

Strategic processes in false recognition memory

EVAN HEIT, NOELLIE BROCKDORFF, and KOEN LAMBERTS
University of Warwick, Coventry, England

In two experiments, the response signal technique (Reed, 1973) was combined with the DRM paradigm (e.g., McDermott & Roediger, 1998) to investigate the time course of false recognition memory—in particular, how this effect varies along the time course of generating a recognition judgment. Across the experiments, in addition to standard instructions, there were forewarning instructions encouraging the participants to avoid this effect, as well as inclusion instructions intended to enhance this tendency. It was found that the false memory effect was at its strongest at earlier response signals, diminishing when more time was given to make a recognition judgment. The forewarning instructions led to a more conservative overall response bias, rather than to a reduction of the effect. However, the participants were able to exaggerate this effect in the inclusion condition. The results are discussed in terms of the role of strategic processing in recognition memory.

In recognition memory, one of the most striking findings is people's tendency to mistakenly recognize words that have not been observed but are semantically related to observed words. Experiments in the DRM paradigm (e.g., McDermott & Roediger, 1998) have shown this effect dramatically. After people have been presented with several interrelated words, such as *door*, *glass*, *pane*, and *ledge*, they are very likely to falsely recognize a highly associated, critical lure item, such as *window*. Often, the false alarm rate for critical lure items is about as high as the hit rate for presented words. The aim of our experiments was to address the role strategic processing plays during recognition judgments in helping to avoid false recognition. By *strategic processing*, we refer to those memory processes that are relatively slow and under intentional control, corresponding roughly to the controlled processing in Schneider and Shiffrin's (1977) distinction between controlled and automatic processing.

This distinction is consistent with other theoretical accounts of memory. For example, Dodson and Schacter (2001) proposed that the distinctiveness heuristic is a general purpose, strategic process for avoiding false memory errors. The distinctiveness heuristic fits within the broader category of monitoring processes for distinguishing experienced events from nonexperienced events in the activation-monitoring framework (see Roediger, Balota, & Watson, 2001, and Roediger, Watson, McDermott, & Gallo, 2001, for a review of evidence for the activation-monitoring framework). More generally, two-process accounts of recog-

nition memory typically draw a distinction between quicker, more automatic familiarity-based processing and slower, more controlled recall-like processing (e.g., Jacoby, 1991; see Rotello & Heit, 1999, and Yonelinas, 2002, for reviews). In addition, single-process accounts of recognition (e.g., Lamberts, Brockdorff, & Heit, 2002) can assume that some featural information is processed very quickly, whereas other details of stimuli are processed later during the course of making a judgment. In Lamberts et al.'s model, the relative influence of features, once they have been processed, is determined by strategic factors. The statistical effect of strategy would be zero at the earliest possible response time, before any stimulus features have been processed, and the effect of strategy would increase at later times, as more features are processed.

We investigated strategic processing in two ways. First, in our experiments, we examined the time course of the false memory effect. Most previous experiments have simply involved self-paced recognition judgments and, therefore, have not distinguished between what happens automatically during a judgment and what occurs after longer deliberation. One exception is a study by Benjamin (2001, Experiment 2), which showed that speeded judgments led to a stronger false memory effect than did unspeeded judgments. Relatedly, Tun, Wingfield, Rosen, and Blanchard (1998) found an association between fast responses and increased levels of false recognition in a self-paced recognition task. In our own experiments, we used the response signal technique (Reed, 1973; see Heit, Brockdorff, & Lamberts, 2003, and Rotello & Heit, 1999, for reviews), in which the time allowed to generate a recognition judgment is under strict experimental control. Although we would not propose a sharp cutoff, in previous research it has been assumed that the processing taking place before 500-msec response signals would be relatively automatic, whereas later processing would increasingly reflect strategic influences.

The second way we investigated strategic processing was to give different instructions to different groups of

Part of this research was presented at the 42nd Annual Meeting of the Psychonomic Society, Orlando, in November 2001. It was supported by a grant from the Biotechnology and Biological Sciences Research Council of the United Kingdom. We are grateful to Aaron Benjamin, David Gallo, Kirsty Gillings, Doug Nelson, Henry Roediger, Caren Rotello, and Jason Watson for comments on this research. Please address correspondence to E. Heit, Department of Psychology, University of Warwick, Coventry CV4 7AL, England (e-mail: e.heit@warwick.ac.uk).

participants (as in, e.g., McDermott & Roediger, 1998). In both experiments, there were standard, or neutral, instructions, as well as instructions that forewarned the participants about the false memory effect and instructed them to try to avoid it. If strategic processing allows participants to avoid the false memory effect, it should be possible to affect the degree of the false memory effect by giving suitable instructions.

Predictions

Our working hypothesis was that false memory errors would arise mainly as a consequence of automatic processing and that the role of strategic processing would be mainly to help correct these errors. This hypothesis arises naturally from previous theoretical accounts in this area, but we emphasize that these points have not previously been articulated or tested. From the working hypothesis followed a number of predictions. First, with standard instructions, the false memory effect should be evident relatively early during the time course of making a recognition judgment. Second, in the standard condition, at later points during the course of making a judgment, strategic processing should lead to a reduction of the effect. Third, because reduction of the effect is under strategic control, giving participants forewarning instructions should lead to a reduction of the effect, particularly at later points during the course of making a judgment.

EXPERIMENT 1

Method

Participants. The paid participants were 40 students from the University of Warwick, 20 in the standard condition and 20 in the forewarning condition.

Apparatus and Stimuli. The experiment was controlled by a computer with a 17-in. (43-cm) monitor. The word lists were based on the 32 categories in the Stadler, Roediger, and McDermott (1999) norms that led to the strongest levels of false recognition. We chose nine members from each category to be used as the studied items, and there was one critical lure item per category. In some cases, the lists were adapted slightly for British usage.

Four experimental study and test lists were used per participant. Each study list contained 72 experimental stimuli and 6 filler stimuli (3 at the start and 3 at the end) to control for primacy and recency effects. The 72 experimental stimuli were nine members from each of eight categories and were presented in random order.

Each test list consisted of 40 words. Sixteen old words (2 randomly selected from each of 8 categories), 16 new (unrelated) words, and the eight nonpresented critical lures from the 8 categories. New words were words selected to be unrelated to any of the 32 categories used for the study and test lists. For each participant, the old, new, and critical lure sets of words were each split randomly into four groups of equal size, assigning them to the four response signals.

Procedure. The experimental session consisted of a practice block followed by the four experimental blocks. Each block consisted of a study phase followed immediately by a test phase. The practice block (not reported here) merely gave the participant some experience with the response signal procedure.

In the study phase, the words in the study list were presented at a rate of 3 sec per word. The participants were instructed to try to remember each word. For the test phase, the participants were in-

structed to press the *old* button if the test word was one previously seen during study and to press the *new* button if the word had not been seen. The response button mappings were counterbalanced across subjects. On each test trial, a cue was shown at the center of the screen for 500 msec. The screen went blank for 100 msec, and then a word appeared, in lowercase. At variable time lags after the stimulus appeared on the screen (200, 400, 600, or 1,100 msec), a tone sounded. The participants were instructed to respond immediately when they heard the tone and to respond as accurately as possible. If no response was made within 300 msec of the onset of the tone or if a response was made before onset, an error message was displayed. Summary feedback about accuracy and proportion of response in time was given after each test phase.

The participants in the forewarning condition were given two paragraphs of additional instructions (derived from McDermott & Roediger, 1998) just before the first experimental block, as a "hint to improve accuracy scores." They were told that groups of words in each list were associated with one common word that would not be presented. The participants were instructed to try not to mistakenly classify the nonpresented common word as old when it was presented during test. They were given an example that the word *king* could be linked to other words, such as *queen*, *crown*, and *palace*, and they would need to avoid saying that *king* was an old word if it had not actually been studied.

Results

The data were trimmed to remove late responses, made more than 350 msec after the response signal (6%), and anticipatory responses, made earlier than 100 msec after the response signal (5%). In terms of the proportion of excluded responses across response signals, there did not seem to be a systematic pattern distinguishing the conditions in either experiment.

First, we present the proportion of *old* responses to old, new, and critical lure test stimuli for the four different response signals, with mean response time added, in Figure 1. (Error bars indicate standard errors of the means.) In the standard condition, the participants were able to distinguish between old and new unrelated items even at the earliest time signal. Furthermore, discrimination between old and new items generally improved at later response signals. The proportion of *old* responses to critical lure items was always between the levels of responding to old items and new items. However, after the first response signal, *old* responses to old items tended to increase, whereas there was a fall in *old* responses to critical lures. In the forewarning condition, it appears that the forewarning instructions did indeed affect responses. The proportion of *old* responses was generally lower in the forewarning condition for all three kinds of stimuli. These results were analyzed with a three-way analysis of variance (ANOVA). There was a significant main effect of instructions [$F(1,38) = 7.87$, $MS_e = 0.14$, $p < .01$], reflecting the lower overall proportion of *old* responses for the forewarning condition. The effect of instructions did not interact with item type [$F(2,76) = 0.24$, $MS_e = 0.05$], implying that the instructions affected old, new, and critical lure items. There was a significant interaction between instructions and response signal [$F(3,114) = 4.98$, $MS_e = 0.09$, $p < .01$], suggesting a greater effect of forewarning at earlier response signals. For completeness, we note that there was a significant main ef-

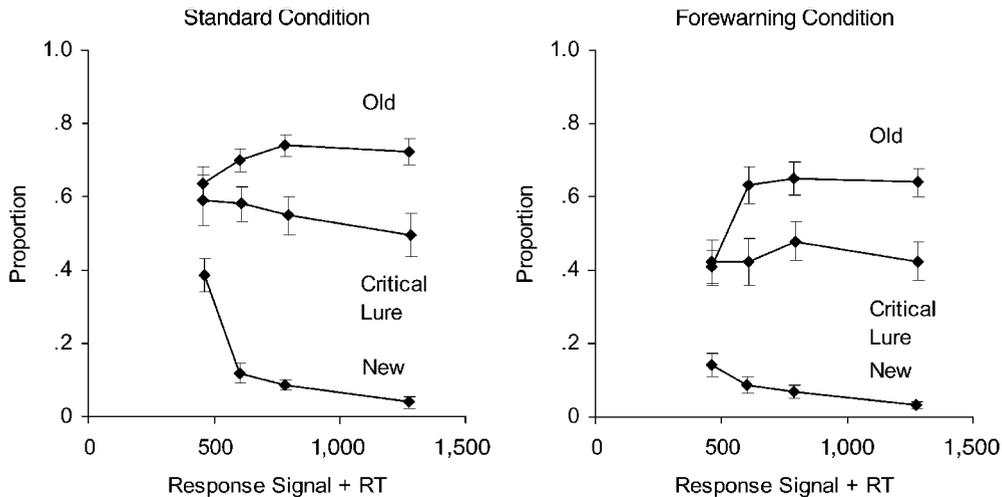


Figure 1. Proportion of *old* responses at varying response signals, with response times (RTs) added, in Experiment 1.

fect of item type [$F(2,76) = 215.54, MS_e = 0.05, p < .001$], and a significant interaction between item and response signal [$F(6,228) = 16.97, MS_e = 0.02, p < .001$]. There was no significant main effect of response signal or three-way interaction.

Raw score measures have some disadvantages, such as not being specifically aimed at a theoretical goal and not correcting for potential changes in response bias at different response signals (Heit et al., 2003; Rotello & Heit, 1999). Therefore, we next will present the key measure that is specifically targeted at the false memory effect and subtracts out effects of response bias. Because the effect is usually described in terms of responses to critical lures being nearly the same as responses to old items, we calculated difference scores between proportions of *old* responses to old items versus critical lures, in d' units.¹ (See also Gallo, Roediger, & McDermott, 2001, and Koutstaal, Schacter, & Brenner, 2001.) In Figure 2, a zero value along the y-axis indicates zero ability to distinguish lures from old items—that is, the strongest possible effect. Higher values in this figure indicate better ability to reject lures, relative to old items—that is, a weaker effect. There is practically no difference in responding to old items versus critical lures at the earliest response signal, suggesting that the effect was at its strongest level. In general, the participants' ability to distinguish between old items and critical lures was better at later response signals. There did not appear to be a systematic influence of instructions on this measure of the false memory effect. A two-way ANOVA showed a significant main effect of response signal [$F(3,114) = 5.85, MS_e = 0.34, p < .001$]. The measures were compared at different pairs of response signals (pooling across the two instructional conditions), using paired t tests (all t tests employed a Bonferroni correction to yield a family-wise error rate of 0.05). With this analysis, two pairs of response signals were

significantly different: the first versus the third and the first versus the fourth. The ANOVA did not show a main effect of instructions [$F(1,38) = 0.00, MS_e = 0.58$], and likewise, there was no significant interaction between response signal and instructions [$F(3,114) = 1.28, MS_e = 0.34$].

Finally, to complement the d' analysis, we calculated response bias on the old and critical lure items, using the c measure (Macmillan & Creelman, 1991).² These measures are presented in Figure 3. The main finding was greater conservatism in the forewarning condition. The effect of instructions was statistically significant [$F(1,38) =$

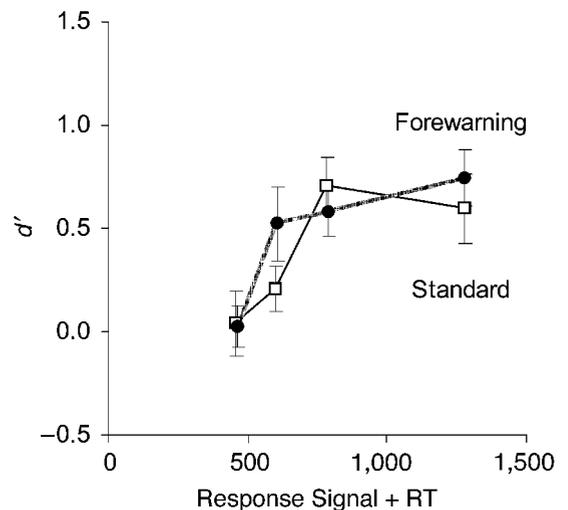


Figure 2. Discrimination between old items and critical lures (d') at varying response signals, with response times (RTs) added, in Experiment 1.

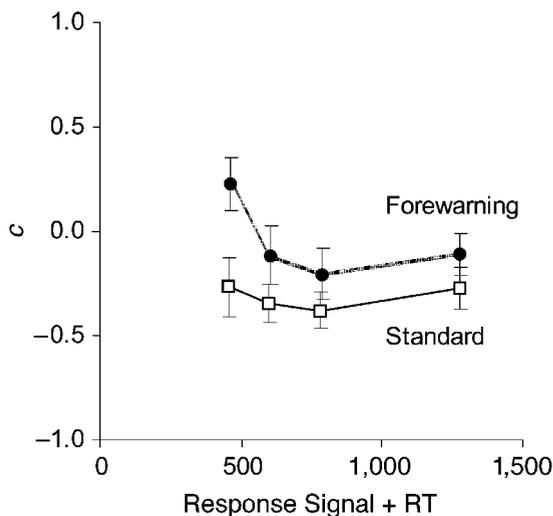


Figure 3. Overall response bias (c , or tendency to say “new”) at varying response signals, with response times (RTs) added, in Experiment 1.

8.30, $MS_e = 0.65$, $p < .01$). There was also a main effect of response signal [$F(3, 114) = 4.02$, $MS_e = 0.13$, $p < .01$], and there was no significant interaction.

Discussion

The results, in terms of the key measure of the false memory effect, gave good support for the first two predictions regarding the standard condition—namely, the false memory effect appeared early during the time course of making a recognition judgment, and it was reduced at later points in time. These results are consistent with the idea that the false memory effect arises from automatic processing and is reduced by strategic processing.

In contrast, the third prediction did not receive obvious support: The forewarning instructions did not reduce the effect in terms of ability to distinguish old items from critical lures. Forewarning merely led the participants to have a more conservative response bias. Several previous studies, in which the response signal procedure was not used, have also shown that the false memory effect remains strong after forewarning, with the warning having, at most, a moderate influence (e.g., McDermott & Roediger, 1998; see Neuschatz, Benoit, & Payne, 2003, for a review). We note that these other studies were also different in some other details, such as length of study list and use of blocked presentation of words from a single category, rather than intermixed categories, so Experiment 1 surely does not rule out the possibility that forewarning can be somewhat, if not completely, effective.

Indeed, one way to reinterpret our working hypothesis, which fits with Experiment 1 as well as with previous studies, is that strategic processing is not completely effective at reducing false recognition. We note that existing theoretical accounts of the false memory effect—namely,

the distinctiveness heuristic and the activation-monitoring framework—would assume that strategic processing takes place even in the standard condition. The lack of difference between the standard condition and the forewarning condition could thus have been due to the imperfect nature of strategic processing, so that the forewarning condition could not improve upon the standard condition.

In Experiment 2, we also compared standard and forewarning conditions, using shorter lists in case strategic processes would be more effective. In addition, there was an inclusion condition in which the participants were instructed to say “old” to the critical lures, rather than “new” (see Brainerd, Wright, Reyna, & Mojardin, 2001, for a similar technique). The fourth prediction for this condition was that the participants would be less likely to engage in strategic processing that would otherwise reduce false recognition. Hence, the reduction of the effect at later response signals should appear in the standard and forewarning conditions but should be less apparent in the inclusion condition. In general, the inclusion condition was a further examination of whether instructions can affect the results in terms of d' and, if so, at what point in time.

EXPERIMENT 2

Method

The method was like that in Experiment 1, except for the following. There were 60 participants, 20 in each of three conditions. In addition to the standard and forewarning conditions, there was an inclusion condition in which the participants were instructed to say “old” to the critical lures, rather than “new.” The instructions for the inclusion condition were in other ways modeled on the instructions for the forewarning condition; indeed, the difference in instructions was less than 10 words from a whole page of printed information.

In addition to the practice block, a total of eight experimental study and test lists were used per participant. Each study list contained 36 experimental stimuli and 6 filler stimuli. The 36 experimental stimuli in each study list were nine members from each of four categories. Each test list consisted of 20 words: 8 old words (2 randomly selected from each of four categories presented during study), 8 new words, and the 4 nonpresented critical lures from the four categories.

Results

The data were trimmed to remove late responses (7%) and anticipatory responses (12%). We attribute the somewhat higher overall proportion of excluded results, relative to Experiment 1, to the shorter test lists, which seemed to disrupt performance slightly.

First, we present the results in terms of raw scores in Figure 4. The graphs for the standard and forewarning conditions are quite like the corresponding graphs for Experiment 1. The inclusion condition had very different results, with the critical lure items obtaining more *old* responses than the old items. Although saying “old” to critical lures would be a correct response, rather than an error, in the inclusion condition, we used the same measures of response to critical lures in all three conditions.

To maintain comparability with Experiment 1, we will present an ANOVA with just the standard and forewarn-

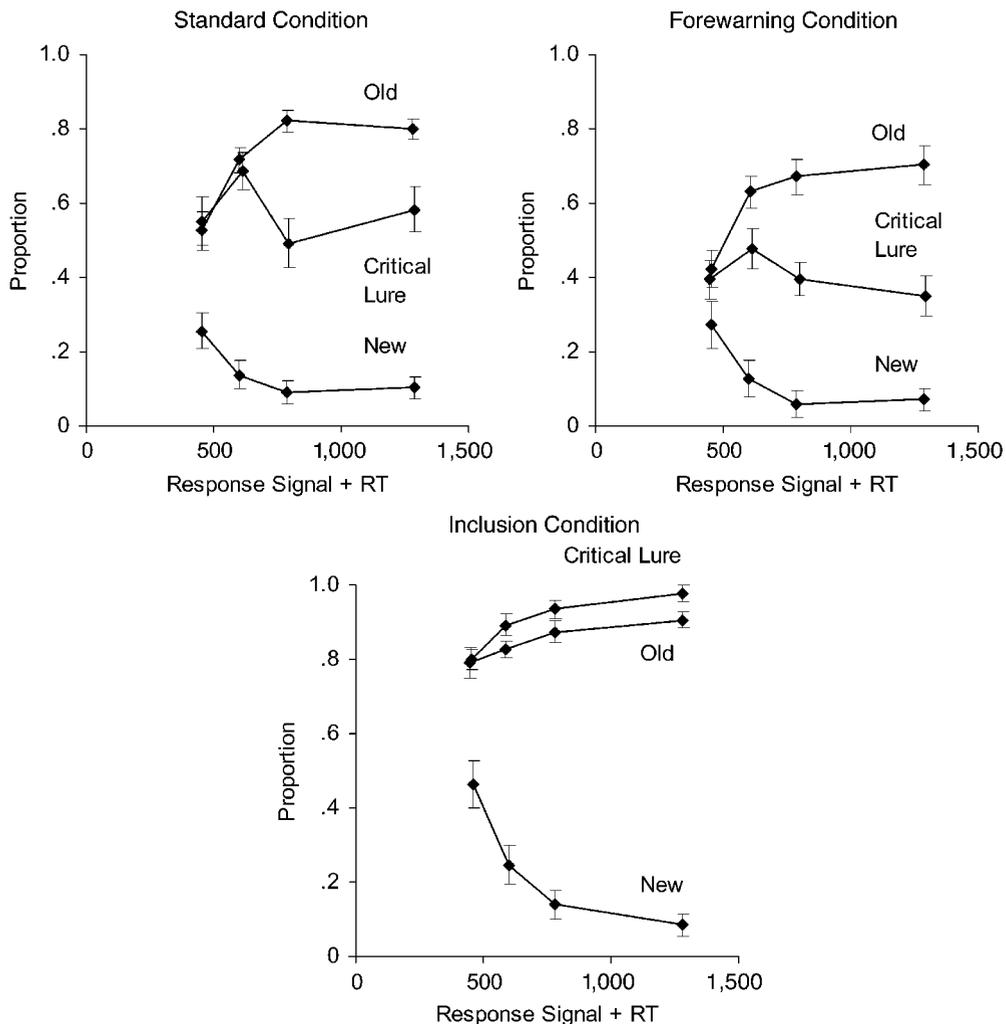


Figure 4. Proportion of *old* responses at varying response signals, with response times (RTs) added, in Experiment 2.

ing conditions. There was a significant main effect of instructions [$F(1,38) = 7.81, MS_e = 0.16, p < .01$], reflecting the lower overall proportion of *old* responses for the forewarning condition. The effect of instructions did interact with item type [$F(2,76) = 4.31, MS_e = 0.06, p < .05$]. Inspection of Figure 4 suggests that this interaction was due to a floor effect on new items, which did not have much of a drop from the standard condition to the forewarning condition. There was a significant main effect of item type [$F(2,76) = 194.36, MS_e = 0.06, p < .001$] but no significant main effect of response signal [$F(3,114) = 2.13, MS_e = 0.06$]. There was no significant interaction between instructions and response signal [$F(3,114) = 0.16, MS_e = 0.04$]. The interaction between item and response signal was significant [$F(6,228) = 24.49, MS_e = 0.02, p < .001$], but the overall three-way interaction was not significant [$F(6,228) = 0.82, MS_e = 0.02$]. Finally, we note that in the standard condition, responses to the critical lure

items showed a sawtooth pattern across response signals; however, none of the pairwise differences was statistically significant, using paired *t* tests. (Also, we note that these raw score measures did not correct for potential changes in response bias at different response signals.)

We also conducted an ANOVA on just the inclusion condition. There was a significant main effect of item type [$F(2,38) = 253.80, MS_e = 0.04, p < .001$]. All three of the pairwise differences between critical lure, old, and new items reached the level of statistical significance, using paired *t* tests. There was no significant main effect of response signal, and the interaction between item and response signal was significant [$F(6,114) = 29.60, MS_e = 0.01, p < .001$].

Next, we will present the key analysis, based on the difference between old items and critical lure items, in d' units. These results are shown in Figure 5. As in Experiment 1, there is nearly no difference in responding to old

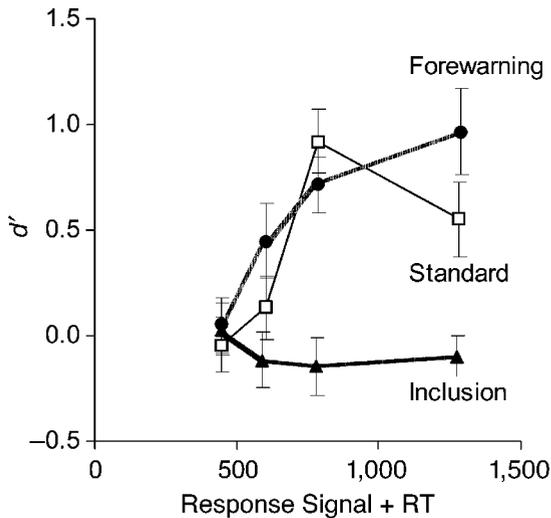


Figure 5. Discrimination between old items and critical lures (d') at varying response signals, with response times (RTs) added, in Experiment 2.

items versus critical lures at the earliest response signal, suggesting that the effect was at its strongest level at that point. For the standard and forewarning instructions, the effect was reduced at later response signals—that is, there was a greater ability to distinguish critical lures from old items. In contrast, for the inclusion condition the participants continued to treat critical lures like old items at later response signals, and indeed, the slightly negative value of the difference score reflects the higher proportion of *old* responses to the critical lures. These results were analyzed with a two-way ANOVA. There was a significant main effect of instructions [$F(2,57) = 12.11, MS_e = 0.71, p < .001$]. The measures were compared at different pairs of instructional conditions (pooling across response signals), using t tests. The standard and the forewarning conditions did not differ significantly; however, the inclusion condition differed significantly from those two conditions. There was a significant main effect of response signal [$F(3,171) = 9.80, MS_e = 0.35, p < .001$], and likewise, there was a significant interaction between condition and response signal [$F(6,171) = 5.03, MS_e = 0.35, p < .001$], which fit with what can be observed from Figure 5, that the difference between the inclusion condition and the other two conditions increased over time.³

Finally, to complement the d' analysis, we calculated response bias on the old and critical lure items, using the c measure. These measures are presented in Figure 6. The main finding is that bias to say “new” was greatest in the forewarning condition, with more tendency to say “old” in the standard condition and even more tendency to say “old” in the inclusion condition. The effect of instructions was statistically significant [$F(2,57) = 52.19, MS_e = 0.46, p < .001$]. Again using t tests, we found significant pairwise differences in response bias between all three instructional conditions. There was a main effect of response

signal [$F(3,171) = 18.25, MS_e = 0.14, p < .001$], and there was no significant interaction.

Discussion

The standard condition had results similar to those in Experiment 1, supporting the first two predictions, that the false memory effect would be strong early in the time course of making a judgment and would be reduced later. As in Experiment 1, forewarning instructions did not lead to a change in the level of the false memory effect, as compared with the standard condition; there was merely a change in response bias.⁴ The results of the inclusion condition fit the fourth prediction—namely, that the participants could avoid the strategic processing that would otherwise reduce the false memory effect. Hence, the inclusion condition showed a much stronger effect, as compared with the standard and forewarning conditions. This difference was greatest at later response times, consistent with a role of strategic processing in reducing the false memory effect in the standard and forewarning conditions.

GENERAL DISCUSSION

In empirical terms, these experiments suggest that the false memory effect is strong at early points during the time course of making a recognition judgment and weaker at later points. Unlike previous studies (Benjamin, 2001; Tun et al., 1998), this finding is not based on any data from unspeeded or self-paced judgments. Likewise, there is further evidence, fitting with previous results (e.g., McDermott & Roediger, 1998; Neuschatz et al., 2003), that the false memory effect is difficult to eliminate, even after instructions to avoid it.

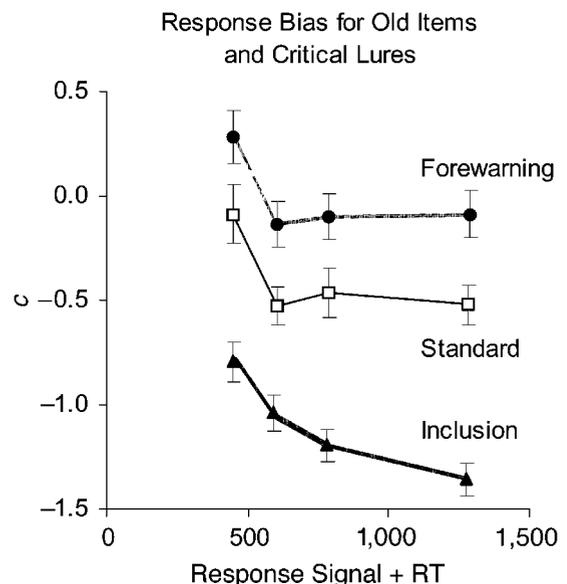


Figure 6. Overall response bias (c , or tendency to say “new”) at varying response signals, with response times (RTs) added, in Experiment 2.

In theoretical terms, our results lead to the following conclusions. On the assumption that earlier processing during recognition judgments is relatively automatic and later processing is more strategic, it appears that the false memory effect arises due mainly to automatic processing. People can reduce false recognition somewhat by applying strategic processing, most evidently at the later response signals. However, the strategic processing that can reduce false recognition is not completely effective and, indeed, was no more effective when explicit forewarning instructions were given. Finally, when given instructions not to avoid the false memory effect, people can do this strategically as well, most notably at later response signals.

REFERENCES

- BENJAMIN, A. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 941-947.
- BRAINERD, C. J., WRIGHT, R., REYNA, V. F., & MOJARDIN, A. H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 307-327.
- DODSON, C. S., & SCHACTER, D. L. (2001). "If I had said it I would have remembered it": Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review*, **8**, 155-161.
- GALLO, D. A., ROEDIGER, H. L., III, & MCDERMOTT, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review*, **8**, 579-586.
- HEIT, E., BROCKDORFF, N., & LAMBERTS, K. (2003). Adaptive changes of response criterion in recognition memory. *Psychonomic Bulletin & Review*, **10**, 718-723.
- JACOBY, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, **30**, 513-541.
- KOUTSTAAL, W., SCHACTER, D. L., & BRENNER, C. (2001). Dual task demands on gist-based false recognition of pictures in younger and older adults. *Journal of Memory & Language*, **44**, 399-426.
- LAMBERTS, K., BROCKDORFF, N., & HEIT, E. (2002). Perceptual processes in matching and recognition of complex pictures. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 1176-1191.
- MACMILLAN, N. A., & CREELMAN, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.
- MCDERMOTT, K. B., & ROEDIGER, H. L., III (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory & Language*, **39**, 508-520.
- NEUSCHATZ, J. S., BENOIT, G. E., & PAYNE, D. G. (2003). Effective warnings in the Deese/Roediger and McDermott false memory paradigm: The role of identifiability. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 35-41.
- REED, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science*, **181**, 574-576.
- ROEDIGER, H. L., III, BALOTA, D. A., & WATSON, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger, III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95-115). Washington DC: American Psychological Association.
- ROEDIGER, H. L., III, WATSON, J. M., MCDERMOTT, K. B., & GALLO, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, **8**, 385-407.
- ROTELLO, C. M., & HEIT, E. (1999). Two-process models of recognition memory: Evidence for recall-to-reject? *Journal of Memory & Language*, **40**, 432-453.
- SCHNEIDER, W., & SHIFFRIN, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, **84**, 1-66.
- SNODGRASS, J. G., & CORWIN, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.
- STADLER, M. A., ROEDIGER, H. L., III, & MCDERMOTT, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, **27**, 494-500.
- TUN, P. A., WINGFIELD, A., ROSEN, M. J., & BLANCHARD, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology & Aging*, **13**, 230-241.
- YONELINAS, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, **46**, 441-517.

NOTES

1. To avoid undefined values in the computation of d' , the hit rates and false alarm rates were adjusted by adding 0.5 to the number of *old* responses and dividing by the number of responses +1.0 (Snodgrass & Corwin, 1988).
2. Higher values of c indicate more conservatism—that is, a greater tendency to say "new" (Macmillan & Creelman, 1991, p. 33).
3. In a supplemental analysis, we conducted an ANOVA on the standard and forewarning conditions for Experiments 1 and 2, pooled together. There was no significant main effect of instructions on d' , and likewise, there was no significant interaction between response signal and instructions. Moreover, a series of paired t tests taken at each response signal did not reveal any significant differences between the standard and the forewarning conditions.
4. We also examined the effect of forewarning instructions specifically on stimuli for which Neuschatz et al. (2003) found some effect of forewarning—namely, those categories for which 70% or more of their participants could identify the central word that linked all the other words together. Even for these high-identifiability lists, there was clearly no effect of forewarning on d' . For example, in Experiment 1, the average d' was 0.33 in the standard condition and 0.28 in the forewarning condition. In Experiment 2, the average d' was 0.37 in the standard condition and 0.36 in the forewarning condition. (Note again that lower values of d' correspond to a stronger false memory effect.)

(Manuscript received November 15, 2002;
revision accepted for publication February 14, 2003.)