In the categorization research, one of the central and the most debated issues has been the nature of category representation. According to the prototype theory, people represent a category by abstracting its central tendency from constituent instances, and make categorization judgements based on similarity to the prototype (…). In contrast, according to the exemplar model, people represent a category by storing individual constituent exemplars in memory, and base categorization judgements on summed similarity to the exemplars (…).

Initially, many experimental results in favor of the prototype model was obtained from the dot pattern paradigm first proposed by Posner and Keele (1968). In a typical dot pattern experiment, prototypes representing different categories were first generated by randomly placing nine dots (connected or not) in a grid, then patterns of various levels of distortions were constructed by displacing each dots of the corresponding prototype according to a statistical-distortion rule. The experiment consists of a learning phase, in which subjects were trained to classify a number of distorted patterns, followed by a transfer phase, in which subjects were tested on classifying patterns including old training distortions, the prototypes and various new distortions. Early studies found that in the transfer phase, the classification accuracy of prototype not presented in the learning phase were often higher than various new distortions, sometimes even exceeding that of old distortions. The prototype enhancement effect was cited as an evidence for the abstraction of a prototype (Posner & Keele, 1968, 1970, …). Moreover, the classification accuracy tended to decrease for new patterns with higher level of distortions. The typicality gradient is also consistent with the prototype model as patterns of higher levels of distortion are less similar to the prototype.

However, these results were also shown to be compatible with the exemplar model (Hintzman, 1986, …). The typicality gradient effect is confirmed as the summed similarity to old exemplars of a category tend to be higher for patterns closer to the center of category. The prototype enhancement effect is confirmed as the prototype is on average more similar to numerous old exemplars compared to a specific old exemplar that exactly matches itself but more different from other old exemplars.

Since then, many pro-prototype researchers argue that prototype abstraction process is in operation only when category size (i.e. number of training exemplars) is large and/or when the transfer phase is delayed (Homa, Sterling & Trepel, 1981, …) Specifically, it was found that the classification accuracy for a new distortion increased as a function of its similarity to a particular high-level old distortion, and more importantly, that the contribution of old-new similarity to the classification performance was attenuated as the category size gets larger. It was also found that the old distortions were better classified than prototypes immediately after the learning phase, but were classified not as well as prototypes after a one-week delay. Though once considered strong evidence for a prototype abstraction process, formal modeling of the old-new similarity × category size interaction and the differential forgetting of old distortion vs. prototype revealed that both phenomena can be quantitatively fit by a pure exemplar model no worse than a more complex mixed model involving a prototype-based process (Shin & Nosofsky, 1992, Busemeyer, Dewey and Medin, 1984, Hintzman and Ludlam, 1980, …).

In a more recent study, Smith argued that even though both exemplar and prototype model perfectly predicted the ordering of classification accuracy of different levels of distortions, the steepness of the typicality gradient and the magnitude of the prototype enhancement effect were better predicted by the prototype model than the exemplar model (Smith, 2002). However, both findings can alternatively be explained by the fact that subjects overlearned the prototype and low distortions that are densely packed around the center of category during the transfer phase, when the category representation was assumed not to change in the formal modeling of a standard dot pattern paradigm. Indeed, the typicality gradient effect and the prototype effect were both magnified when the prototype and low distortions were presented with higher frequency in the transfer phase, but were both diminished when more patterns surrounding a particular high distortion were presented in the transfer phase (Zaki & Nosofsky, 2004, 2007).

A previous study by Homa et al. (2019) argued that the classification and recognition performance following a learning phase in which each old exemplar was presented once can be better accounted for by a prototype-based process whereas an exemplar-based process took place when fewer number of old exemplar was presented multiple times. The purpose of the present study is to show that a more parsimonious pure exemplar model can well predict most of their experimental results, especially in the condition where the prototype-based process was supposed to take place, and also show that a high-powered replication of their experiments produced results consistent with the qualitative predictions by a pure exemplar model.

Before conducting theoretical analyses, we will briefly review their experimental design. a) six dot-pattern prototypes were randomly generated, three of which was chosen to construct various distortions. b) For each category, low, medium and high-level distortions were constructed by displacing each dot of the corresponding prototype by a random distance. The random distances for the three levels of distortions were sampled from a zero-centered normal distribution with standard deviation of 1.2, 2.8 and 4.6 respectively. Foils were constructed as medium-level distortions from each of the three remaining prototypes. c) The learning phases were similar across three experiments. Two different learning phases were employed in two conditions. In the repeating condition (REP), the same 15 medium-level distortions (5 per category) were presented in each trial blocks. In the non-repeating condition (NREP), different 15 medium-level distortions (5 per category) were presented in each trial block. The experiment 1 contained 15 learning blocks and the experiments 2 and 3 each contained 20 learning blocks. d) During the transfer phases, subjects were instructed to classify patterns in the experiment 1 and to recognize patterns as old (presented in the learning phase) or new (not presented in the learning phase) in the experiment 2. The patterns used in the transfer phase consisted of the prototypes, low, medium and high-level distortions in the experiment 1, old, new medium-level distortions and foils in the experiment 2, old, new medium-level distortions and the prototypes in the experiment 3.

Homa et al. stated three main results that can be well predicted by their mixed model: 1) there was no difference between the speed of learning in the REP and NREP conditions. 2) old and new medium-level distortions were hardly discriminated in the NREP condition. 3) novel patterns were classified with high accuracy even in the NREP condition, with the new medium-level distortions better classified in the NREP condition than the REP condition. We intend to show that results 2 and 3 can also be well predicted by a pure exemplar model. To be fair, we adopted the same set of free parameters as in the exemplar model used in the REP condition by Homa et al.

The second result can be demonstrated by modeling recognition transfer phase. In the experiments 2 and 3, the probability with which a transfer pattern i is recognized as old can be expressed by the equation (1):

(1)

where is the background noise present at the beginning at the learning phase, and is the response-scaling factor (subjects make classification decisions exactly based on the probability matching when , but towards the category with the largest probability more deterministically when ). The criterion parameter takes the value of in the REP condition and of in the NREP condition to reflect different decision criteria. The difference between the old response probabilities of old and new distortions is essentially determined by the difference between the summed similarities of the two item types to all three categories.

To mirror the modeling approach used by Homa et al., summed similarities of an old and new distortion to stored patterns in each category can be approximated in terms of “within-category similarity” between any two medium-level distortions within the same category, and “between-category similarity” between any two medium-level distortions belonging to different categories, as in equations (6) and (7). It is assumed that pattern i belongs to category A.

(2)  
 (3-1)  
 (3-2)

where the total number of learning blocks is equal to 20 in experiments 2 and 3. Equation 3-1 represents the summed similarity from an old medium-level distortion o to all stored exemplars in category A. Its self-similarity is maximal at 1 and its similarity to 99 other old distortions is defined to be . Equation 3-2 represents the summed similarity from a new medium-level distortion n to all stored exemplars in the category A. Its similarity to all 100 old distortions is defined to be . Intuitively, tend to be much smaller on average than , so the summed similarities from either the old or new distortion to category B and C (equation 2) is negligible compared to the summed similarity to category A (equations 3) when computing the grand total. is greater than by , derived from equation (3-1) minus equation (3-2). However, the difference attributed to the single self-similarity term is swamped by all other terms shared by both item types. In other words, and are almost the same, thus the old response probabilities of the two item types are very close, implying that subjects failed to discriminate between old and new patterns.

By contrast, the old response probability is much higher for old than new distortions in the REP condition. , and stay the same as in the NREP condition. Equation 3-3 represents the summed similarity from an old medium-level distortion to all learned patterns in category A:

(3-3)

Its self-similarity of 1 is multiplied by 20 as it was repeated once in each learning block and its similarity to 80 other old distortions is defined to be . is greater than to category A by . The difference attributed to the self-similarity terms is now magnified to produce a substantial difference between and thus the old response probabilities between the two item types, implying that subjects can easily discriminate between old and new distortions in the REP condition.

Another reason causing the old response probability between the two item types to be smaller in the NREP condition than in the REP condition is different values of in the two conditions. As illustrated in figure 1, with more old distortions densely packed in the NREP condition than the REP condition, a new distortion is more likely to be surrounded by old distortions highly similar to it, so the , or the expected similarity between any two distortions, tend to be higher in the NREP condition than the REP condition. The lower in the NREP condition further diminishes the differences between and . Notably, Homa et al. failed to capture this effect even in their exemplar model, which used derived from the same medium-to-medium distortion distances in both conditions.

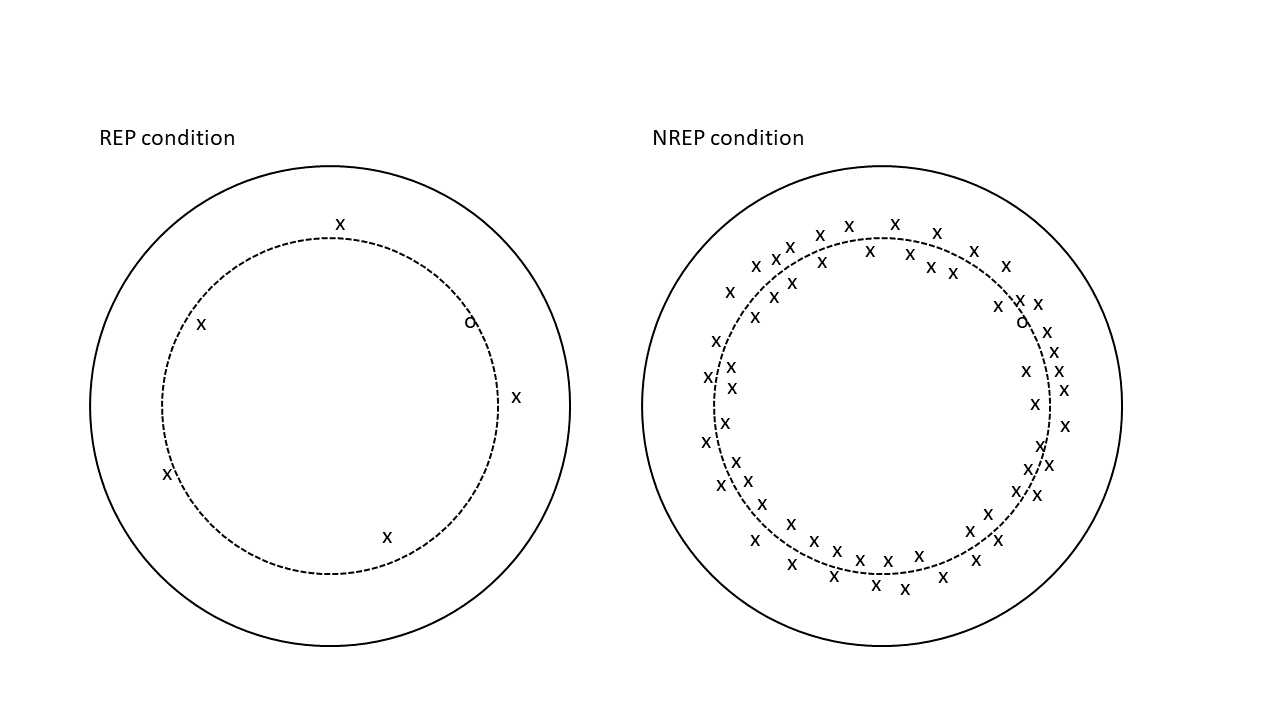


Figure 1 schematic illustration of similarity structure of stored exemplars in the REP and NREP conditions. A solid circle represents a category and a dashed circle represents the average medium-level distortion distances from its prototype. The symbol o denotes old medium-level distortions and the symbol x denotes new medium-level distortions. Smaller spatial distance between any two symbols indicates higher similarity between the two.

The third result can be demonstrated by modeling classification transfer phase. In the experiment 1, the probability with which a transfer pattern i is correctly classified as belonging to category A can be expressed by the equation (4):

In both REP and NREP conditions, the summed similarities from any novel pattern to each category can be estimated by equations (2) and (3-2). The classification accuracy is primarily determined by the proportion of the summed similarity to all three categories constituted by category A. As shown before, regardless of the item types, and are much smaller than , so the classification accuracy is generally very high. Although the summed similarity to category A can be expressed by the same equation (3-2) in both conditions, is higher in the NREP condition than the REP condition, so thus the classification accuracy is higher in the NREP condition.

To capture the subtle effect of category density on , we represent dot patterns as points in a six-dimensional feature space and compute pairwise similarities from Euclidean distances between the points. According to an MDS analysis of the similarity ratings of dot patterns (Shin & Nosofsky, 1992), six psychological dimensions can well account for variability in the perceived similarity among dot patterns. The six-coordinates of the prototypes, various distortions and foils are randomly generated in a way analogous to the statistical-distortion procedure used to generate the dot patterns. Two freely varying parameters are adopted in generating dot patterns. A “within” parameter is multiplied by distortion distances to generate various distortions from the prototype, and a “between” parameter defines the maximum distance between any two prototypes on each dimension. The psychological distance between two dot patterns i and j can be expressed by equation (5).

(5)

where represents the coordinate of pattern i on dimension m in the six-dimensional psychological space. The similarity measure is an exponential decay function of the distance, as in equation (6)

(6)

where the sensitivity parameter c reflects subjects’ overall discriminability in the psychological space.

The transfer phases of the three experiments are simulated 10000 times to obtain a reliable prediction of the mean proportions across trials and subjects. In each simulation, a single point is generated to represent each item type of the transfer patterns\*, and its similarities to old distortions are derived from equations (5) and (6). For each category, ( = 15 in the experiment 1, = 20 in the experiments 2 and 3) different training patterns were generated in the REP condition, and 5 different training patterns, each to be repeatedly counted in each training block, were generated for each category in the NREP condition. The probabilities of correct response and old recognition are computed from individual similarities by equations (4) and (1) respectively. After 10000 iterations, the predicted probabilities in each experiment by learning conditions and item types were averaged across iterations.

\* For simplicity, the transfer patterns always belong to the first category as defined in each iteration. The old distortions in the two learning conditions were separately defined as the first exemplar in the first category of the respective condition.

As can be seen from the figure 2, the exemplar model perfectly predicts the ordering of classification accuracies and old response probabilities for the different item types in each condition. The quantitative fit was also reasonably well (SSE = .009). Specifically, there was little difference between the old response probabilities of old and new distortions in the NREP condition. The classification accuracies in the experiment 1 also tend to quite high, with the accuracy of medium-level distortion higher in the NREP condition than the REP condition.

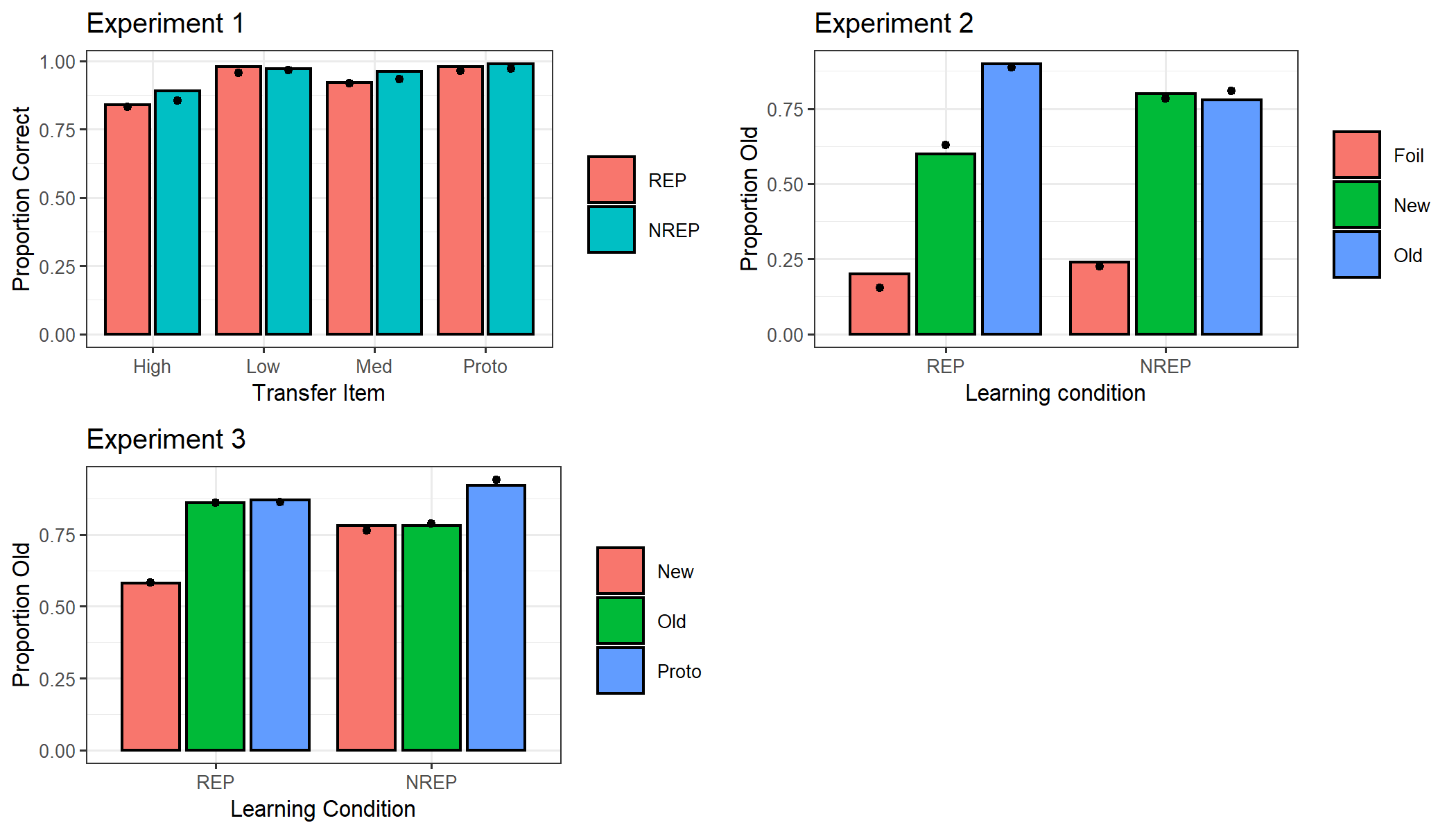


Figure 2 observed and predicted probabilities of correct classification in experiment 1 and of old responses in experiments 2 and 3 for each item types, shown for the REP and NREP conditions separately. The bars represent observed data and the dots on each bar represent predicted data.

Lastly, we acknowledge that the first result severely challenges a pure exemplar model. Indeed, Homa et al. showed that the speed of learning was predicted by a pure exemplar model to be faster in the REP condition than the NREP condition. The learning rate was demonstrated by the rate at which classification accuracy increases across learning blocks. At the beginning of the learning phase, the classification accuracies between the two conditions were very similar as they was dominated by the same background noise ; as the learning progresses across blocks, the classification accuracies became increasingly different in the two conditions as caused by the differential magnitudes of relative to (equation 8). To summarize their argument, in each learning block, the main difference between the classification accuracy for the two learning conditions lies in the difference between s in the two conditions. s for the REP and NREP conditions can be estimated by equations (11-1) and (11-2) respectively.

(11-1)

(11-2)

where B indexes the learning block and begins at B = 2, and o represents an old distortion used in the block B of the learning phase.

As the learning progresses, increased at the rate of in the REP condition as the same pattern o was presented one more time in the previous block, while increased at the rate of in the NREP condition as the pattern o were different from all five patterns presented in the previous block. Therefore, the learning rate was greater in the REP condition than the NREP condition by per block.

We intend to replicate Homa et al.’s study to confirm whether or not the speed of learning is different in the REP and NREP conditions, and whether or not there is any above-chance discrimination of old vs. new medium distortions in the NREP condition.