9009_CH20_Gluck_LEA 3/9/07 12:17 PM Page 325

CHAPTER
20

Categorization, Recognition, and Unsupervised Learning

Evan Heit
University of California–Merced

Noellie Brockdorff
University of Malta, Msida, Malta

Koen Lamberts
University of Warwick, Coventry, England

In this chapter, we review several strands of research that share one common theme: the close conceptual and empirical links between categorization and recognition memory. We demonstrate that these links are strong, and often surprising. We explore the relation between recognition and unsupervised learning, and we investigate how perceptual processes can provide a unifying link between categorization and recognition.

In Gordon Bower's laboratory in the late 1980s and early 1990s, unsupervised learning was in the air. Several of his students were working on categorization as well as memory, and it appeared that dominant ideas in categorization research were being overturned. In particular, the traditional supervised-learning paradigm of category learning (e.g., Bruner, Goodnow, & Austin, 1956; Medin & Schaffer, 1978) was being challenged. In this paradigm, which was closely connected to paired-associate learning in the study of cued-recall memory (e.g., Bower, 1961), subjects would learn to link stimuli with category labels. On a learning trial, a subject would categorize a stimulus as, say, an A or a B, then receive corrective feedback. Although some work in the Bower lab during this period did use the supervised-learning paradigm (e.g., Gluck & Bower, 1988), his students were also asking what could be learned without feedback.

Intuitively speaking, although it is readily possible for people to be explicitly trained to associate stimuli with category labels, by no means is this procedure necessary. People can observe different kinds of houses, different kinds of landscapes, or different kinds of social interactions, and sort them into categories, even without having to label them and wait for corrective feedback from some external tutor. Experimental evidence supported the idea that people could learn categories without feedback. For example, Fried and Holyoak's (1984) subjects were able to learn to distinguish between categories of geometric designs even when category labels were not provided and in fact no information was given about the number of categories to be learned (see also Homa & Cultice, 1984). Lewicki (1986) presented a number of surprising results on the topic of nonconscious social information processing, for example, that subjects could observe personality descriptions of a set of people and successfully make predictions about additional persons, without being able to report the rules relating different personality traits.

Around this time, several models were proposed that could learn without feedback. The array model of Estes (1986) and the Minerva 2 model of Hintzman (1986) described category learning as simply storing memory traces, with or without presented labels or feedback. Connectionist models such as auto-associators (McClelland & Rumelhart, 1985) were being developed. These models learned the internal structure of stimuli without any training signal needed. Likewise, Billman and Heit (1988) presented an associationist model that generated its own internal feedback, making an external signal unnecessary. Anderson (1990) proposed a rational model of categorization, making the explicit claim that the most important function of categories is not to facilitate labeling but to enable any kind of prediction among stimulus features.During this period, two of Gordon Bower's graduate students conducted research on unsupervised learning. In experiments by John Clapper (Clapper & Bower, 1991, 1994), subjects learned about fictitious categories of insects by simply observing a series of pictures. Their sensitivity to implicit categorical structure was shown in an attribute-listing task. Over the course of learning, subjects were less likely to list attributes that correlated with each other and were predictable based on category structure, and more likely to list uncorrelated features that helped identify individual items. In experiments by Evan Heit (Heit, 1992), subjects observed featural descriptions of people and were able to make predictions about missing features for transfer items, with no category labels presented at all. These predictions were fairly elaborate, involving transitive inferences across chains of observed descriptions. Heit modified Medin and Schaffer's (1978) exemplar model of categorization to account for this unsupervised learning.

Over the past two decades, it has been increasingly accepted that category learning is not just a matter of learning from corrective feedback about explicit category labels, and the role of unsupervised learning is now emphasized (e.g., Love, Medin, & Gureckis, 2004; Markman & Ross, 2003). In general, categorization research has shown that people are able to learn about the detailed structure of what they observe, even without explicit tutoring, labeling, or feedback. Unsupervised learning allows people to adapt to complex and dynamic environments, without explicit feedback. Unsupervised learning can therefore be expected to play a role in numerous tasks that have been used in psychological studies, and accounting for this role can improve our understanding of a wide range of cognitive processes. In this chapter, we illustrate this point in the context of memory research, more specifically in terms of recognition memory.

Unsupervised Learning of Recognition List Structure

What can people learn implicitly about the structure of what they observe, when they try to remember the items they come across? Is recognition memory mainly passive storage and retrieval, or is there active, unsupervised adaptation to the structure of the information and the context in which it occurs? There is great intuitive appeal to the idea that people's response strategies can adapt to the structure of the stimulus environment. After leaving Stanford, Evan Heit dallied in the American Midwest for several years before moving to the Midlands of England, where he found colleagues Noellie Brockdorff and Koen Lamberts at the University of Warwick. As part of a larger project on categorization and recognition, they studied the issue of whether people can adapt strategically to the proportion of new versus old items on a recognition memory test. It seems logical that if the test list contains mostly new items, then people should maintain a conservative criterion for saying "old," because responding "new" will be correct most of the time. Furthermore, given the success of subjects in learning the finer details of stimulus structure in unsupervised category-learning tasks, it would seem straightforward to simply respond in accord with the proportion of new test items on a recognition test. Yet despite the inherent plausibility, there has been little empirical evidence that people can actually adapt their response criteria to the proportions of old and new items in a recognition test.As reviewed by Estes and Maddox (1995, p 1076), few recognition experiments have addressed this issue systematically. Ratcliff, Sheu, and Gronlund (1992) reported significant effects of test list composition, using a study–test recognition procedure and word stimuli. However, subjects were informed in advance of the proportion of new items before each test list, so it is not clear whether they were actually influenced by the content of the test list. Indeed, other studies have found that subjects are sensitive to misinformation about the content of test lists (Hirshman & Henzler, 1998; Strack & Foster, 1995), which implied a lack of sensitivity to list composition itself. In contrast, Estes and Maddox varied the test list without informing subjects about the proportion of new items. They used a continuous recognition procedure, with digits, letters, and words as stimuli. Surprisingly, for words, Estes and Maddox did not report significant effects of proportion of new test items (33% or 67%) on a measure of response bias, *c*. For digits and letters, proportion of new test items had an effect only when trial-by-trial feedback was provided. In the absence of this detailed feedback, subjects showed no sensitivity to composition of the test list. Therefore, previous research presents a fairly pessimistic picture of the extent to which people can adapt their responses strategically according to what is on the test list.

Although our focus here is just on adaptation to proportion of new test items, the question of whether people can vary their response criterion strategically is quite general in memory research. For example, Gillund and Shiffrin (1984) suggested that subjects have different biases for high-frequency versus low-frequency words on recognition tests, allowing them to take account that lure items that are high-frequency words might still seem quite familiar. This issue has also come up in interpreting results from the false-recognition paradigm (e.g., Gallo, Roediger, & McDermott, 2001; Miller & Wolford, 1999; Wixted & Stretch, 2000) in which people show a high false-alarm rate to nonpresented lure items that are semantically associated to a set of studied words. Here, again, the question is whether people's response biases are implicitly responding to the structure of test stimuli.

We now present two previously unpublished experiments directly addressing the issue of how people can vary their response criterion strategically, as a result of unsupervised learning in a recognition memory experiment. In Experiment 1, as in Ratcliff et al. (1992), we used a study–test recognition procedure. We varied, between groups of subjects, the proportion of new items on the test list, but we did not inform subjects in advance about the composition of the test list. We compared two kinds of stimuli, words (for which Estes and Maddox, 1995, did not find an effect of test list composition) and pictures (which we expected would lead to better overall performance). Unlike Estes and Maddox, we did not provide trial-by-trial feedback. It is possible that the continuous recognition procedure, employed by Estes and Maddox, made it particularly difficult for subjects to track the proportion of new items on the test list, because in this procedure the same test item can be both new and old depending on how many times it is tested. Therefore we predicted that even without detailed feedback, subjects' response criteria would be influenced by the composition of the test list, when using a study–test procedure. In our Experiment 2, we examined whether individual subjects could vary response criterion strategically within a test list. Other research (Hintzman, Caulton, & Curran, 1994; Rotello & Heit, 1999) has come to conflicting conclusions about whether response criterion varies within a list, particularly when the response signal technique is used, requiring subjects to respond to different stimuli at different time lags.

EXPERIMENT 1

Method

Two stimulus sets were used, pictures and words. Picture stimuli consisted of 360 color photographs, showing a variety of outdoor scenes, and looking much like postcards. Word stimuli consisted of 360 common English nouns, between five and seven letters in length.

Each subject attended one session, consisting of two old–new recognition memory tests. In each study phase, 90 stimuli were presented on the computer screen, 1 stimulus at a time, at a rate of 4 s per stimulus. Immediately following the study phase, a recognition test was given on 135 stimuli. Overall feedback about accuracy was given at the end of each test block.

The 80 subjects were randomly assigned to four different conditions, with 20 subjects in each condition. The four conditions used were: picture stimuli with 33% new test stimuli (45 new stimuli, 90 old stimuli), picture stimuli with 67% new test stimuli (90 new stimuli, 45 old stimuli), word stimuli with 33% new test stimuli (ratios same as for pictures), and word stimuli with 67% new test stimuli. The study phase in all four conditions was the same, aside from using pictures or words.

Results

The main results from Experiment 1, in terms of sensitivity (d') and bias (c), are shown in Figures 20–1 and 20–2. Note that higher values of d' indicate greater sensitivity, and higher values of c indicate more conservatism, that is, greater tendency to say “new.” Visual inspection of Figure 20–1 suggests that accuracy was high for both picture and word stimuli, with better performance for pictures. There does not seem to be much of a systematic effect of proportion of new items or list number on performance. An analysis of variance (ANOVA) supported these observations. There was a main effect of stimulus type, $F(1,76) = 12.60$, $p < .001$, $MSE = 15.53$. Most important, there was no significant effect of proportion of new items, $F(1,76) = 0.01$, $MSE = 0.01$, suggesting that overall sensitivity was comparable in the 33% new and 67% new conditions. (In general, results not meeting a .05 criterion of statistical significance are not reported.)

Next, turning to the response bias results in Figure 20–2, there are clear, systematic differences between the first and second lists. The first list does not seem to show an effect of proportion of new items. In contrast, on the second list there is a consistent pattern for pictures and words. It appears that subjects were more

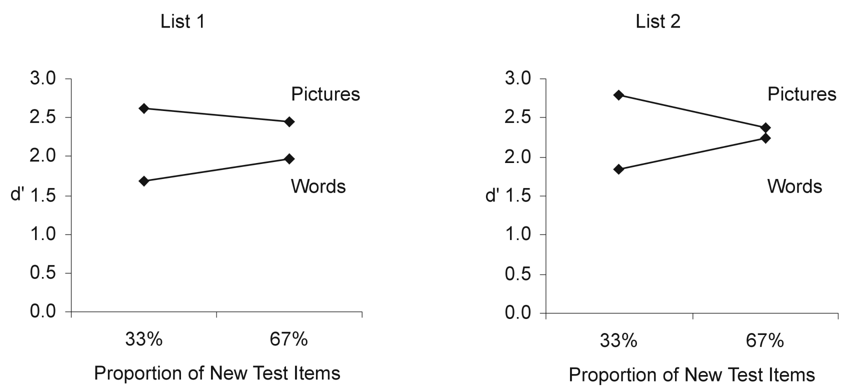

Figure 20–1. Sensitivity (*d'*) for test lists containing 33% and 67% new words, Experiment 1.

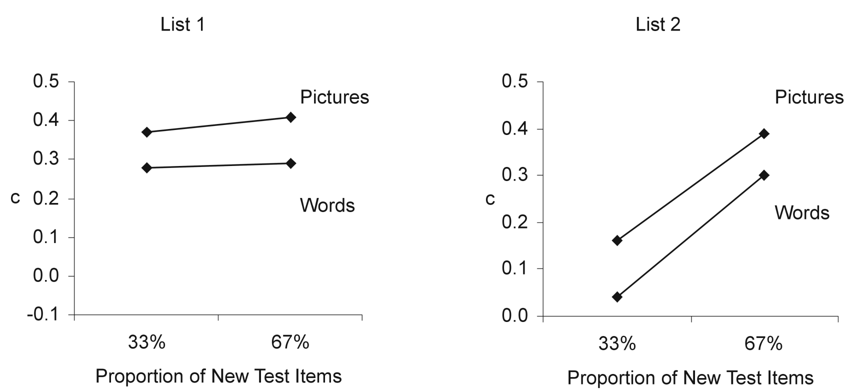

Figure 20–2. Bias (c) for test lists containing 33% and 67% new words, Experiment 1.

conservative in the 67% new conditions compared to the 33% new conditions. In an ANOVA, this was reflected by a significant list by proportion new interaction, $F(1,76) = 17.09$, $p < .001$, $MSE = 0.50$. However, the overall main effect of proportion new did not reach the level of significance. The effect of list number was significant, $F(1,76) = 18.48$, $p < .001$, $MSE = 0.54$, reflecting somewhat lower conservatism overall in the second list.

To clarify the interaction between list and proportion of new items, we conducted an additional ANOVA on response bias for just the second list. Here we found a significant main effect of proportion of new items, $F(1,76) = 9.20$, $p < .01$, $MSE = 1.20$, such that subjects were more conservative in the 67% new conditions compared to the 33% new conditions.

Discussion

In this experiment, it was clear that by the time of the second block, subjects had become sensitive to the proportion of new items on the test list, showing a more conservative response bias when there were more new items on the test list. We found these results consistently for both kinds of stimuli. These findings contrast with those of Estes and Maddox (1995) who, using a continuous recognition procedure and the same measure *c* of response bias, reported no effects of test list composition for words, and for other stimuli reported that the effect depended on the presence of detailed feedback.

Following the success of subjects in Experiment 1 in varying response criteria strategically in response to proportion of new items on a test list, we turned to a more challenging test of this ability. In particular, we looked at whether people could adopt different response criteria for different test items within a list. This is related to the issue raised by Miller and Wolford (1999), who claimed that people had different criteria for lure items versus studied items. In contrast, results by Stretch and Wixted (1998) suggested that subjects did not adopt varying criteria within a test list, for high- versus low-linguistic frequency words that were marked with different colors.

Furthermore, Rotello and Heit (1999) have raised the issue of varying response criteria in a different context, with respect to the response signal technique. With this technique, the subject is interrupted at various points in time while making a recognition judgment. The subject is instructed to respond quickly after an interrupting signal, making a judgment based on whatever assessment has been completed up until that point. The response signal comes at some point from, say, 100 ms to 2,000 ms from stimulus onset, determined randomly for each judgment trial.

Rotello and Heit (1999) reported some significant changes in response bias, as a function of time of response signal, for a subset of conditions in their own experiments and those of Hintzman and Curran (1994). The general finding in Rotello and Heit's experiments was a more conservative response bias at later times, when more information would be available to the subject. Therefore, Rotello and Heit argued that for response signal experiments, raw scores need to be interpreted with caution, because changes in false-alarm rate could reflect mere changes in response bias. Still, the empirical picture is unclear, because most other studies using this procedure have not reported bias measures. One exception was Hintzman et al. (1994), which reported no significant changes in response bias as a function of time to process test stimuli, for three experiments.

Still, null results such as those in Hintzman et al. (1994) could be due to averaging responses from subjects who each took their own strategic approaches to varying response bias. (Indeed, Hintzman et al. reported that individual subjects' response biases were correlated from one response signal to another, fitting the idea that there were systematic individual differences.) Therefore in Experiment 2, we manipulated the pattern of test stimuli to encourage different patterns of response bias in different conditions, to document people's ability to vary response bias for items tested at early time signals versus items tested at late test signals.

In Experiment 2, we compared two main configurations in which the proportion of new items tested varied as a function of processing time. In the *rising* configuration, the proportion of new items tested was low at early response signals but this proportion increased at later signals. In the *falling* configuration, the proportion of new items tested was high at early response signals, falling at later response signals. We predicted that subjects in the two groups would show different patterns of response bias over processing time, reflecting different strategies based on the distribution of old and new test items at different signals. For comparison, we also included a third configuration in which the proportion of new test items was *constant* across response signals. The test phase was organized in terms of two series of blocks, and the test proportions shifted from the first series to the second series. Therefore we investigated whether subjects could quickly adopt one response bias strategy for the first series, then take on a different strategy for the second series. (See Heit, Brockdorff, & Lamberts, 2003, for a related experiment in which there was only a single series of test blocks.)

There were four conditions in Experiment 2. In the rising/falling condition, subjects were first tested on a series of rising test blocks. For the rising test blocks, there was a low proportion of new items at early response signals and a high proportion of new items at late response signals. Then subjects in the rising/falling condition were tested on a series of falling test blocks, that is, with a high proportion of new items at early response signals and a low proportion of test items at late test signals. The second condition was rising/constant, in which a rising test series was followed by a series in which the proportion of new items was constant at different test lags. One purpose of the rising/constant condition was to see whether a response bias strategy acquired in the first series would carry over to the second series that had a more neutral character, or would subjects take a probability-matching strategy in the second, constant test series. The remaining two conditions were falling/rising and falling/constant.

EXPERIMENT 2

Method

This experiment used picture stimuli only. For each subject, the 360 stimuli were randomly assigned to six study–test blocks. Each study list consisted of 30 experimental pictures. In each study phase the pictures were presented, one at a time, at a rate of 3 s per picture. Immediately following the study phase, a recognition test was given on 30 old and 30 new pictures, using the response signal procedure. On each test trial, a cue (a cross) was shown at the center of the screen for 500 ms. The screen went blank for 100 ms and then a picture appeared. At variable time lags after the stimulus appeared on screen (100, 250, 500, or 1,250 ms) a tone sounded and the stimulus was replaced by a mask made up of small multicolored squares. Subjects were instructed to make a recognition judgment immediately after they heard the tone, and to respond as accurately as possible. If no response was made within 300 ms of the onset of the tone, or if a response was made before the onset of the tone, an appropriate error message was displayed. In addition, summary feedback about accuracy and proportion within time was given at the end of each test block.

The 32 subjects were assigned to four different conditions, with 8 subjects in each condition: rising/falling, rising/constant, falling/rising, falling/constant. The study phases in the four conditions were identical, but the four conditions differed in the test phase. Each test block in all four conditions consisted of 60 trials. In a falling test block, for the two earliest time lags the ratio of new stimuli to old stimuli was 10 new:2 old, in the middle time lag the ratio was 6 new:6 old, and in the last two time lags the ratio was 2 new:10 old. In a rising test block, for the two earliest time lags the ratio of new stimuli to old stimuli was 2 new:10 old, in the middle time lag the ratio was 6 new:6 old, and in the last two time lags the ratio was 10 new:2 old. In a constant test block, 6 new stimuli and 6 old stimuli were assigned to each response signal. Depending on the condition, each subject had three test blocks of one type followed by three test blocks of another type. For example, in the falling/rising condition, subjects faced a series of three falling test blocks followed by a series of three rising test blocks. See Table 20–1 for a summary of the conditions in this experiment.

Table 20–1
Ratio of New to Old Test Stimuli Across Time Lags And Blocks in the Four Conditions of Experiment 2

Condition	Series I (Blocks 1 to 3)					Series II (Blocks 4 to 6)				
	Time lag (ms)					Time lag (ms)				
	100	250	500	900	1,500	100	250	500	900	1,500
Falling/Rising	10:2	10:2	6:6	2:10	2:10	2:10	2:10	6:6	10:2	10:2
Falling/Constant	10:2	10:2	6:6	2:10	2:10	6:6	6:6	6:6	6:6	6:6
Rising/Falling	2:10	2:10	6:6	10:2	10:2	10:2	10:2	6:6	2:10	2:10
Rising/Constant	2:10	2:10	6:6	10:2	10:2	6:6	6:6	6:6	6:6	6:6

RESULTS

The data were trimmed to remove responses in which subjects either failed to respond within 350 ms of a response signal (8%), or made anticipatory responses, earlier than 100 ms after the response signal (2%).

The mean accuracy in each condition, in terms of *d'*, is shown in Figure 20–3. The clear results are that *d'* increased over the time course of judgment, reaching a high level of performance, and that *d'* did not vary as a function of condition (rising/falling, rising/constant, falling/rising, and falling/constant). The conclusions were supported by an ANOVA. There was only a significant main effect of time of response signal, $F(4,28) = 360.16$, $p < .001$, $MSE = .24$.

To improve the power of our response bias analyses, and more important, to improve their interpretability, we pooled the response bias scores, combining results from the first and second test signals, and the fourth and fifth test signals, for each subject. Within each of these pairs of response signals, a subject saw the same proportion of new versus old items. For example, a subject in the falling/rising condition saw five/six new items at response signals 1 and 2 in the first series, and saw one/six new items at response signals 4 and 5. We dropped the data from the third test signal, for which every subject always saw 50% new items.

The pooled bias results are presented in Figure 20–4. At the left are the response bias patterns for the first series, for all four conditions. Clearly, the falling/rising and falling/constant conditions showed falling patterns in the first series, whereas the rising/falling and rising/constant conditions showed rising patterns. Next, turning to the panels at the right, for the second series, it is clear that there is strong influence of the test configuration observed during this second series of test blocks. Subjects in the falling/rising condition showed a rising pattern in the second series, whereas subjects in the falling/constant condition continued to show a falling pattern (perhaps attenuated). In the rising/falling

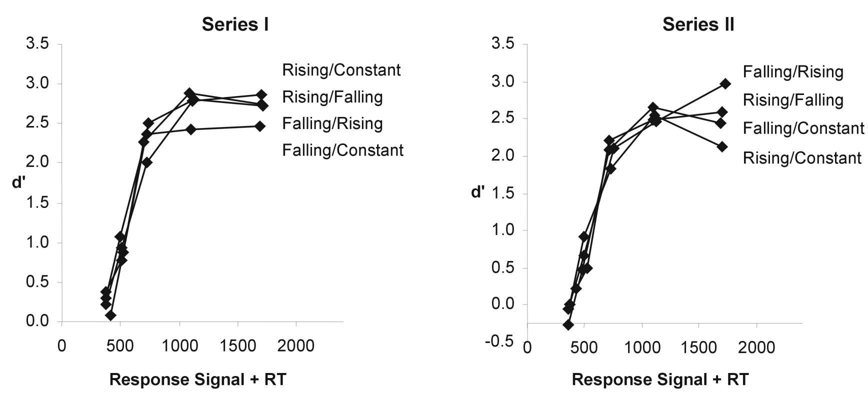

Figure 20–3. Sensitivity (*d'*) at varying response signals, Experiment 2.

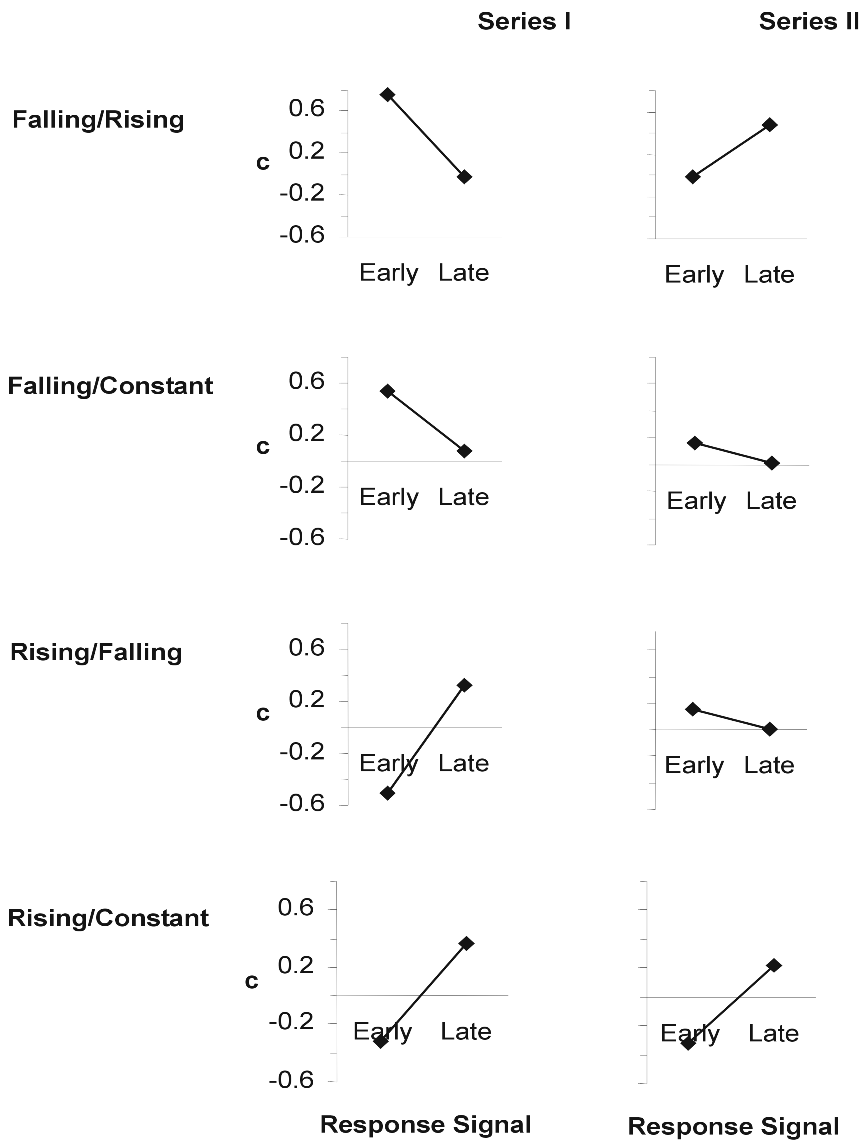

Figure 20–4. Bias (*c*) at varying response signals, Experiment 2.

condition, subjects showed a falling pattern in the second series. Finally, in the rising/constant condition, subjects continued to show a robust rising pattern in the second series.

Our analyses of the pooled response bias results targeted specific aspects of the data. We looked at each condition separately, to see when subjects truly changed their response bias patterns from the first series to the second series. That is, we were looking for series by response signal interactions. In the falling/rising condition, the interaction between series and response signal was significant, $F(1,7) = 46.80$, $p < .001$, $MSE = .07$. Likewise in the rising/falling condition, the interaction between series and response signal was significant, $F(1,7) = 28.46$, $p < .001$, $MSE = .07$. In comparison, for the falling/constant and rising/constant conditions, the interactions were not significant. Therefore, there was substantial evidence that subjects changed their pattern of response biases when the test configuration reversed from the first series to the second series, either rising to falling or falling to rising. However, there was no evidence that subjects changed their response patterns when the test configuration was constant (50% new at each time signal) in the second series, as if subjects in the falling/constant and rising/constant conditions carried over their response strategies from the first series to the second series.

Discussion

It was clear that in Experiment 2, people could not only adopt a response bias strategy quickly, in just the first three test blocks, but they could also change their strategy quickly, for the next three test blocks. Just as this experiment showed dramatic changes in the pattern of response biases from the first series to the second series in the rising/falling and falling/rising conditions, it is also notable that no significant changes were found in the rising/constant and falling/constant conditions. There was no statistical evidence that subjects in either condition changed their response strategies for the second series, even though the test configuration had changed from the first series. If subjects had simply tried to match the test configuration, then they should have shown a constant response bias pattern for the second series. Instead they showed a rising pattern in the rising/constant condition and a falling pattern in the falling/constant condition. Putting together the results from all four conditions, it appears that subjects quickly adopted a strategy for varying response bias over time, and carried on with this strategy unless it was greatly discrepant with the test configuration in the second series.

It is also interesting to compare the results on the second series for the rising/constant and falling/constant conditions. Subjects in both conditions saw the same test configuration, with a constant proportion of new items at different response signals. This configuration is also comparable to past studies using the response signal technique, in which the proportion of new test items was not manipulated across response signals. Despite seeing the same configuration of test items, subjects in the rising/constant and falling/constant conditions showed completely opposite patterns of response bias on the second series. The reason for this dramatic difference is that the two groups of subjects brought different expectations to this neutral situation, leading them to take different strategies.

These experiments demonstrate the importance of unsupervised learning in a recognition memory task. The subjects in the experiments showed exquisite sensitivity to statistical properties of the task environment, and adapted their response criterion rapidly to these properties without ever receiving explicit feedback on the outcome of these adaptive processes. The results thus provide a powerful demonstration of the ubiquity of unsupervised learning processes. Moreover, they demonstrate that a full understanding of recognition performance can be achieved only if unsupervised adaptation is taken into account.

GENERAL DISCUSSION

Strategic Changes in Response Criterion

In related research conducted at about the same time at the University of Warwick (Heit et al., 2004), we examined people's ability to vary response criterion strategically as a result of explicit instruction. Ratcliff et al. (1992) had found that response criterion had varied with explicit information about test list composition. However, more generally in recognition memory research, particularly in the false-recognition memory paradigm, there was an unanswered question of what happens when people are warned about a memory illusion—can they avoid it? For example, after studying the words *pillow, rest, slumber, nap,* and *rest,* people are likely to falsely recognize the lure item *sleep*. In the extreme, people will respond at the same level to nonpresented lures as to words that had actually been presented. But what if they are forewarned about this illusion, namely that there will be a central word, closely related to presented words but not actually studied, and that they should avoid saying that this word had been presented? Can this kind of memory error be prevented?

We answered this question using the response signal technique and word lists from previous research (Stadler, Roediger, & McDermott, 1999). The main comparison was between a standard condition, with standard instructions for a recognition memory experiment, and a forewarning condition, in which people were told about the illusion and instructed to avoid it. By using the response signal technique, we were able to pinpoint not only what effect the forewarning instructions had, but at what point in time the instructions had an effect. We assumed that responses to earlier signals, say within 500 msec of test item presentation, would reflect mainly automatic processing, and that responses to later signals increasingly would reflect strategic processing. The analyses in Heit et al. (2004) were very similar to the analyses of Experiments 1 and 2 in this chapter. We assessed sensitivity, in this case using a *d'* measure of whether people responded different to lure items versus presented items. And we assessed response bias, using the *c* measure.

The standard condition on its own revealed some interesting points about the false-memory illusion. At the earliest response signals (200 msec and 400 msec), the illusion was at its strongest, in that people were insensitive to the difference between lure items versus presented items. At later response signals (600 msec and 1,100 msec), the illusion was reduced, in that people responded differently to the two kinds of items, correctly recognizing presented items with a lower rate of falsely recognizing lure items. Just based on the standard condition, we concluded that the false-memory illusion arises due to automatic processes and is reduced somewhat as strategic processes become available.

Next, we compared the standard condition to the forewarning condition. The surprising result was that the forewarning instructions had no effect at all on the extent of the false-memory illusion! That is, forewarning did not increase sensitivity to the distinction between presented items and lures. However, forewarning did have a substantial effect on response criterion, at all response signals. In particular, people became more conservative overall—they were less likely to recognize any item as having been studied. Although the false-memory illusion was reduced in absolute terms—that is, people were less likely to say they recognized lure items—this was accompanied by a reduction in correctly recognizing presented items. Hence there was a change only in response criterion and not in sensitivity to the distinction between lure items and presented items.

Although there are some notable differences between Heit et al. (2004) and Experiments 1 and 2 in this chapter, especially in regard to whether results were due to explicit instruction or unsupervised learning, both lines of work do highlight the adaptability of response criterion and the importance of considering it when interpreting recognition performance.

Perceptual Processes in Recognition and Categorization

A central theme in our research project at the University of Warwick was the relation between perceptual processes in categorization and in recognition memory. This relation was made explicit in the FESTHER (FEature-Sampling THEory of Recognition) model proposed by Brockdorff and Lamberts (2000), derived from previous modeling work in perceptual categorization (Lamberts, 2000; Nosofsky & Palmeri, 1997). A key assumption in the FESTHER model of recognition is that recognition performance depends on the perceptual information that is available. To inform a recognition judgment, stimulus information must be perceived, which takes time. FESTHER indicates how the characteristics of feature perception affect the time course of recognition judgments.

A testable prediction from the model is that information about the time course of perception can be used in the prediction of the time course of recognition judgments. If we know how long it takes to perceive specific stimulus features, and if we understand the relation between features and recognition decisions, we can predict recognition judgments at different response latencies. Lamberts, Brockdorff, and Heit (2002) carried out four experiments, in which they first measured perceptual processing times of various features, and then used these estimates to predict old–new recognition performance with the same stimuli. The stimuli in the four experiments were drawings of various objects and scenes. Some of the stimuli had been obtained by modifying a single feature in one of the other stimuli, such that there were pairs of highly similar stimuli in the set.

The first three experiments were designed to measure the perceptual processing rates of the features that distinguished the stimuli in these similar pairs. In the first experiment, the participants carried out a perceptual same–different matching task. Two stimuli were presented simultaneously on each trial, and the participants were instructed to judge their identity as quickly as possible. Response times for "different" judgments on trials with similar stimulus pairs were taken as an index of the perception time of the critical feature that distinguished the members of the pair. In Experiments 2 and 3, the response signal technique was used with perceptual matching tasks to obtain further measurements of the processing rates of the critical features. In Experiment 4, the participants carried out an old–new recognition task using the response signal technique, with the same stimuli as in the matching experiments. The accuracy results for the different stimulus types showed strong effects of response signal interval. Most important, it proved possible to account for the data from the recognition experiment using the perceptual processing rates estimated from the matching experiments. Therefore, the results demonstrated that it was possible to predict speeded recognition performance from performance in perceptual matching, by adapting a categorization model to the recognition task.

The relation between the time course of categorization and recognition memory was also studied in another project by the same authors (Lamberts, Brockdorff, & Heit, 2003). The three experiments in this study used a new method for collecting and analyzing response times in recognition experiments. The usual tradition in categorization research is to examine performance on individual items. In contrast, in most studies of recognition, the data are obtained by averaging across stimuli. The reason for aggregation across stimuli is that each test stimulus in a recognition experiment can be presented only once to each subject (to avoid interference). However, averaging data, including response times, across items is potentially misleading, and response times to individual items can provide critical information about the underlying processes. In three experiments, Lamberts et al. used a technique that offered a compromise between the need for repeated observations for modeling and analysis, and the need to avoid interference due to repetition of individual test stimuli. Instead of repeating individual stimuli, they tested recognition using classes of structurally equivalent stimuli. The subjects repeatedly carried out a study–test procedure with pseudowords as the stimuli. All the lists used in an experiment were generated from a template that constrained the structure of the stimuli. The lists were created by mapping the template onto different sets of characters. As a result, the study and test lists in an experiment had a different appearance, but at the same time represented exactly the same underlying structure. The common structure across the lists allowed us to treat the different instantiations of a template item as repetitions of a single stimulus. There were robust and systematic differences in response times between items. Most important, these differences were in agreement with the predictions from process models of recognition memory that were derived from process models of perceptual categorization.

CONCLUSION

In the most general, and perhaps most important sense, all of our results point to the close conceptual and empirical links that exist between memory and category learning. Whereas few modern researchers would dispute the fundamental connections between processes of memory and category learning, we believe that the research areas of memory and category learning are still treated too much as separate topics; for example, there are memory phenomena and categorization phenomena, memory models and categorization models, memory workshops and categorization workshops, memory chapters in textbooks and categorization chapters in textbooks. Still, there are many positive examples of modeling memory and categorization together (e.g., Estes, 1994; Heit & Hayes, 2005; Lamberts et al., 2003; Nosofsky, 1988). In Gordon Bower's own work on category learning, models were developed that embodied learning principles that were so universal that they applied in a wide range of settings. For instance, Gluck and Bower's (1988) associative model of category learning uses learning principles, based on error correction in response to feedback, that are universal, covering a wide range of tasks and processes that involve explicit feedback. Similarly, the principles of unsupervised learning proposed by Clapper and Bower (1991, 1994) should apply to all contexts in which structured information is presented, but in which no explicit feedback is available. Indeed, unsupervised learning of detailed stimulus structure is a powerful phenomenon illustrating the importance of studying the relations between memory and category learning.

REFERENCES

Billman, D., & Heit, E. (1988). Observational learning without feedback: A simulation of an adaptive method. *Cognitive Science, 12,* 587–625.

Bower, G. H. (1961). Application of a model to paired-associate learning. *Psychometrika, 26,* 255–280.

Brockdorff, N., & Lamberts, K. (2000). A feature-sampling account of the time course of old–new recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 77–102.

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

Clapper, J. P., & Bower, G. H. (1991). Learning and applying category knowledge in unsupervised domains. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 27, pp. 65–108). San Diego, CA: Academic Press.

Clapper, J. P., & Bower, G. H. (1994). Category invention in unsupervised learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 443–460.

Estes, W. K. (1986). Array models for category learning. *Cognitive Psychology, 18,* 500–549.

Estes, W. K. (1994). *Classification and cognition*. New York: Oxford University Press.

Estes, W. K., & Maddox, W. T. (1995). Interactions of stimulus attributes, base rates, and feedback in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1075–1095.

Fried, L. S., & Holyoak, K. J. (1984). Induction of category distributions: A framework for classification learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10,* 234–257.

Gallo, D. A., Roediger, H. L., & McDermott, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review, 8,* 579–586.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–67.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117,* 227–247.

Heit, E. (1992). Categorization using chains of examples. *Cognitive Psychology, 24,* 341–380.

Heit, E., Brockdorff, N., & Lamberts, K. (2003). Adaptive changes of response criterion in recognition memory. *Psychonomic Bulletin & Review, 10,* 718–723.

Heit, E., & Hayes, B. (2005). Relations among categorization, induction, recognition, and similarity. *Journal of Experimental Psychology: General, 134,* 596–605.

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

Hintzman, D. L., Caulton, D. A., & Curran, T. (1994). Retrieval constraints and the mirror effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 275–289.

Hirshman, E., & Henzler, A. (1998). The role of decision processes in conscious recollection. *Psychological Science, 8,* 61–65.

Homa, D., & Cultice, J. (1984). Role of feedback, category size, and stimulus distortion on acquisition and utilization of ill-defined categories. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8,* 37–50.

Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review, 107,* 227–260.

Lamberts, K., Brockdorff, N., &. Heit, E. (2002). Perceptual processes in matching and recognition of complex pictures. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1176–1191.

Lamberts, K., Brockdorff, N., & Heit, E. (2003). Feature-sampling and random-walk models of individual stimulus recognition. *Journal of Experimental Psychology: General, 132,* 351–378.

Lewicki, P. (1986). *Nonconscious social information processing*. New York: Academic Press.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review, 111,* 309–332.

Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin, 129,* 592–613.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114,* 159–197.

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

Miller, M. B., & Wolford, G. L. (1999). The role of criterion shift in false memory. *Psychological Review, 106,* 398–405.
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., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104,* 266–300.
Ratcliff, R., Sheu, C.-F., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review, 99,* 518–535.
Rotello, C. M., & Heit, E. (1999). Two-process models of recognition memory: Evidence for recall-to-reject? *Journal of Memory and Language, 40,* 432–453.
Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition, 27,* 494–500.
Strack, F., & Forster, J. (1995). Reporting recollective experience: Direct access to memory systems? *Psychological Science, 6,* 352–358.
Stretch, V., & Wixted, J. T. (1998). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 137–196.
Wixted, J. T., & Stretch, V. (2000). The case against a criterion-shift account of false memory. *Psychological Review, 107,* 368–376.