093	.095	.821	.085	.094
A_2	.751	.090	.158	.771	.085	.144	.763	.135	.103	.756	.142	.102
A_3	.870	.045	.085	.884	.028	.088	.841	.067	.091	.850	.061	.089
A_4	.825	.051	.124	.845	.065	.090	.848	.076	.076	.857	.070	.073
A_5	.819	.102	.079	.831	.071	.099	.830	.078	.092	.839	.071	.091
A_6	.881	.051	.068	.890	.042	.068	.857	.069	.073	.869	.061	.070
A_a	.661	.124	.215	.715	.088	.198	.721	.118	.161	.746	.100	.154
A_b	.576	.158	.266	.551	.178	.271	.671	.158	.170	.669	.154	.177
A_c	.509	.113	.379	.568	.076	.356	.494	.192	.314	.510	.178	.312
Category B												
B_P	.141	.785	.073	.102	.833	.065	.156	.732	.112	.114	.808	.079
B_1	.113	.746	.141	.105	.757	.138	.113	.765	.122	.106	.783	.111
B_2	.062	.898	.040	.057	.893	.051	.102	.799	.099	.101	.809	.090
B_3	.164	.763	.073	.130	.799	.071	.119	.783	.099	.108	.808	.085
B_4	.040	.751	.209	.040	.763	.198	.105	.725	.171	.111	.733	.156
B_5	.085	.701	.215	.073	.726	.201	.105	.688	.207	.111	.671	.218
B_6	.130	.689	.181	.127	.706	.167	.122	.703	.174	.129	.699	.172
B_a	.062	.751	.186	.071	.768	.161	.119	.724	.157	.127	.722	.152
B_b	.130	.735	.136	.124	.757	.119	.167	.686	.147	.141	.741	.118
B_c	.085	.729	.186	.082	.737	.181	.138	.702	.160	.138	.718	.144
Category C												
C_P	.305	.215	.480	.274	.172	.554	.142	.222	.636	.122	.167	.712
C_1	.141	.147	.712	.150	.122	.729	.088	.112	.800	.092	.095	.814
C_2	.141	.107	.751	.130	.099	.771	.091	.099	.811	.099	.088	.814
C_3	.096	.181	.723	.110	.147	.743	.098	.132	.770	.098	.109	.793
C_4	.136	.102	.763	.138	.096	.766	.096	.121	.783	.110	.111	.779
C_5	.085	.113	.802	.073	.105	.822	.090	.102	.808	.101	.098	.800
C_6	.073	.141	.785	.096	.107	.797	.081	.116	.803	.087	.100	.812
C_a	.073	.158	.768	.090	.153	.757	.100	.150	.750	.098	.120	.781
C_b	.136	.147	.718	.170	.153	.678	.119	.152	.729	.117	.124	.759
C_c	.170	.096	.735	.127	.099	.774	.142	.122	.735	.147	.107	.746

Note: A_P = prototype of Category A; A_1 = old instance of Category A; A_a = new instance of Category A.

apparently were extremely similar to one another, as indicated by the compact clustering of points along the three dimensions shown in Figure 9. Perhaps the lack of variance in the category instances limited people's tendency to abstract whatever emergent dimensions might cause the prototypes to be conceived as extreme points.

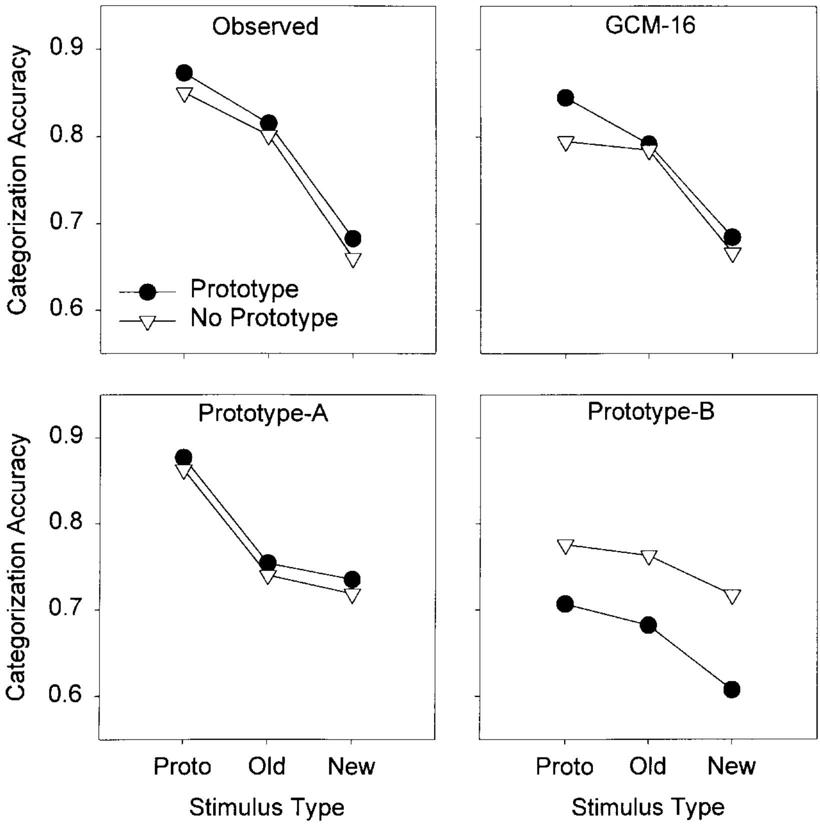

Figure 8. The upper left panel displays observed categorization accuracy, the upper right panel displays GCM predicted categorization accuracy, the lower left panel displays Prototype-A predicted categorization accuracy, and the lower right panel displays Prototype-B predicted categorization accuracy collapsed across individual stimuli, for prototypes, old items, and new items in Experiment 2 (only categorization accuracy for Categories A and B are included). The no prototype condition is indicated by open triangles, and the prototype condition is indicated by filled circles.

Categorization theoretical analyses

As in Experiment 1, the full version of the GCM had sixteen parameters, eight for the prototype and eight for the no prototype condition (five free attention weights, w_m, two free response biases, b_J, and one free sensitivity parameter, c, in each condition: γ could be set equal to one without much influencing the fit of the model). The predicted response probabilities for each individual stimulus from all three categories for the full model are given in Table 3; the best-fitting parameters and fit values are given in Table 4 (GCM–16). The GCM fit quite well, accounting for 97.8% of the variance in the observed data in Table 3 (predictions for just Categories A and B are summarized in Figure 8). For Categories A and B, the model predicted correctly the prototypes to be classified more accurately than the old items, and for Category C, the model predicted correctly the prototypes to be classified less accurately than the old items.

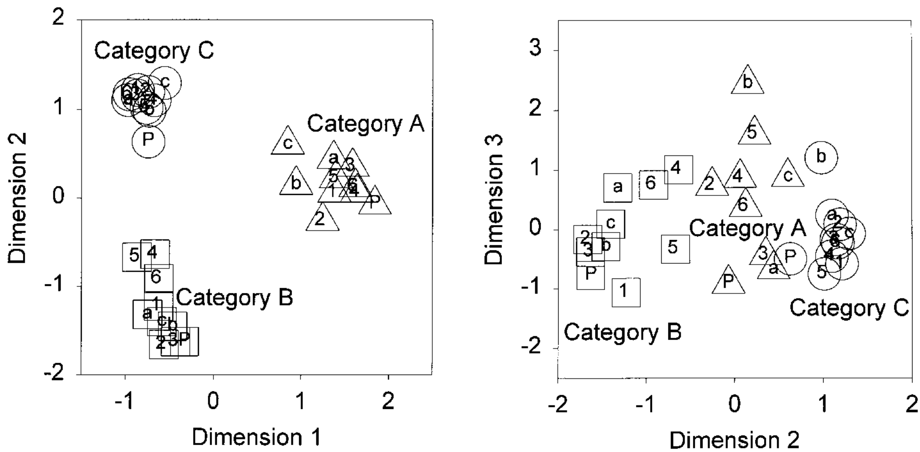

Figure 9. The left panel displays Dimensions 1 and 2, and the right panel displays Dimensions 2 and 3 of the six-dimensional MDS solution given in Appendix B for Experiment 2. Category A exemplars are indicated by triangles, Category B exemplars are indicated by the squares, and Category C exemplars are indicated by circles. Numbers 1–6 indicate old exemplars, letters a–c indicate new exemplars, and P indicates the prototype of each category.

Because no statistical difference was found between the prototype and no prototype conditions, we expected that restricted versions that constrained parameters to be the same in both conditions would also fit the data quite well. First, an eight-parameter version was fitted to the data in which all eight parameters were constrained to be the same in both conditions (GCM–8). This model did not fit the data significantly worse than the full version of the GCM, $\chi^2(8) = 5.86$, $p > .10$. Second, a one-parameter version, with attention weights set equal to the INDSCAL weights from the MDS solution, equal biases for each category, and the same sensitivity parameter for both conditions, fitted the data quite well, accounting for 96.9% of the variance; however, this model did fit significantly worse than the full version, $\chi^2(15) = 41.90$, $p < .001$.

The two versions of the prototype model were also fitted to the observed data. Best fitting parameters and fit values for both models are given in Table 4. Prototype-A, which assumes the MDS coordinates of the prototypes, fitted the data quite well, accounting for 95.5% of the variance in the observed data (predictions for just Categories A and B are summarized in Figure 8). However, note that this 16-parameter prototype model fitted the data worse than the 1-parameter version of the GCM. Furthermore, this version of the prototype model is not the same version of the prototype model that fared well in Experiment 1 (in that experiment, it was Prototype-B that provided a reasonably good fit). Thus, considerations of parsimony favour the exemplar-based interpretation of the data.

Prototype-B, which assumes the prototypes to be central tendencies of the old exemplars, fitted the data quite poorly, accounting for only 81.8% of the variance (predictions for just Categories A and B are summarized in Figure 8). This model performed most poorly in the prototype condition, accounting for only 67.9% of the variance (note that the abstracted prototype in the prototype condition was assumed to be an average of the old distortions and

TABLE 4
Maximum-likelihood parameters and summary fits of the categorization data for the full parameter GCM, restricted versions of the GCM, and prototype models in Experiment 2

	Parameter	*GCM-16*	*GCM-8*	*GCM-1*	*Prototype-A*	*Prototype-B*
No prototype	c	1.856	1.890	1.875	1.291	1.329
	w_1	.354	.344	<u>.707</u>	.521	.516
	w_2	.320	.337	<u>.600</u>	.479	.441
	w_3	.000	.000	<u>.158</u>	.000	.000
	w_4	.007	.006	<u>.154</u>	.000	.027
	w_5	.156	.171	<u>.133</u>	.000	.002
	w_6	.163	.142	<u>.127</u>	.000	.015
	b_A	.315	.324	<u>.333</u>	.402	.345
	b_B	.315	.317	<u>.333</u>	.344	.314
	b_C	.371	.359	<u>.333</u>	.255	.342
Fit	−lnL	227.94	230.36	248.89	290.56	259.09
	RMSE	0.046	0.047	0.054	0.068	0.057
	%Var	97.7	97.7	96.9	95.0	96.5
Prototype	c	1.933	1.890	1.875	1.365	2.561
	w_1	.329	<u>.344</u>	<u>.703</u>	.486	.331
	w_2	.353	<u>.337</u>	<u>.613</u>	.514	.351
	w_3	.000	<u>.000</u>	<u>.161</u>	.000	.000
	w_4	.002	<u>.006</u>	<u>.156</u>	.000	.169
	w_5	.186	<u>.171</u>	<u>.146</u>	.000	.102
	w_6	.130	<u>.142</u>	<u>.126</u>	.000	.047
	b_A	.333	<u>.324</u>	<u>.333</u>	.402	.245
	b_B	.319	<u>.317</u>	<u>.333</u>	.346	.394
	b_C	.347	<u>.359</u>	<u>.333</u>	.253	.362
Fit	−lnL	222.46	225.00	241.92	267.90	922.24
	RMSE	0.045	0.047	0.050	0.063	0.179
	%Var	97.9	97.8	97.5	96.0	67.9
Overall Fit	−lnL	450.40	455.36	490.81	558.45	1181.33
	RMSE	0.046	0.047	0.052	0.066	0.133
	%Var	97.8	97.7	97.2	95.5	81.8

Note: %Var = variance accounted for (in percentages). c = general sensitivity parameter; w_m = attention weight given to dimension m; b_j = bias for making category j response; −lnL = negative value of log-likelihood; *RMSE* = root mean squared error between observed and predicted categorization probabilities. GCM-16 = full parameterized GCM; GCM-8 = GCM constrained with weights, biases and sensitivities common between the no prototype and prototype conditions; GCM-1 = GCM constrained with weights equal to INDSCAL weights, equal category biases, and sensitivities common between the no prototype and prototype conditions; Prototype-A = prototype model using MDS-derived "prototypes"; Prototype-B = prototype model using average of old exemplars. Underlined values are constrained parameters.

the presented prototype). The combined exemplar-plus-prototype model was also investigated. As in Experiment 1, additional weighted prototypes were assumed to be represented in memory. Neither a model that assumed the MDS coordinates of the prototype, $-\ln L = 450.39$, $\chi^2(2) = 0.0001$, nor a model that assumed the average prototype, $-\ln L = 448.67$, $\chi^2(2) = 2.99$, fitted significantly better than the full version of the GCM.

Discussion

In Experiment 2, we found that novel, vertically symmetric, dot-pattern prototypes gave rise to an extreme prototype enhancement effect (for two of the categories tested) and that the prototypes tended to be represented as extreme points rather than as central tendencies. Again, although the prototypes were indeed physical central tendencies of their distortions, they did not emerge as psychological central tendencies. Thus, the finding that dot-pattern prototypes may often have extreme-point representations appears to have some generality.

Although the category prototypes tended to have a “special” representational status as relative extremes in the psychological space, they did not have a special status with regards to making categorization decisions. Extending the theoretical results of Experiment 1, a pure exemplar model (the GCM) provided a good account of the categorization data, whereas simple prototype models fared less well. Furthermore, the combined exemplar-plus-prototype models, which supplement exemplar generalization with prototype abstraction, did not provide a significantly better account of the data than did the GCM.

EXPERIMENT 3

The final experiment in this study examined a finding of extreme prototype enhancement reported by McLaren et al. (1995). These researchers had participants learn two categories of checkerboard patterns that were constructed so that the prototypes were physical central tendencies of the category instances (see also McLaren, 1997; McLaren, Leevers, & Mackintosh, 1994; Wills & McLaren, 1998). The prototype of Category A was a random configuration of white and black squares, as shown in Figure 10. The prototype of Category B was created by randomly switching a relatively large proportion of squares of the Category A prototype from white to black or from black to white, as shown in Figure 10. From these category prototypes, two types of instance were generated with differing relationships to the prototype of the other category. *Close* patterns were generated by switching approximately 10% of the unique squares of one prototype to the colour of the other prototype. This procedure created patterns that were physically more similar than was the prototype to members of the *contrast* category (an example from each prototype is shown in Figure 10). *Far* patterns were generated by switching approximately 10% of the squares common to both prototypes to the opposite colour. This procedure created patterns that were physically more dissimilar than was the prototype to members of the contrast category (an example from each prototype is shown in Figure 10). Thus, physically, this procedure produces patterns with roughly the following schematic arrangement:

FAR–A PROTOTYPE–A CLOSE–A CLOSE–B PROTOTYPE–B FAR–B

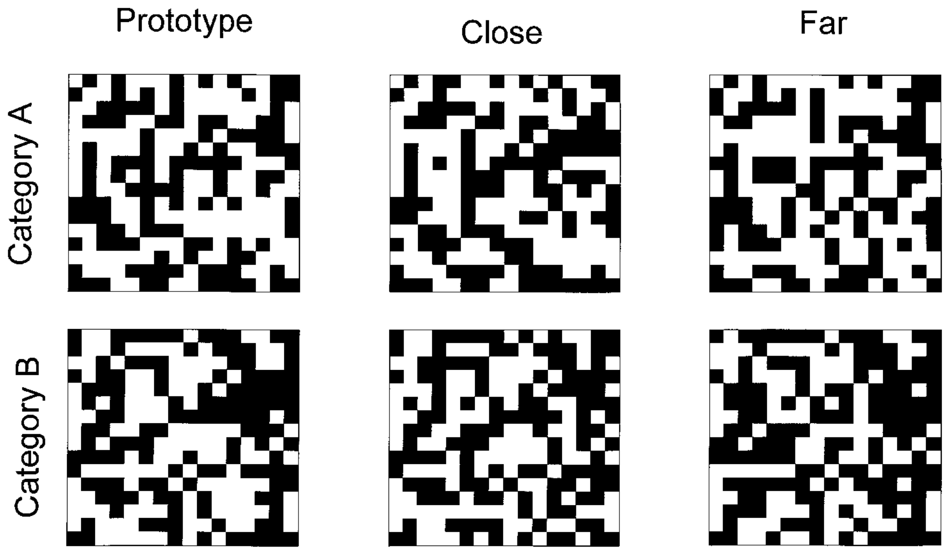

Figure 10. Stimulus used in Experiment 3. The top part of the figure shows the Category A prototype, a close example, and a far example. The bottom part of the figure shows the Category B prototype, a close example, and a far example.

McLaren et al. (1995) found the category prototypes to be the best classified items, a clear case of extreme prototype enhancement. Indeed, McLaren et al. conducted formal theoretical analyses involving the GCM that assumed that the distance between checkerboard patterns was directly related to the number of mismatching squares computed on a city-block metric. In these analyses, McLaren et al. demonstrated that the GCM failed to account for the data and argued that the results posed a serious challenge to exemplar models in general. Lamberts (1996) subsequently demonstrated that by making alternative metric assumptions for calculating distances among the physically defined patterns, exemplar models, such as the GCM, could account for McLaren et al.'s data.

We hypothesized, however, that an even more fundamental issue might also be involved. In particular, the configuration of checkerboard-pattern stimuli in psychological space may be quite different from the physical arrangement defined by the number of overlapping white and black squares. Once again, although the prototypes are central tendencies in the physically defined space, they may exist as extreme points in a psychologically defined space.

McLaren et al. (1995) acknowledged the possibility that the individual black and white squares might give rise to higher order emergent features in which the prototypes were no longer central tendencies. To explore this possibility, they tested a small subset of participants in a separate identification training session. In this session, participants learned a unique label for two prototypes, two close patterns, and two far patterns; the identification confusion data were then used as input for a multidimensional scaling analysis. Although McLaren et al. found some evidence for a psychological arrangement of the stimuli comparable to the physical arrangement, they reported only a *one-dimensional* scaling solution. It seems much more likely that these relatively complex checkerboard patterns should give rise to higher dimensional psychological scaling solutions. Moreover, the fact that only six of the original patterns were examined in the McLaren et al. study potentially limits the generalizability of their results.

Our goal was to replicate the essential features of the McLaren et al. experiments while also conducting a comprehensive similarity rating task in which similarities among all patterns were measured. Unlike in the McLaren et al. study, all participants in our study learned the same set of stimuli. By using this method, the psychological scaling solution reflects an arrangement of the complete set of actual stimuli that were learned rather than an approximate arrangement of a subset of a general class of stimuli. As was the case in Experiments 1 and 2, we predicted that extreme prototype enhancement would be observed and that the prototypes patterns would emerge in the scaling solution as extreme points rather than central tendencies.

Method

Participants

Participants were 133 undergraduates who received course credit. All participants were tested individually.

Stimuli

Unlike the original McLaren et al. (1995) study, all participants in this experiment learned to categorize the same set of patterns. Therefore, the stimulus generation procedure, which follows, was performed only once before the experiment was conducted.

The stimuli were 16×16 checkerboard patterns of white and black squares, 4 pixels on a side (see Figure 10 for examples). The stimuli were presented at a video resolution of 640×480 on 14-in computer monitors. The prototype for Category A was a completely random pattern of white and black squares. The prototype for Category B was created by randomly changing 6 squares per row from white to black or from black to white (96 squares changed). A fixed set of 16 "close" patterns was generated from each prototype by switching a randomly selected 25 of those 96 squares not in common with the other prototype from white to black or from black to white. This procedure produces patterns that have greater physical overlap with the other prototype than does their prototype (although they still have physically more overlap with their own prototype). For each category, 12 of the close patterns were designated training patterns, and the remaining 4 were designated transfer patterns. A fixed set of 4 "far" patterns was generated from each prototype by switching a randomly selected 25 of the 160 squares in common with the other prototype from white to black or from black to white. This procedure produces patterns that have less physical overlap with the other prototype than does their prototype.

Procedure

A standard category learning/transfer paradigm was used. During training, participants learned to classify the 12 designated close patterns from each category. Each of these 24 close patterns was presented once per block for five training blocks. Corrective feedback was supplied for 3 s after every response. After an intertrial interval (ITI) of 500 ms the next pattern was displayed. Participants were urged to respond more quickly if their response times exceeded 4 s. The assignment of category to response key was randomized for every participant. Responses were made by pressing labelled keys on a computer keyboard.

During transfer, all 42 patterns were tested (12 old close, 4 new close, 4 far, and the prototype from each category). Each pattern was presented once per block for two transfer blocks. No corrective feedback was provided. The computer merely reported "OK" for 2 s to indicate that a response had been recorded. After an ITI of 500 ms the next pattern was displayed.

Participants also rated the pairwise similarities among patterns. Because of the large number of possible pairs, each participant rated only one third of them (randomly selected for each participant). On each trial, two patterns were presented side by side. Participants rated similarities using a 10-point scale (1 = very dissimilar, 10 = very similar).

Results

Categorization data analyses

Of the 133 participants, 25 were removed for failing to reach a fairly lax criterion of at least 60% correct on the last block of training. The observed categorization probabilities for each pattern in the transfer phase are shown in Table 5. As summarized in Figure 11, on average, the prototypes were classified the best, followed by the old close patterns, the far patterns, and the new close patterns. A one-way repeated measures ANOVA revealed a significant main effect of stimulus type, $F(3, 321) = 60.40$, $MSE = 0.01$. Planned comparisons revealed that prototype classification accuracy was significantly the best, old close and far pattern classification accuracies were, on average, not significantly different from one another, and new close pattern classification accuracy was, on average, the worst. Extreme prototype enhancement was observed in which the prototypes were classified better than any other pattern.

Overall, these results largely replicate those of McLaren et al. (1995): The prototypes were classified better than any other patterns, even though the statistical distortion algorithm produced prototypes that were roughly physical central tendencies of the category instances. (Unlike McLaren et al., however, we did not find that new far patterns were classified significantly better overall than old close patterns, although this result appears to depend on the specific category being considered—see Figure 11.) The key question now is to examine the psychological scaling solution for these checkerboard patterns and to use it in conjunction with the GCM to test the exemplar-model predictions.

Multidimensional scaling analyses

The average similarity rating matrix was used as input to the KYST scaling model (Kruskal, Young, & Seery, 1973).[6] The resulting scaling solution had a stress of .090. The first two dimensions of the solution, which are most informative, are shown in Figure 12, and the complete scaling solution is reported in Appendix C. Along Dimension 1, the first principal component of the KYST solution, the category prototypes have extreme values

[6] Because each subject only rated 1/3 of the possible pairs of checkerboard patterns, an average similarity rating matrix had to be used for the scaling analyses. Because only a single experimental condition was tested, INDSCAL could not be used in the present study.

TABLE 5
Observed and predicted categorization response accuracy for Experiment 3

Category A			*Category B*		
Stimulus	*Observed*	*Predicted*	*Stimulus*	*Observed*	*Predicted*
PA	0.944	0.898	PB	0.935	0.882
CA_1	0.796	0.803	CB_1	0.620	0.785
CA_2	0.870	0.745	CB_2	0.857	0.836
CA_3	0.833	0.799	CB_3	0.838	0.860
CA_4	0.843	0.863	CB_4	0.829	0.842
CA_5	0.796	0.772	CB_5	0.685	0.742
CA_6	0.926	0.888	CB_6	0.759	0.750
CA_7	0.875	0.884	CB_7	0.847	0.824
CA_8	0.880	0.855	CB_8	0.866	0.874
CA_9	0.815	0.821	CB_9	0.810	0.829
CA_{10}	0.810	0.838	CB_{10}	0.690	0.758
CA_{11}	0.866	0.841	CB_{11}	0.759	0.721
CA_{12}	0.893	0.845	CB_{12}	0.755	0.816
Ca_1	0.681	0.727	Cb_1	0.722	0.760
Ca_2	0.815	0.872	Cb_2	0.625	0.776
Ca_3	0.852	0.802	Cb_3	0.745	0.795
Ca_4	0.875	0.819	Cb_4	0.574	0.576
FA_1	0.662	0.784	FB_1	0.884	0.748
FA_2	0.713	0.726	FB_2	0.810	0.751
FA_3	0.759	0.790	FB_3	0.894	0.767
FA_4	0.722	0.767	FB_4	0.861	0.806

Note: CA_1 = old close pattern of Category A; Ca_1 = new close pattern of Category A; FA_1 = far pattern of Category A; PA = prototype of Category A.

and are not central tendencies of the old patterns. However, the prototypes had roughly equal values along the other dimensions, and calculations of the within- and between-category distance measures that we introduced in Experiment 1 produced intermediate values for the prototypes. Although the prototypes for these grid patterns are not strictly extreme points by our criteria proposed earlier, the extreme values rather than central tendencies along Dimension 1 may allow the exemplar model to adequately predict the extreme prototype enhancement effects and the other aspects of the categorization data.

Categorization theoretical analyses

A version of the GCM with only two free parameters was fitted to the observed data. Parameters were a free sensitivity parameter, c, and a free response scaling term γ. Unlike INDSCAL, the orientation of the psychological dimensions is arbitrary in the KYST model, so incorporating the dimension-weighting parameters was not possible. The predicted categorization response accuracies for the model are shown in Table 5 and are

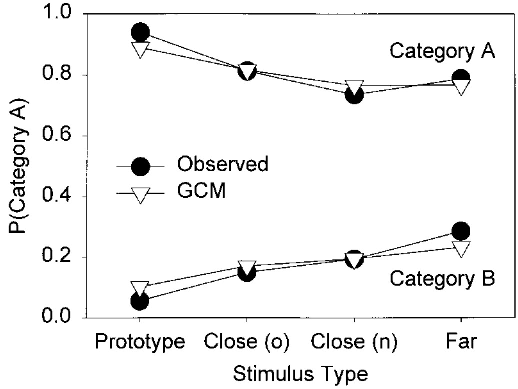

Figure 11. Observed (filled circles) and GCM predicted (open triangles) categorization accuracy, collapsed across individual stimuli, for prototypes, close (old), close (new), and far patterns in Experiment 3.

summarized in Figure 11. This two-parameter version of the GCM fit fairly well, accounting for 95.9% of the variance in the observed data, $-\ln L = 228.11$, $RMSE = 0.063$. The best fitting parameter values were $c = 1.336$ and $\gamma = 4.228$.

Most importantly, the prototypes were predicted to be the best classified items, as was observed. Contrary to McLaren et al. (1995), when combined with the multidimensional scaling solution derived from the similarity ratings participants made, a pure exemplar model was able to account for the qualitative finding of extreme prototype enhancement effects observed in the present experiment.

Inspection of Figure 11 does reveal that the predicted prototype classification probabilities are somewhat smaller than observed. As in the previous experiments, we examined a combined exemplar-plus-prototype model in which evidence for a given category was equal to the summed similarity of an item to all category exemplars plus a weighted similarity to the category prototype. The combined model did provide a significantly better fit to the observed data than did the pure exemplar model, $\chi^2(1) = 22.47$, accounting for 96.6% of the variance, $-\ln L = 205.64$, $RMSE = 0.058$. Although this mixed model somewhat more accurately predicted the prototype classification probabilities, it did not provide noticeably improved fits for the other stimulus types.

Discussion

In this experiment, we replicated the extreme prototype enhancement effects observed in a series of studies by McLaren et al. (1995).[7] From these results, McLaren et al. argued that "an exemplar theory of the type considered here is constrained to predict that far exemplars will always be categorized at least as well as prototypes" (p. 671; but see Lamberts, 1996).

[7] We performed two additional replications and extensions of the McLaren et al. (1995) study using different stimulus sets. These studies provided additional converging evidence for the results reported here. In both studies, extreme prototype enhancement was observed, and the prototypes were psychological extreme points. Moreover, in one of the two studies, the far stimuli were categorized significantly more accurately than the close stimuli, replicating the original McLaren et al. results.

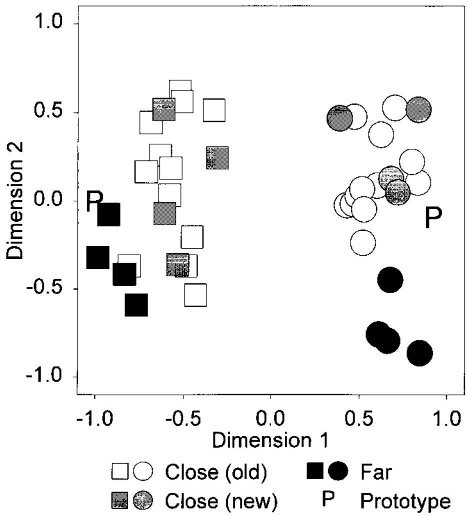

Figure 12. Dimensions 1 and 2 of MDS solutions for Experiment 3. Category A patterns are indicated by circles, and Category B patterns are indicated by squares. White symbols indicate old close patterns, grey symbols indicate new close patterns, black symbols indicate far patterns, and P's indicate prototypes.

McLaren et al.'s claim, however, assumes that the psychological space for the stimuli has a direct correspondence with the physical one. In particular, they assume that in addition to being physically defined central tendencies, the prototypes are psychological central tendencies as well. By contrast, our MDS analyses revealed that the prototypes gave rise to extreme values along one of the emergent psychological dimensions. Furthermore, when a particular exemplar model, the GCM, was used in combination with this derived MDS solution, it provided a reasonably good account of the extreme prototype enhancement that was observed (but see McLaren, 1997, and Wills & McLaren, 1998, for other experimental situations involving these checkerboard stimuli that may involve prototype abstraction following categorization or preexposure).

An assumption made by McLaren et al. (1995) was that each 16 × 16 checkerboard pattern was represented within a 256 dimensional space. In this space, each psychological dimension corresponds directly to the primitive feature of some given square being white or black. We contend that this assumption is probably unjustified, as our MDS analyses reveal. Inspecting the patterns in Figure 10, it seems entirely reasonable to suppose that participants might extract higher level configurations of black and white squares. Clearly, no simple feature-based account could capture properties such as the degree of symmetry in a pattern or the "clumpiness" of a pattern. Also, depending on whether white is foreground or background, different kinds of configurations can emerge. For example, informal discussions with a subset of participants revealed that some noticed configurations such as the white "staircase" in the upper left corner of the Category A prototype or the black "island" in the upper right hand corner of the Category A prototype. The extreme value of the prototypes along Dimension 1 may in part reflect a fairly sophisticated feature creation process, which a simple elemental approach to understanding psychological dimensions overlooks (see Schyns, Goldstone, & Thibaut, 1998).

GENERAL DISCUSSION

The present article was initially motivated by a set of results from the classic dot-pattern categorization paradigm (Homa, 1984; Posner & Keele, 1968, 1970). In such experiments, prototypes are often classified as well as and often times better than old training patterns. Although these results were originally taken as solid evidence for the existence of a prototype abstraction process, these effects have since been shown to be entirely consistent with exemplar models of categorization (e.g., Busemeyer et al., 1984; Hintzman, 1986; Hintzman & Ludlam, 1980; Nosofsky, 1988; Shin & Nosofsky, 1992). Although prototype enhancement effects are regularly observed in such studies, they are typically not very large, and, almost always, at least some of the individual old exemplars are classified better than the prototypes. Whereas innumerable dot-pattern studies have reported prototype enhancement effects in which the prototypes are classified as well as or better than the average of the old distortions, few have reported extreme prototype enhancement effects in which the prototypes are classified better than all instances of a category.[8] As described earlier, finding extreme prototype enhancement for prototypes that are psychological central tendencies of old distortions may prove extremely difficult, if not impossible, for an exemplar model, such as the GCM, to account for—such a finding could potentially falsify the GCM.

Therefore, our goal in developing these studies was to attempt to empirically create extreme prototype enhancement effects that might prove exceptionally challenging to exemplar models. In Experiment 1, we did so by using dot-pattern prototypes that were simplified representations of highly familiar objects. This method is precisely the one that Posner and Keele used in their very first studies. It seemed reasonable to hypothesize that if extreme prototype enhancement were ever to be found in a dot-pattern paradigm, it would be found using such prototypes. In Experiment 2, we used symmetric dot-pattern prototypes. Again, it seemed reasonable that such prototype patterns might exhibit a special status vis-à-vis their distortions, which would cause them to be classified more accurately. Finally, in Experiment 3, we borrowed an experimental paradigm recently reported by McLaren et al. (1995), using checkerboard patterns of white and black squares, that was also shown to give rise to very high classification accuracy for the category prototypes. In Experiments 1 and 3, we found evidence for extreme prototype enhancement—category prototypes were classified better than any of the old category examples on which participants were trained. Tendencies for extreme prototype enhancement were also observed in Experiment 2.

Recall, these results pose a serious challenge to exemplar models only if the prototypes, which are physical central tendencies of category instances, are psychological central tendencies as well. For the dot-pattern stimuli used in Experiments 1 and 2, and for the checkerboard patterns used in Experiment 3, the categories were constructed in such a way that the prototypes were physical central tendencies of the category instances. That is, a composite image made by averaging together all of the category instances looks almost identical to the category prototype. The fact that the prototypes are physical central tendencies

[8] Unfortunately, only quite recently have studies reported classification probabilities for all of the individual stimuli used in dot-pattern experiments. Typically, only average classification accuracies for various types of stimulus (prototype, old distortion, new distortion) were reported, making it impossible to assess whether extreme prototype enhancement was observed.

does not necessarily imply that they are psychological central tendencies as well. For stimuli with clearly defined psychological dimensions, such as semicircles of varying sizes containing radial lines of varying angles (e.g., Nosofsky, 1986; Shepard, 1964), tones varying in loudness and pitch (Melara & Marks, 1990), or perhaps even schematic faces varying the location of facial features (e.g., Nosofsky, 1991), a fairly direct mapping between physical and psychological dimensions may exist. However, the mappings between physical properties and psychological dimensions of fairly complex stimuli, such as artificial dot-patterns or checkerboard patterns and perhaps more natural stimuli, are not so clearly defined (Hock et al., 1988; Shin & Nosofsky, 1992).

To determine the location of the category prototypes relative to the other category instances, participants provided similarity ratings between pairs of stimuli. These ratings were then subjected to a multidimensional scaling analysis to recover the underlying psychological space of the stimuli used in each experiment. Across all three experiments, the scaling analyses revealed the category prototypes to be relative extremes in the psychological space rather than central tendencies. Across all three experiments, when coupled with this obtained psychological scaling solution, one particular exemplar model, the GCM, could both qualitatively and quantitatively account for the observed extreme prototype enhancement effects, as well as other aspects of the observed data. By contrast, various types of prototype model either failed to account for the observed data, or the best-fitting version of the prototype model was not of the same type in different experiments. In addition, little or no evidence was found that pointed to the need to supplement an exemplar model with an additional prototype abstraction process.

It should be stressed that the MDS solutions incorporated in the GCM modelling were derived from similarity judgements following category learning. Although research on this issue is needed, we think it is likely that the modelling approach would be considerably less successful if the similarity judgements were obtained in a completely independent context. The present types of complex dot-pattern can probably be coded and represented in a variety of highly flexible ways. Combined with the fact that similarity judgements are themselves highly context dependent (Medin et al., 1993; Tversky, 1977), this flexibility of coding suggests a need to derive the MDS representation in the context of the category-learning situation. The most challenging future goal is to understand how the psychological representations are formed—how and why did the prototypes come to be represented psychologically as extreme points? Much of the work in categorization has assumed representations to remain relatively unchanged with experience, apart from changes in how dimensions are selectively attended (Kruschke, 1992; Nosofsky, 1986). Recent work, however, has begun to suggest that an important component of categorization involves a more complex form of perceptual learning—not only must a person learn what features are diagnostic, they must also create new diagnostic features (e.g., Lesgold et al., 1988; Schyns et al., 1998; Schyns & Murphy, 1991; Schyns & Rodet, 1997). Rather than assume that the perceptual system provides a set of fixed features to be used as inputs to higher level categorization processes, Schyns et al. (1998) have suggested that flexible, functional features may be created as part of the process of category learning. A reasonable hypothesis is that the extreme-point prototype representations in the present studies arose from the emergence of these diagnostic, functional features in the context of learning these particular dot-pattern categories. One possible starting point for incorporating perceptual learning mechanisms into theories of categorization might be an associative model proposed by McLaren and colleagues (McLaren, Kaye, & Mackintosh, 1989; McLaren, 1997).

Given the nature of the dot-pattern stimuli and the flexibility of coding that arises, it seemed necessary to collect the similarity judgement data only following the completion of category learning. A potential concern that arises is that the prototypes emerged as extreme points in the similarity representations *because* they were judged as the best examples of their categories. According to this view, the prototypes are indeed central tendencies in the "true" psychological space in which the patterns are embedded. Because observers may allow judged category goodness to influence their similarity ratings, however, the MDS analyses of the similarity data revealed a distorted psychological space. To the extent that functional features are indeed created by the act of categorizing, as hypothesized by Schyns and his colleagues (Schyns et al., 1998; Schyns & Rodet, 1997), these alternatives become extraordinarily difficult to disentangle. Nevertheless, we can point to some indirect evidence that poses problems for the prototype-as-central-tendency view. First, consider category-learning situations in which stimuli vary along a few salient psychological dimensions and where creation of new functional features tends not to take place. Examples include learning to classify colours varying in their brightness and saturation, or simple geometric forms varying in their size and angle of orientation. If the view is that the central tendency of the distribution is privileged, and similarity judgements among patterns simply follow judged category goodness, then when using such stimuli, the central tendency should be the best classified item and an MDS analysis based on similarity data should place it at the extremes of the category distribution. A vast literature, however, indicates that such is not the case. Instead, when stimuli with these types of clear-cut psychological dimension are used, the central tendency tends not to be among the most accurately or rapidly classified objects, and scaling solutions based on similarity judgement data place it in the centre of the category distribution. Thus, although we cannot rule out the possibility that in the present experiments the prototype was a "true" psychological central tendency and that the similarity judgements led to a distorted representation, this view appears to be severely limited in its generality.

The bottom line is that we have provided evidence in favour of one particular theoretical approach to understanding extreme prototype enhancement effects that are observed for stimuli composed of highly flexible and complex psychological dimensions. In this approach, the prototypes are to be interpreted as extreme points in the psychological similarity space that emerges from categorization experience. The MDS-based exemplar model has been highly successful at accounting for details of classification performance in numerous domains involving stimuli varying along clear-cut and salient psychological dimensions. The present research demonstrates generality for the approach by using the same types of similarity scaling method as in this previous work, but where the nature of the psychological dimensions that compose the objects is far more flexible and complex. It is an open question whether or not alternative models can be developed that match the generality and precision of this MDS-based exemplar-modelling approach.

REFERENCES

Anderson, J.R. (1980). *Cognitive psychology and its implications*. San Francisco: W.H. Freeman and Company.

Ashby, F.G., & Gott, R.E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 33–53.

Barresi, J., Robbins, D., & Shain, K. (1975). Role of distinctive features in the abstraction of related concepts. *Journal of Experimental Psychology: Human Learning and Memory*, *104*, 360–368.

Barsalou, L.W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 629–654.

Barsalou, L.W. (1991). Deriving categories to achieve goals. *The Psychology of Learning and Motivation*, *27*, 1–64.

Baylis, G.C., & Driver, J. (1995). Obligatory edge assignment in vision: The role of figure and part segmentation in symmetry detection. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1323–1342.

Bornstein, M.H., & Krinsky, S.J. (1985). Perception of symmetry in infancy: The salience of vertical symmetry and the perception of pattern wholes. *Journal of Experimental Child Psychology*, *39*, 1–19.

Breen, T.J., & Schvaneveldt, R.W. (1986). Classification of empirically derived prototypes as a function of category experience. *Memory & Cognition*, *14*, 313–320.

Busemeyer, J.R., Dewey, G.I., & Medin, D.L. (1984). Evaluation of exemplar-based generalization and the abstraction of categorization information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 638–648.

Carroll, J.D., & Wish, M. (1974). Models and methods for three-way multidimensional scaling. In D.H. Krantz, R.C. Atkinson, R.D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 2, pp. 57–105). San Francisco: Freeman.

Goldstone, R.L. (1993). Evidence for interrelated and isolated concepts from prototype and caricature classifications. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 498–503). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Hintzman, D.L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.

Hintzman, D.L., & Ludlam, G. (1980). Differential forgetting of prototypes and old instances: Simulations by an exemplar-based classification model. *Memory & Cognition*, *8*, 378–382.

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

Homa, D. (1978). Abstraction of ill-defined form. *Journal of Experimental Psychology: Human Learning and Memory*, *4*, 407–416.

Homa, D. (1984). On the nature of categories. *Psychology of Learning and Motivation*, *18*, 49–94.

Homa, D., & Chambliss, D. (1975). The relative contribution of common and distinctive information on the abstraction of ill-defined categories. *Journal of Experimental Psychology: Human Learning and Memory*, *104*, 351–359.

Homa, D., Cross, J., Cornell, D., Goldman, D., & Schwartz, S. (1973). Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. *Journal of Experimental Psychology*, *101*, 116–122.

Homa, D., Dunbar, S., & Nohre, L. (1991). Instance frequency, categorization, and the modulating effects of experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 444–458.

Homa, D., Goldhardt, B., Burruel-Homa, L., & Smith, J.C. (1993). Influence of manipulated category knowledge on prototype classification and recognition. *Memory & Cognition*, *21*, 529–538.

Homa, D., Rhoads, D., & Chambliss, D. (1979). Evolution of conceptual structure. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 11–23.

Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *7*, 418–439.

Homa, D., & Vosburgh, R. (1976). Category breadth and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 322–330.

Knapp, A.G., & Anderson, J.A. (1984). Theory of categorization based on distributed memory storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 616–637.

Knowlton, B.J., & Squire, L.R. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science*, *262*, 1747–1749.

Kolodny, J.A. (1994). Memory processes in classification learning: An investigation of amnesic performance in categorization of dot patterns. *Psychological Science*, *5*, 164–169.

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

Kruskal, J.B., Young, F.W., & Seery, J.B. (1973). *How to use KYST: A very flexible program to do multidimensional scaling and unfolding*. Murray Hill, NJ: Bell Telephone Laboratories.

Lamberts, K. (1996). Exemplar models and prototype effects in similarity-based categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1503–1507.

Lesgold, A., Glaser, R., Rubinson, H., Klopfer, D., Feltovich, P., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing X-ray pictures. In M.T.H. Chi, R. Glaser, & M.J. Farr (Eds.), *The nature of expertise* (pp. 311–342). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Maddox, W.T., & Ashby, F.G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, *53*, 49–70.

Massaro, D.W. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Massaro, D.W., & Friedman, D. (1990). Models of integration given multiple sources of information. *Psychological Review*, *97*, 225–252.

McClelland, J.L., & Rumelhart, D.W. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*, 159–188.

McKinley, S.C., & Nosofsky, R.M. (1995). Investigations of exemplar and decision bound models in large, ill-defined category structures. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 128–148.

McLaren, I.P.L. (1997). Categorization and perceptual learning: An analogue of the face inversion effect. *The Quarterly Journal of Experimental Psychology*, *50A*, 257–273.

McLaren, I.P.L., Bennet, C.H., Guttman-Nahir, T., Kim, K., & Mackintosh, N.J. (1995). Prototype effects and peak shift in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 662–673.

McLaren, I.P.L., Kaye, H., & Mackintosh, N.J. (1989). An associative theory of representation of stimuli: Applications to perceptual learning and latent inhibition. In R.G.M. Morris (Ed.), *Parallel distributed processing–implications for psychology and neurobiology*. Oxford: Oxford University Press.

McLaren, I.P.L., Leevers, H.L., & Mackintosh, N.J. (1994). Recognition, categorization, and perceptual learning. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV*. Cambridge, MA: MIT Press.

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

Medin, D.L., & Schaffer, M.M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.

Melara, R.D., & Marks, L.R. (1990). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 398–414.

Metcalfe, J. (1982). A composite holographic associative recall model. *Psychological Review*, *89*, 627–661.

Metcalfe, J., & Fisher, R.P. (1986). The relation between recognition memory and classification learning. *Memory & Cognition*, *14*, 164–173.

Nosofsky, R.M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 104–114.

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

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

Nosofsky, R.M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 700–708.

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

Nosofsky, R.M., & Palmeri, T.J. (1997). An exemplar-based random walk model of speeded classification, *Psychological Review*, *104*, 266–300.

Nosofsky, R.M., & Zaki, S.R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science*, *9*, 247–255.

Oden, G.C., & Massaro, D.W. (1978). Integration of featural information in speech perception. *Psychological Review*, *85*, 172–191.

Omohundro, J. (1981). Recognition vs. classification of ill-defined category exemplars. *Memory & Cognition*, *9*, 324–331.

Palmeri, T.J., & Flanery, M.A. (1999). Learning about categories in the absence of training: Profound amnesia and the relationship between perceptual categorization and recognition memory. *Psychological Science*, *10*, 526–530.

Posner, M.I. (1969). Abstraction and the process of recognition. In G.H. Bower & J.T. Spence (Eds.), *The psychology of learning and motivation* (pp. 43–100). New York: Academic Press.

Posner, M.I., Goldsmith, R., & Welton, K.E., Jr. (1967). Perceived distance and the classification of distorted patterns. *Journal of Experimental Psychology*, *73*, 28–38.

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

Posner, M.I., & Keele, S.W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, *83*, 304–308.

Reber, P.J., & Squire, L.R. (1997). Implicit learning in patients with Parkinson's disease. *Abstracts of the Psychonomic Society*, *2*, 9.

Reber, P.J., Stark, C.E.L., & Squire, L.R. (1998). Cortical areas supporting category learning identified using functional MRI. *Proceedings of the National Academy of Sciences USA*, *95*, 747–750.

Reed, S.K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, *3*, 382–407.

Rosch, E. (1975a). Cognitive reference points. *Cognitive Psychology*, *7*, 532–547.

Rosch, E. (1975b). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*, 192–233.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B.B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Schyns, P.G., & Murphy, G.L. (1991). 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.

Shepard, R.N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, *1*, 54–87.

Shepard, R.N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, *210*, 390–398.

Shepard, R.N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*, 1317–1323.

Shepard, R.N., & Chang, J.J. (1963). Stimulus generalization in the learning of classifications. *Journal of Experimental Psychology*, *65*, 94–102.

Shin, H.J., & Nosofsky, R.M. (1992). Similarity-scaling studies of dot-pattern classification and recognition. *Journal of Experimental Psychology: General*, *121*, 278–304.

Smith, E.E., & Medin, D.L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Squire, L.R., & Knowlton, B.J. (1995). Learning about categories in the absence of memory. *Proceedings of the National Academy of Sciences USA*, *92*, 12470–12474.

Strange, W., Kenney, T., Kessel, F., & Jenkins, J. (1970). Abstraction over time of prototypes from distortions of random dot patterns. *Journal of Experimental Psychology*, *83*, 508–510.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.

Tversky, A., & Hutchinson, J.W. (1986). Nearest neighbor analysis of psychological space. *Psychological Review*, *93*, 3–22.

Vandierendonck, A. (1984). Concept learning, memory and transfer of dot pattern categories. *Acta Psychologica*, *55*, 71–88.

Wagemans, J. (1993). Skewed symmetry: A nonaccidental property used to perceive visual forms. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 364–380.

Wagemans, J., Van Gool, L., & d'Ydewalle, G. (1992). Orientational effects and component processes in symmetry detection. *Quarterly Journal of Experimental Psychology*, *44A*, 475–508.

Weyl, H. (1952). *Symmetry*. Princeton, NJ: Princeton University Press.

Wickens, T.D. (1982). *Models for behavior: Stochastic processes in psychology*. New York: Freeman.

Wills, A.J., & McLaren, I.P.L. (1998). Perceptual learning and free classification. *The Quarterly Journal of Experimental Psychology*, *51B*, 235–270.

Original manuscript received 11 October 1999
Accepted revision received 22 March 2000

APPENDIX A

Six-dimensional scaling solution for the dot patterns in Experiment 1

	Dimension					
Stimulus	*1*	*2*	*3*	*4*	*5*	*6*
Triangle						
T_P	2.106	−0.190	−1.490	1.566	0.151	−0.234
T_1	1.602	−0.009	0.174	0.861	−0.110	−0.265
T_2	1.687	−0.323	−0.219	0.681	0.051	−1.020
T_3	0.003	−0.044	0.717	0.978	−0.123	−2.586
T_4	1.428	0.342	−0.076	0.454	0.496	0.092
T_5	1.850	−0.097	−0.401	1.079	0.247	−0.372
T_6	0.964	−0.076	0.316	0.773	−0.978	−0.248
T_a	1.233	0.392	0.016	−0.004	0.278	−0.800
T_b	1.095	−0.359	0.067	−0.235	−1.552	−1.751
T_c	1.054	0.586	−0.035	−0.122	0.881	0.402
Pluses						
P_P	−1.082	1.921	−3.199	−0.873	0.432	0.535
P_1	−0.528	1.223	0.041	0.826	−0.745	1.308
P_2	−0.614	0.648	0.726	−0.749	0.815	1.640
P_3	−0.532	0.643	0.061	−0.129	1.941	−0.847
P_4	−0.733	1.258	−0.101	−1.298	−0.676	0.363
P_5	−1.107	0.330	0.437	0.737	1.919	−0.260
P_6	−0.791	1.076	−0.436	−1.547	0.465	0.036
P_a	−0.702	0.782	0.653	0.701	−1.584	0.385
P_b	−0.713	0.552	0.441	0.036	−0.299	−2.039
P_c	−0.447	0.991	0.231	−1.754	1.148	−0.468
Fs						
F_P	−1.200	−2.527	−2.646	−0.442	−0.022	0.980
F_1	−0.784	−0.110	0.759	1.713	1.706	−0.341
F_2	−0.457	−0.620	1.490	−1.659	1.459	1.360
F_3	−0.499	−1.835	−0.044	0.010	−1.127	−0.400
F_4	−0.292	−0.099	1.129	0.417	−1.428	1.111
F_5	−0.724	−0.330	0.959	0.744	−1.513	0.201
F_6	−0.756	−2.285	−0.857	−0.561	−0.848	0.351
F_a	−0.151	−1.489	−0.652	0.736	−0.507	0.890
F_b	−0.687	−0.708	1.130	−0.646	−0.595	1.584
F_c	−0.222	0.357	0.811	−2.293	0.119	0.395

Note: INDSCAL weights for the no prototype group are .674, .412, .412, .315, .191, and .156, and for the prototype group are .714, .474, .262, .281, .206, and .170. T_P = prototype of triangle category; T_1 = old instances of triangle category; T_a = new instance of triangle category. Category T = triangles; Category P = pluses; Category F = Fs.

APPENDIX B

Six-dimensional scaling solution for the dot patterns in Experiment 2

	Dimension					
Stimulus	*1*	*2*	*3*	*4*	*5*	*6*
Category A						
A_P	1.834	−0.070	−0.913	0.702	0.051	0.101
A_1	1.374	0.064	−3.266	−0.726	1.154	−0.162
A_2	1.240	−0.257	0.738	−0.012	−1.406	0.735
A_3	1.578	0.354	−0.434	−0.090	0.943	0.304
A_4	1.628	0.062	0.870	0.095	−0.396	0.086
A_5	1.391	0.227	1.597	0.613	0.166	−0.200
A_6	1.611	0.125	0.385	−0.332	0.283	−0.211
A_a	1.363	0.441	−0.672	−0.462	0.066	−1.461
A_b	0.941	0.151	2.445	−0.381	−1.078	0.279
A_c	0.846	0.605	0.883	0.068	−1.681	−1.958
Category B						
B_P	−0.333	−1.627	−0.767	1.078	−1.144	0.211
B_1	−0.628	−1.226	−1.058	−1.136	−1.889	−0.647
B_2	−0.567	−1.661	−0.165	0.428	0.336	−0.730
B_3	−0.434	−1.630	−0.373	0.280	−0.332	0.498
B_4	−0.668	−0.636	1.010	−0.202	2.943	−1.491
B_5	−0.869	−0.672	−0.329	−2.380	−0.062	−0.154
B_6	−0.624	−0.917	0.752	−1.497	0.528	0.412
B_a	−0.752	−1.319	0.704	−1.357	1.090	−0.707
B_b	−0.467	−1.455	−0.284	0.244	−1.197	1.103
B_c	−0.590	−1.403	0.094	1.358	0.894	0.689
Category C						
C_P	−0.742	0.634	−0.498	2.547	−0.317	0.072
C_1	−0.863	1.213	−0.577	1.118	0.429	−0.133
C_2	−0.746	1.191	0.085	0.356	1.041	1.369
C_3	−0.866	1.150	−0.171	−0.321	−0.801	−0.114
C_4	−0.675	1.098	−0.452	1.098	−0.127	−2.612
C_5	−0.763	1.015	−0.751	−1.437	0.171	2.397
C_6	−0.957	1.166	−0.225	0.153	1.264	0.084
C_a	−0.971	1.101	0.237	−1.208	−0.445	0.427
C_b	−0.733	0.982	1.205	0.767	−0.006	0.189
C_c	−0.559	1.292	−0.068	0.634	−0.476	1.625

Note: INDSCAL weights for the no prototype group are .707, .600, .158, .154, .133, and .127, and for the prototype group are .703, .613, .161, .156, .146, and .126. T_P = prototype of Category A; A_1 = old instances of Category A; A_a = new instance of Category A.

APPENDIX C

Six-dimensional scaling solution for the checkerboard patterns in Experiment 3

	Dimension					
Stimulus	*1*	*2*	*3*	*4*	*5*	*6*
Category A						
C1	0.429	−0.023	0.214	−0.490	−0.175	0.307
C2	0.462	−0.015	−0.440	0.597	−0.082	0.323
C3	0.489	0.028	0.481	−0.051	0.755	0.133
C4	0.602	0.089	0.335	−0.160	0.173	0.288
C5	0.520	−0.236	−0.122	−0.471	0.359	0.171
C6	0.836	0.107	0.050	0.267	0.144	−0.035
C7	0.802	0.225	−0.122	0.057	−0.146	0.323
C8	0.624	0.377	0.361	0.004	0.164	−0.315
C9	0.471	0.477	0.179	0.088	−0.668	0.098
C10	0.515	0.068	0.353	−0.298	−0.199	−0.259
C11	0.526	−0.046	0.168	0.235	−0.412	0.350
C12	0.705	0.529	−0.128	0.176	−0.256	−0.350
C13	0.388	0.468	−0.140	−0.440	−0.334	0.520
C14	0.839	0.519	0.137	0.447	0.045	0.150
C15	0.682	0.127	−0.236	−0.505	0.017	−0.215
C16	0.719	0.050	−0.186	0.016	0.250	−0.335
F1	0.677	−0.447	0.078	0.529	0.008	−0.562
F2	0.847	−0.864	−0.899	−0.060	−0.121	0.141
F3	0.611	−0.758	0.471	−0.003	0.028	−0.220
F4	0.662	−0.794	−0.085	0.254	0.211	0.046
P	0.924	−0.114	0.166	0.041	−0.079	−0.086
Category B						
C1	−0.453	−0.203	−0.444	0.530	0.031	−0.196
C2	−0.636	0.262	−0.020	0.165	0.608	−0.022
C3	−0.576	0.187	−0.540	0.170	0.422	0.108
C4	−0.715	0.167	−0.385	−0.028	−0.450	0.062
C5	−0.433	−0.534	0.126	−0.594	0.358	0.088
C6	−0.487	−0.368	0.070	−0.542	−0.213	0.073
C7	−0.688	0.446	0.099	0.363	0.128	0.243
C8	−0.809	−0.367	−0.298	−0.388	0.293	0.078
C9	−0.522	0.619	−0.379	0.273	0.375	−0.177
C10	−0.327	0.515	−0.475	−0.303	0.156	−0.426
C11	−0.511	0.563	0.204	0.059	−0.459	0.138
C12	−0.583	0.034	−0.268	−0.086	−0.039	0.484
C13	−0.610	−0.070	−0.106	−0.269	−0.360	−0.631
C14	−0.535	−0.361	−0.519	−0.087	−0.393	−0.322
C15	−0.616	0.520	0.063	−0.217	0.332	0.143
C16	−0.308	0.246	0.517	−0.597	0.156	−0.335
F1	−0.987	−0.320	0.843	0.261	−0.325	0.043
F2	−0.769	−0.590	0.345	0.457	0.291	0.564
F3	−0.836	−0.415	0.199	0.114	−0.725	−0.058
F4	−0.925	−0.075	0.391	0.481	0.177	−0.459
P	−1.006	−0.024	−0.054	0.008	−0.048	0.127

Note: C1–C12 = old close patterns; C13–C16 = new close patterns; F1–F4 far patterns; P = prototype.