

Integrating Episodic Memories and Prior Knowledge
at Multiple Levels of Abstraction

Pernille Hemmer

Mark Steyvers

University of California, Irvine

Address for correspondence:

Pernille Hemmer

University of California, Irvine

Department of Cognitive Sciences

3151 Social Sciences Plaza

University of California Irvine, CA 92697-5100

E-mail: phemmer@uci.edu **Phone:** (949) 824-9486 **Fax:** (949) 824-2307

Word count: 4223 (excluding abstract, tables and legends)

Abstract

Prior knowledge can have a large influence on recall when the memory for the original event is error-prone or incomplete. We investigated the interaction between memory and prior knowledge in a recall task involving natural objects such as fruits and vegetables. We first quantified prior knowledge for the sizes of objects in a norming experiment. We then assessed the influence of prior knowledge in a memory experiment where we compared the actual size of objects shown during a study phase and the reconstructed size of an object during the test phase. Recall was biased both by the mean size of the specific object that was studied as well as the mean size of all objects in the category. This suggests that the influence of prior knowledge can come from multiple, hierarchically related levels of representation such as the object category and the superordinate category level.

Integrating Episodic Memories and Prior Knowledge at Multiple Levels of Abstraction

Reconstructing events from memory involves the coordination of multiple sources of information. Recall of past events such as the time we went to work or the size of the coffee we ordered last week, might be based partially on vague recollections of the events themselves but also on prior knowledge -- perhaps we usually go to work at a specific time and tend to order coffee of a specific size. Such knowledge might provide useful cues when recalling past events. Bartlett's (1932) seminal research on reconstructive memory initiated many research paradigms that demonstrated how cultural and social norms as well as cognitive expectations influence our recall of past events. For example, height judgments can be biased by gender (Biernat, 1993) and face recognition can be influenced by ethnicity (Rehman & Herlitz, 2006). Sometimes the expectations based on prior knowledge can lead to systematic errors and intrusions in recall. An office without books can later be remembered as having books presumably because most offices are expected to have books (Brewer & Treyns, 1981). Similarly, words can be falsely recalled from a list of words especially when the list is structured to have strong associations to a word not present on the list (Roediger & McDermott, 1995).

On the other hand, prior knowledge can also be an aid to memory. Recall for abstract information typically improves when additional meaningful information is available (Bartlett, 1932; Bower, Karlin & Dueck, 1975; Bransford & Johnson, 1972). Prior knowledge can also be advantageous in recall when the episodic memories for the original event are noisy or incomplete. Huttenlocher and colleagues presented a model of category effects in which reconstruction from memory is based on a weighted average of

episodic memory traces and prior knowledge in the form of category information (Crawford, Huttenlocher & Engebretson, 2000; Huttenlocher, Hedges & Duncan, 1991; Huttenlocher, Hedges & Vevea, 2000). By combining both episodic and category information in recall, large fluctuations in reconstructions due to noisy episodic memories can be prevented. The result of this weighted average is that reconstructions are systematically biased towards the category center. Such systematic biases have also been found in memory psychophysics experiments that reveal a compression of estimated stimulus attributes when the attributes are recalled from memory (Moyer, Bradley, Sorensen, Whiting & Mansfield, 1978; Kerst & Howard, 1978).

Previous studies (e.g., Huttenlocher et al., 1991, 2000) focused on categorization, using relatively short-term memory tasks where participants were tested immediately after stimulus exposure. With such short lags, retrieval from episodic memory is less error prone and it might be more difficult to observe the influence of prior knowledge. Previous studies have also relied on estimation of relatively artificial stimuli, such as the location of dots placed in a circle, lengths of lines, or sizes of blue and red circles (e.g., Huttenlocher et al., 1991; 2000; Sailor & Antoine, 2005). Participants are unlikely to have strong pre-experimental prior knowledge for such stimuli and instead, the prior knowledge is created during the course of the experiment by training participants on different stimulus distributions. In addition, some researchers have argued that any central bias can simply be explained by sequential effects (Sailor and Antoine, 2005). Because the smallest (largest) stimuli are always preceded by a larger (smaller) stimulus temporal contiguity can result in a central bias and no explanation in terms of prior knowledge is needed.

In this research, we investigate the interaction between prior knowledge and recall, using an experimental paradigm that differs in three important respects from previous studies. First, our memory tasks emphasize memory on a longer time-scale which we expect to lead to a decrease in the accuracy of the episodic memory representations, and a concomitant greater reliance on prior knowledge. This allows us to better investigate the influence of prior knowledge on recall, and to emphasize the memorial process. Second, unlike previous studies, we employ naturalistic stimuli such as fruits and vegetables for which we expect relatively well established pre-experimental prior knowledge. We also expect that such naturalistic stimuli are associated with more structured knowledge representations where knowledge can be described at several levels of abstraction. In our experiments, we test memory for one-dimensional stimulus attributes, namely the size of an object. Prior knowledge for such attributes might not only exist at the superordinate level (e.g. “I expect fruits to be roughly of this size”) but also at the object level (e.g. “I expect an apple to be of this size”). Therefore, the inclusion of naturalistic stimuli allows us to study the influence of prior knowledge at multiple levels of abstraction. Our experiments are further designed to evaluate systematic deviations that are a result of the influences of prior knowledge, and not an artifact of sequential dependencies.

In two experiments we investigated the interaction between prior knowledge and episodic memory. In Experiment 1, we collected norms for the sizes of natural objects such as fruits and vegetables. In this norming experiment we measured the expected sizes of common natural objects. Participants were asked to make judgments both about the mean expected size as well as the minimum and maximum expected sizes. In Experiment

2, we measured performance in a reconstructive memory task involving the normed stimuli.

Experiment 1

Participants

Eighteen undergraduate students at the University of California, Irvine participated in the experiment. They were compensated with course credit.

Materials and Procedure

We sampled 24 high-resolution color images of fruits and vegetables from a large image database (PhotoObject.net). All images were photographed against a white background. The fruit and vegetable images had an average size of 1126 x 1078 and 1175 x 878 pixels respectively. All images were presented on two computer screens. Each fruit and vegetable target image was presented on the right screen along with a slider used to manipulate the size of the object. A comparison image containing a display of a plate and utensils was presented on the left screen. Participants were asked to make the following three judgments relative to the comparison image: “What is the average size of an object like this?”, “What is the smallest size of an object like this?” and “What is the largest size of an object like this?” Participants resized the image using the slider. Responses were measured on a scale from zero to one, where zero corresponds to an object scaled to less than one pixel on the screen and one corresponds to the maximum position of the slider where the target object fills either the entire height or width of the screen. The object remained on the screen until the participant was satisfied with the current size judgment and clicked a button to continue. Participants were shown real life versions of the comparison object (a set of utensils) to ensure agreement on the common

size of the object. The images were blocked by category (fruits and vegetables) and the 24 images within each category were presented in random order. The initial size for each image was chosen at random from sizes 0.2, 0.4, 0.6 and 0.8.

Results

Figure 1 shows the mean ratings for 24 objects from the fruit and vegetable category ordered by the mean ratings for the average size¹. The ordering of sizes reflects intuitive notions participants have about fruits and vegetables. For example, in the fruit category the average size judgment for raspberries is smaller than strawberries, which is smaller than apples, all of which in turn are smaller than a pineapple. The range of the size judgments also increases with the magnitude of the objects, such that the range for the raspberry is the smallest, and the range for the pineapple is the largest. Participants expressed a reasonable degree of agreement in their judgments. For the “average” judgment, the mean pairwise rater correlation was .65 with a 90% confidence interval of .48 and .81.

Experiment 2

In this experiment, we assessed how prior knowledge for the sizes of objects influences the recalled size of a particular object studied in the context of the experiment. We tested recall for the natural objects that were normed in Experiment 1 and also a set of artificial shapes. We predicted that the effect of prior knowledge for fruits and

¹ Data from four participants was discarded because they made ‘large’ (‘small’) judgments that were smaller than their ‘small’ (‘large’) judgments indicating that they did not understand the task.

vegetables would occur at two levels of abstraction: the object category and superordinate category level. We hypothesized that recall would be biased towards the mean of the distribution associated with all size variations of the particular object (e.g., an apple) and also towards the mean of the size distribution associated with all objects within the object's superordinate category (e.g. fruits). Therefore, we predicted that an object that is studied at a relatively small/large size relative to the mean of the object would be overestimated/underestimated at test. Similarly, we predicted that an object that is studied at a relatively small/large size relative to the mean of the superordinate category distribution would be overestimated/underestimated at test. These predictions are consistent with the category effects found by Huttenlocher et al. (1991, 2000), where small things were made larger at reconstruction, and large things were made smaller at reconstruction. Figure 2 illustrates how the contribution of prior knowledge at the object and superordinate knowledge levels can be teased apart. It illustrates the predicted recall for two example objects, a strawberry and an apple, shown at the same study size. An object-level bias would lead to an underestimation for the strawberry and an overestimation for the apple towards the centers of their respective object size distributions. A superordinate-level bias would lead to an overestimation for both the apple and the strawberry. Because both of these effects might operate simultaneously it is unclear what the combined result is of the object and superordinate-level biases. However, the key prediction is the objects that are studied at the same absolute study size can be differentially biased when the objects are associated with different object-level prior knowledge. This prediction also excludes the possibility that the bias is due to sequential

effects or edge effects whereby participants are reluctant to use the edges of the response scale.

In addition to the natural objects, we also tested recall for artificial shapes that have no natural size scale (see figure 3) and presented these at the same sizes as the vegetable stimuli. The inclusion of such shapes allowed additional comparisons of the relative influences of prior knowledge for artificial and naturalistic stimuli. We expected that participants would not have any strong pre-experimental expectations about size at either the object or the superordinate level. However, it is possible for participants to gain knowledge about the overall sizes of artificial shapes during the course of the experiment. Therefore, we hypothesized that recall would be biased towards the mean size of the overall size distribution associated with artificial shapes.

Participants

There were 25 participants, undergraduate students at the University of California, Irvine. The participants were not involved in the norming study in Experiment 1. They were compensated with course credit.

Materials and Procedure

We used the 24 objects from the fruit and vegetable categories in Experiment 1. We also created a category of 24 artificial shapes shown in Figure 3. These images were created by drawing outlines of objects and filling the interior with a blue color. We employed a recall paradigm analogous to continuous recognition where study and test events are randomly interleaved. The images were blocked by category (fruits, vegetables and shapes) and each of the 24 images within each category was studied and tested three times for a total of 144 trials per block. Trials were presented in random order, but were

constrained such that all test images had been previously seen as a study image and there was always an intervening test trial between each repeated study trial. Furthermore, the study size for a given image remained constant across the three repetitions. Each study image was presented for two seconds at the center of the right screen. The same comparison image that was used in Experiment 1 remained on the left screen at all times. Each test image was presented on the right screen and participants responded by clicking on one of two buttons on the screen: “Remember” or “Don’t Remember”. After this recognition judgment, participants were then asked to scale the size of the test object using the mouse to move the slider presented on the right screen. The following question appeared with the slider: “What was the size of this object, compared to the object on the left, when you saw it at study?” Responses were measured on a scale from zero to one. Participants were given clear verbal instructions to ensure that they understood the task, and were shown real life versions of the comparison object (a set of utensils) to ensure agreement on the frame of reference.

On study trials, the size of a fruit or vegetable object was selected from the ranges shown in Figure 1. For each object, the range extended from the mean of the small judgments to the mean of the large judgments. This range was divided into 8 bins, separately for each object. Five percent of study sizes were drawn from bins 1 and 8 (smallest and largest), 10% from each of bins 2 and 7, 15% from each of bins 3 and 6, and 20% from each of bins 4 and 5 (middle bins). This procedure insured that the majority of objects were presented at a size close to the mean size of the object and few objects were shown at size close to the smallest and largest acceptable sizes. It also guaranteed that objects would never be presented outside their acceptable range. For the

artificial shapes, the study sizes were sampled from the same distributions as the vegetables. Each artificial shape was yoked to a vegetable object such that the study size was the same as the paired vegetable. On test trials, the images were shown randomly at one of four sizes: 0.2, 0.4, 0.6 and 0.8. The slider was initiated at the corresponding location.

Results

To investigate the effect of prior knowledge at the object category and superordinate category level, we measured recall error as the difference between the recalled size and studied size. We restricted the analysis to hits only (i.e., “remember” responses). For visual clarity, we divided the study objects in the vegetable and fruit category into four classes based on the study size relative to the size ranges obtained in the norming study. We named the four classes “very small”, “small”, “large”, and “very large”. These sizes do not refer to the absolute size at which an object was studied but rather the size relative to the distribution at the object level (e.g., the example in Figure 2 illustrated two objects presented at the same absolute study size but different relative study size – a very small apple and a very large strawberry). Figure 4A shows the recall error as a function of absolute study size for each superordinate category. The lines in the figure have two important aspects: slope and intercept. A negative slope indicates a bias to the center of the category range, such that small objects are overestimated (positive bias) and large objects are underestimated (negative bias). A *difference* in intercepts tests our key prediction that objects studied at the same absolute study size can be differentially biased depending on the prior knowledge at the object level. An intercept difference indicates a bias to the center of the object range, such that objects that are

presented at a size that is small relative to the object size distribution have a greater overestimation error than objects presented at a relative large size. To statistically assess these effects, a linear regression model was fitted separately at the individual subject level. The regression model contained five parameters for each category; four intercept parameters corresponding to the four relative object classes and a single slope parameter (See figure 4B for fits of this regression model). Table 1 shows the mean estimated slopes and intercepts across superordinate category and relative object size. T-tests revealed that the slope for each category was significantly smaller than zero (fruits: $t(24)=-4.714, p=0.000$, vegetables: $t(24)=-5.657, p=0.000$, shapes: $t(24)=-10.754, p=0.000$). This is consistent with a superordinate-level influence of prior knowledge. An object-level effect is revealed by the intercept differences between relative object sizes. A repeated measures ANOVA for the shape category found no effect of intercept ($F(3,72)=0.453, p=0.716$). This lack of difference between the intercepts in the shape category indicates that there is no object-level effect for the objects for which the participants have no pre-experimental knowledge. A 2(category) X 4(intercept) repeated measures ANOVA was performed for the fruit and vegetable categories. There was no significant main effect of category ($F(1,24)=3.927, p=0.059$) indicating there was no difference between the fruit and vegetable categories. There was a significant main effect of intercept ($F(3,72)=14.359, p=0.000$). Contrasts revealed that the bias for relatively very small objects was significantly larger than that of relatively small objects ($F(1,24)=4.403, p=0.037$), as was that of relatively small to relatively large objects ($F(1,24)=7.902, p=0.010$), and relatively large to relatively very large objects ($F(1,24)=6.022, p=0.022$). These differences in intercept confirm the prediction that

objects studied at the same size are recalled differently depending on the relative size of the studied object. This is consistent with an object-level influence of prior knowledge.

The shape category shows no such effect.

We also determined the relationship between recognition response (“remember” or “don’t remember”) and recall performance. In this analysis, recall error was assessed by RMSE (root mean squared error) which measures the absolute deviations between the studied and recalled sizes of the objects by subject². Larger RMSE values indicate worse performance. Figure 5 shows the RMSE by category and recognition response. A 2(remember) X 3(category) repeated measures ANOVA was performed. There was a significant main effect of remembering ($F(1,6)= 10.712, p=0.017$). Recall error was significantly lower for hits (i.e., “remember” responses) than misses (“don’t remember” responses). There was a significant main effect of category ($F(2,12)=14.572, p=0.001$). Contrasts revealed that recall error was greater for artificial shapes compared to fruits ($F(1,6)=32.155, p=0.001$) and vegetables ($F(1,6)=8.571, p=0.026$). There was no significant difference in recall error between fruits and vegetables ($F(1,6)=4.91, p=0.069$). The worst performance was observed for the misses in the artificial shapes category. In this condition, performance was similar to chance performance as measured by the mean of the artificial shapes size distribution (indicated in Figure 5 by the dotted line). Note that this distribution was by design identical to the vegetable size distribution. There was a significant interaction effect between ‘remember’ and category ($F(2,12)=4.721, p=0.031$). Even when participants indicated that they had no conscious recollection of

² $RMSE = \sqrt{\sum_i (x_i^{recall} - x_i^{study})^2 / N}$, where x^{recall} and x^{study} are the recalled and studied sizes respectively and N is the number objects in the calculation.

seeing the object at study, they performed better in both the fruit and the vegetable category compared to the artificial shape category.

General Discussion

We have shown that episodic memories are influenced by prior knowledge and that such influences appear to be hierarchically structured. Objects above the superordinate category mean (e.g., a pineapple) were reconstructed smaller and objects below the superordinate category mean (e.g. a raspberry) were reconstructed larger. At the same time, an object studied at the small end of that object range (e.g. a small apple) was overestimated while an object studied at the large end of that range (e.g. a large apple) was underestimated, independent of the absolute study size. The biases in size recall are consistent with prior knowledge exerting an influence at both the object category and the superordinate category level.

Our experimental design allowed us to rule out various alternative hypotheses about the object-level effects. Sailor and Antoine (2005) hypothesized that temporal contiguity (i.e., sequential effects) produced the center bias effects of the Huttenlocher studies (1991, 2000) and that reconstruction did not vary systematically as a function of category membership. We, however, have demonstrated that not only does reconstruction vary systematically as a function of category membership, but also that these findings can not be due to sequential effects. Sequential effects cannot be responsible for the object-level effects because we have demonstrated that objects studied at the exact same size can be recalled differently depending on the prior knowledge associated with the object. Our findings also preclude the possibility that the object-level bias is due to edge effects and slider start location. These effects would not predict differential bias between vegetables

and shapes which are studied at the same size and they are unable to explain the object level effects. Therefore, our results show that prior knowledge exerts strong effects on long term memory, and that prior knowledge corrects for noisy episodic memories at the object level. On the other hand, it is not entirely clear whether the negative slopes in Experiment 2 can be solely attributed to prior knowledge at the superordinate category level. It is possible that additional factors have contributed to this effect. For example, sequential effects can lead to a bias toward the center of the overall size distribution. Similarly, it is not clear whether participants were utilizing prior knowledge at the superordinate level (e.g. fruits) or an even more abstract level encompassing all objects encountered during the experiment.

Overall, our results are consistent with the view that the interaction between episodic memory and prior knowledge occurs at multiple levels of abstraction. The nature of combination of different sources of prior knowledge might depend on the familiarity of the object. Very specific prior knowledge associated with the object might be utilized during recall, when the participant is very familiar with the object. On the other hand, if the object is unfamiliar, prior knowledge at a higher level of abstraction (such as the superordinate level) might be engaged.

In contrast to previous studies, our experimental paradigm allowed us to disentangle recall error for objects that were or were not recognized. One striking finding was that even in the absence of recognition, a smaller reconstruction error was found for objects associated with pre-experimental prior knowledge (fruits and vegetables) compared to objects without such knowledge (shapes). This difference cannot be attributed to differences in the size distribution as the shapes and vegetables were studied

at the exact same sizes. In the absence of episodic information about a studied artificial shape, the lowest recall error (as assessed by RMSE) is achieved by responding with the mean of the superordinate category distribution observed in the experiment. Participants appeared to have followed this strategy (see Figure 4). However, in the absence of episodic information about a studied vegetable or fruit, a lower error can be achieved by using both the superordinate as well as the object size distribution to generate a response. Such knowledge about sizes of specific objects could only have been attained outside the experiment because in the experimental context, participants saw a particular object only at a single size. This supports findings by Huttenlocher et al. (1991; 2000) that prior knowledge can serve a useful role to improve average performance in recall when episodic memory is weak, incomplete, or error-prone.

Using prior knowledge at multiple levels of abstractions is an efficient strategy that allows generalizing over experiences and correction of noisy memories. Bartlett's (1932) finding that his British participants tended to recall the natives sailing in a boat rather than a canoe is in line with the predictions of a hierarchical effect that recall for an unfamiliar object, like a canoe, is influenced by a higher, more abstract level of knowledge, namely that people sail in boats. In essence participants had weak memory traces when recalling the story months or even years after initial exposure, but were able to use prior knowledge at a higher level of abstraction to aid recall.

Acknowledgments

The authors wish to thank Ken Malmberg and Scott Brown for discussions that helped shape this research.

References

- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.
- Biernat, M. 1993. Gender and height: development patterns in knowledge and use of accurate stereotypes. *Sex roles*, 29, 691-713.
- Bower, G. H., Karlin, M. B. & Dueck, A. (1975). Comprehension and memory for pictures. *Memory and cognition*, 3, 216-220.
- Brewer, W. F. & Treyens, J. C. (1981) Role of schemata in memory for places. *Cognitive Psychology*, 13, 207-230.
- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 11, 717-726.
- Crawford, E., Huttenlocher, J., & Engebretson, P. H. (2000). Category effects on estimates of stimuli: Perception or reconstruction? *Psychological Science*, 11, 280-284.
- Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in establishing spatial location. *Psychological Review*, 98, 352-376.
- Huttenlocher, J., Hedges, L. V., & Vevea, J. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General*, 129, 220-241.
- Kerst, S. M., & Howard, J. H. (1978). Memory psychophysics for visual area and length. *Memory & Cognition*, 6, 327-335.
- Moyer, R. S., Bradley, D., Sorensen, M., Whiting, J., & Mansfield, D. (1978). Psychophysical functions for perceived and remembered size. *Science*, 200, 330-332.

Rehman, J., & Herlitz, A. (2006). Higher face recognition ability in Girls: Magnified by own-sex and own-ethnicity bias. *Memory, 14*, 289-296.

Roediger, H. L. & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and cognition, 21*, 803-814.

Sailor, K. M., & Antoine, M. (2005). Is memory for stimulus magnitude Bayesian? *Memory & Cognition, 33*(5), 840-851.

Table 1. *Mean slopes and intercepts by category and relative object size for Experiment 2*

	Fruits		Vegetables		Shapes	
	Mean	SD	Mean	SD	Mean	SD
Slopes	-0.120	0.127	-0.191	0.169	-0.415	0.201
Intercepts						
<i>Relative object size</i>						
Very small	0.057	0.049	0.089	0.062	0.144	0.075
Small	0.048	0.045	0.075	0.058	0.148	0.072
Large	0.040	0.044	0.065	0.059	0.142	0.072
Very large	0.035	0.054	0.047	0.065	0.143	0.016

Note. $N=25$

Figure Caption

Figure 1. Mean rating results from Experiment 1 for the fruit (A) and vegetable (B) objects. The objects are in reading order by the mean of the “average” rating. The bar graph shows the range of size judgments for individual objects. For each bar, the mean of the “small”, “average”, and “large” sizes are indicated by the bottom line, middle line and top line respectively. The vertical lines correspond to the 25%-75% confidence interval across participants.

Figure 2. Predicted influences of prior knowledge at the object and superordinate level. The distributions illustrate the hypothetical range of sizes associated with the fruit category and two particular objects: a strawberry and an apple. If the two objects are shown at the same study size marked by “x”, the strawberry and apple are presented at a very large and very small size respectively, relative to their object distributions. This leads to object bias in opposite directions, towards the center of the respective object distributions, as shown by the light colored arrows. In contrast, the superordinate bias operates in the same direction towards the center of the fruit distribution as shown by the dark colored arrows.

Figure 3. Artificial shapes used in Experiment 2.

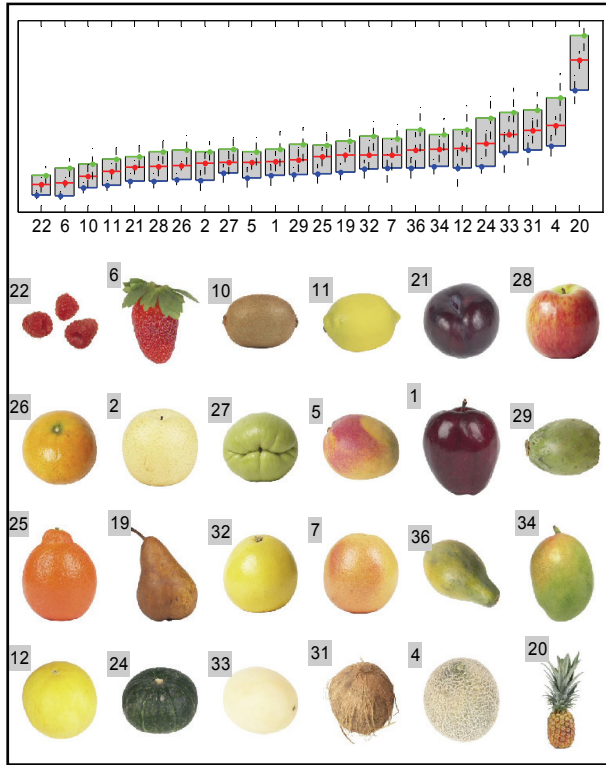
Figure 4. Recall biases as a function of absolute and relative study size. (A) Experimental data. Results are grouped by superordinate category and relative object size. (B) Linear

regression model fits. Negative slopes are consistent with biases toward the center of superordinate category range. Intercept differences are consistent with object-level knowledge effects.

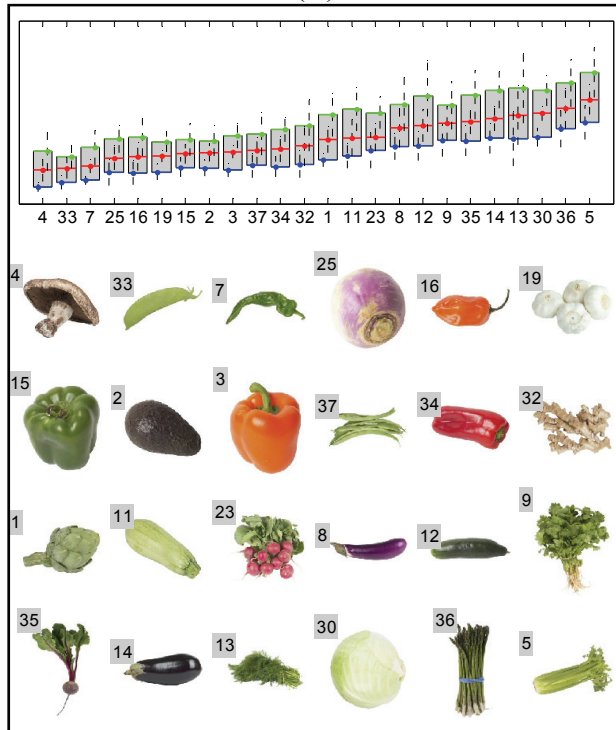
Figure 5. Recall error in Experiment 2 measured by root mean squared error (RMSE).

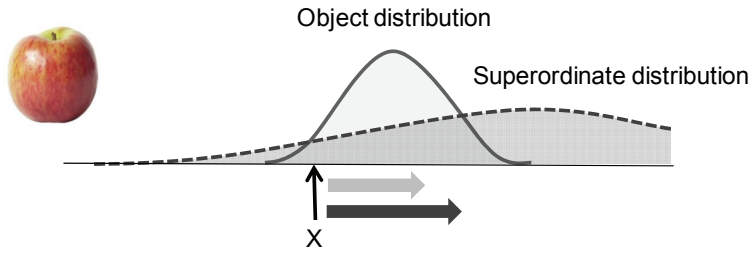
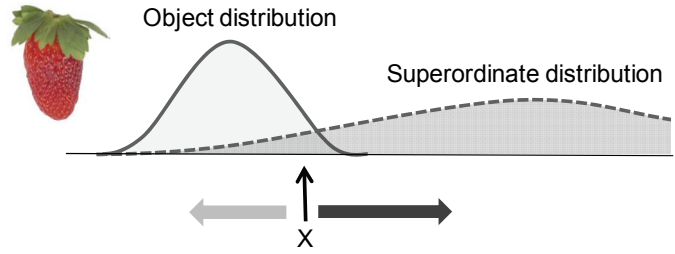
The error is grouped by superordinate category and by recognition response. The baseline error is equivalent to responses based on the mean of the vegetable, or equivalently the shapes category. Error bars indicate standard error of the mean.

(A)

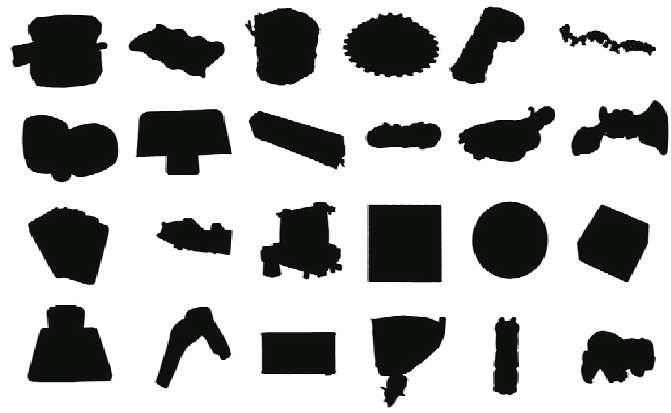


(B)

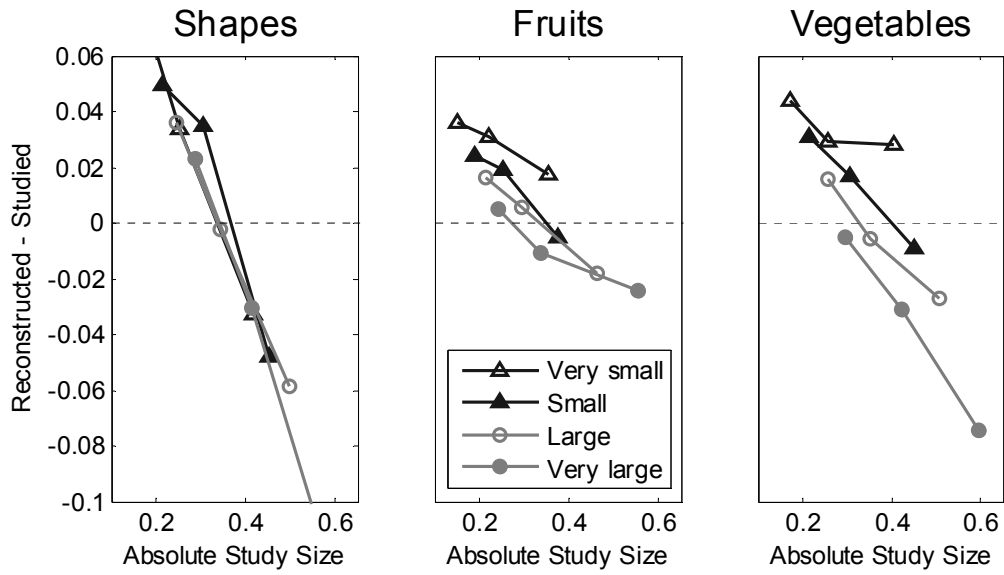




———— Absolute Study Size —————>



(A)



(B)

