

Retrieval-induced forgetting in recognition is absent under time pressure

Michael F. Verde · Timothy J. Perfect

Published online: 3 August 2011
© Psychonomic Society, Inc. 2011

Abstract We examined retrieval-induced forgetting (RIF) in recognition from a dual-process perspective, which suggests that recognition depends on the outputs of a fast familiarity process and a slower recollection process. In order to determine the locus of the RIF effect, we manipulated the availability of recollection at retrieval via response deadlines. The standard RIF effect was observed in a self-paced test but was absent in a speeded test, in which judgments presumably depended on familiarity more than recollection. The findings suggested that RIF specifically affects recollection. This may be consistent with a context-specific view of retrieval inhibition.

Keywords Recognition · Memory · Retrieval-induced forgetting

Recovering a specific memory is complicated, because retrieval cues tend to simultaneously activate many memories. For example, trying to recall our most recent encounter with a dog may activate memories of many encounters with dogs and other animals. According to retrieval inhibition theory, the memory system deals with this complication by suppressing unwanted memories, rendering them temporarily inaccessible. Inhibition during the act of retrieval has been proposed as one of the primary mechanisms of forgetting (Anderson, 2003).

Retrieval inhibition theory was developed in studies of recall using the retrieval-induced forgetting (RIF) paradigm. In a typical experiment, the study materials consist of category–exemplar pairs (e.g., *FRUIT–pear*, *FRUIT–apple*,

TOOL–hammer, *TOOL–wrench*, etc.). The study phase is followed by a memory test in which some of the categories are given retrieval practice, meaning that some category members are presented as targets to be recalled (e.g., *FRUIT–pe__?*). The practiced items are designated *RP+* items. Sometime later, a final test asks for the recall of all studied items. The typical finding in the final test is that nonpracticed items from practiced categories (*apple*), designated *RP–* items, are less likely to be recalled than items from nonpracticed categories (*wrench*), designated *NP* items. According to the inhibition account, retrieving *RP+* items during the retrieval practice phase causes the suppression of the nonpracticed but categorically related *RP–* items, making them more difficult to retrieve.

The paradigm has since been extended to a variety of other memory tasks in order to gain insight into the nature of the inhibitory mechanism. Theories of forgetting have often focused on the role of interference among memories associated with a common retrieval cue. In the example above, *FRUIT* is associated with both *pear* and *apple*, and difficulty recalling one of the associates might arise because both become activated by the cue and compete for retrieval. Proponents of inhibition theory have argued that, contrary to this notion, RIF is not tied to any specific retrieval cue but is a form of general inhibition that renders a memory representation inaccessible in any context. This *cue independence* hypothesis has become one of the primary means of differentiating the inhibition account from explanations based on competitor interference (Anderson, 2003; Anderson & Levy, 2007). It is often suggested that an important test of this hypothesis is whether RIF appears in a task such as recognition, where the retrieval cues differ greatly from those present during retrieval practice involving cued recall.

The empirical findings with recognition have been somewhat inconsistent. In several studies, retrieval practice

M. F. Verde (✉) · T. J. Perfect
School of Psychology, University of Plymouth,
Plymouth PL4 8AA, England
e-mail: michael.verde@plymouth.ac.uk

has led to decreased recognition accuracy for unpracticed RP– items from practiced categories, as compared to NP items from unpracticed categories (Gómez-Ariza, Lechuga, Pelegrina & Bajo 2005; Hicks & Starns, 2004; Racsmány, Conway, Garab & Nagymáté 2008; Saunders & MacLeod, 2002; Spitzer & Bäuml, 2007, 2009). Other studies have failed to observe an effect on accuracy, but found instead that response latencies were slower for RP– items (Racsmány et al. 2008; Veling & van Knippenberg, 2004). Finally, some studies have failed to find any effect of RIF on recognition (Koutstaal, Schacter, Johnson & Galluccio 1999) or have found that the effect depended in part on the encoding conditions (Verde, 2004).

Despite these inconsistencies, the majority of studies have produced evidence of RIF in recognition in one form or another. However, we argue that the implications for cue independence are not straightforward. According to dual-process theories, recognition judgments are based on the output of two distinct processes, familiarity and recollection (Yonelinas, 2002). *Familiarity* is a relatively fast process that produces a context-free sense of “oldness.” *Recollection* is a slower search for specific episodic details and context. If both processes access a common memory representation, cue independence predicts that inhibition of the representation should have a negative impact on both processes. Unfortunately, this cannot be determined by examining overall recognition performance, where any change could be due to an effect located in only one of the component processes. Two studies have considered the dual-process perspective in detail.

Verde (2004), noting the conceptual similarities between recall and recollection, hypothesized that RIF should be observed in recollection and might be isolated to that component of recognition. Limiting the role of recollection, such as with a manipulation of study duration, should therefore reduce RIF. In an associative recognition task, RIF was observed when word pairs were studied slowly, but not when the pairs were studied much more quickly. The lack of RIF when judgments would have depended to a greater extent on familiarity suggested that familiarity may be relatively immune to RIF.

Spitzer and Bäuml (2007) argued against the viability of a dual-process account on the basis of tests of two formal dual-process models: Yonelinas’ (1994) high-threshold signal detection (HTSD) model and Rotello, Macmillan and Reeder’s (2004) STREAK model. The best-fitting parameters for both models pointed to an influence of RIF on familiarity but not on recollection. This conclusion is puzzling, given that no theory predicts such a pattern. Inhibition theory, for example, predicts a negative effect on both processes if cue independence is assumed. Furthermore, observations of RIF in source recognition (Hicks & Starns, 2004; Spitzer & Bäuml, 2009) contradict the notion

that RIF is absent in recollection. These points, while problematic for the HTSD and STREAK models, are not sufficient to rule out a dual-process account altogether, because these two models do not represent the range of views in the dual-process literature. The high-threshold nature of HTSD, for example, has been widely criticized (e.g., Wixted, 2007). STREAK, was developed to explain subjective states of experience, and its ability to model retrieval processes has never been rigorously tested.

As additional evidence against a dual-process account, Spitzer and Bäuml (2007) noted that neither HTSD nor STREAK fit their data as well as a unidimensional signal detection (USD) model, which they referred to as a “single-process” model. The USD model is associated historically with single-process strength theories of recognition, and it has often been used to test strength models against specific dual-process models. However, the implication that the USD model is limited in application to single-process theories conflates the decision variable with underlying retrieval processes. The USD model assumes that the recognition decision is based on a unidimensional variable that represents the strength of memory evidence. This decision variable, although unidimensional, could be based on multiple sources of information (Swets, Tanner & Birdsall 1961). It could, for example, summarize the output from the recollection and familiarity processes (e.g., Wixted & Mickes, 2010).

We argue that on theoretical grounds, the findings of Spitzer and Bäuml (2007) do not constitute strong evidence against a dual-process account of RIF. In fact, the present study offers new evidence to support the viability of such an account. Following a standard RIF design, participants studied exemplars from a variety of semantic categories, some of which were later given retrieval practice. During a final recognition test, the availability of recollection was manipulated with a response deadline: One group made judgments at their own pace, while another was given only 750 ms to respond. Speeded judgments are commonly used to limit the availability of recollection (Jacoby, Jones & Dolan 1998), and we expected a greater reliance on familiarity in the speeded test. If inhibition affects both recollection and familiarity, as cue independence predicts, RIF should be observed in both conditions. On the other hand, if RIF is specific to recollection, it should be absent or greatly attenuated in the speeded condition.

Method

Participants

A group of 60 undergraduates from the University of Plymouth participated for course credit. These participants

were randomly assigned in equal numbers to two test conditions.

Materials and design

The stimuli consisted of 12 exemplars from each of eight critical semantic categories (metals, animals, colors, body parts, fruits, weapons, professions, and instruments) selected from the Battig and Montague (1969) category norms. Exemplars within a category differed in their initial two letters. Half of the exemplars from each category were presented in the study list, and half were reserved as lures for the final recognition test. For the four *practiced* categories, half of the studied exemplars appeared in both the retrieval practice phase and the recognition test (RP+ items). The other half appeared only in the recognition test (RP– items). For the remaining, *unpracticed* categories, studied exemplars appeared only in the recognition test (NP items). Three additional categories (gemstones, clothing, and tools) provided filler items.

The 60-item study list consisted of six exemplars from each critical category, with an additional six from the filler categories added to the beginning and end of the list. The 36-item retrieval practice list consisted of two halves, each containing 3 RP+ items from each practiced category, with 3 filler words added before and after the critical items. Each RP+ was tested twice during practice. The 108-item recognition list consisted of the studied items and an equal number of lures, with an additional 12 filler items at the beginning of the list to serve as practice trials.

Procedure

At the beginning of the session, participants were seated at individual computers and told that they would see a list of words belonging to various categories that they would have to remember for an upcoming memory test. The words of the study list were then presented on the screen, each for 1,750 ms with a 500-ms ISI. The retrieval practice phase followed study. Each cued recall probe consisted of the category label and the initial two letters of the target (*FRUIT-ap_?*). Participants typed the appropriate word from the study list, entering “no” if they could not remember the word.

Following retrieval practice, the final recognition test began with 12 practice trials, and then the remaining critical trials. Each trial began with a fixation line shown in the center of the screen for 500 ms, which was replaced with a probe word. Participants pressed the “Z” or “M” key to indicate whether the probe was a “new” word or an “old” word from the study list. Participants in the self-paced condition were instructed to respond accurately. Those in the speeded condition were instructed to always respond quickly enough to meet the response deadline. If they failed

to press a response key by the 750-ms deadline, the warning “Too Slow!” was shown for 2,000 ms. Trials were separated by 2,000-ms intervals.

Results

Retrieval practice

During retrieval practice, the proportions of items correctly recalled did not differ between the self-paced ($M = .87$, $SE = .02$) and speeded ($M = .88$, $SE = .02$) test conditions, $t(56) = 0.21$, $p = .839$.

Self-paced recognition

Our initial analysis of recognition performance consisted of pair-wise comparisons of hit and false alarm rates between the item types (Table 1). The hit rate was significantly higher for items given retrieval practice (RP+ items) than for either the unpracticed items from the same categories (RP– items), $t(29) = 12.24$, $p < .001$, $d = 2.92$, or the items from unpracticed categories (NP items), $t(29) = 8.68$, $p < .001$, $d = 2.20$. The hit rate advantage for RP+ items reflected the fact that earlier recall facilitated subsequent recognition. Of critical interest, the hit rate was lower for RP– items than for NP items, $t(29) = 2.13$, $p = .042$, $d = 0.35$. Earlier recall of other items from the same category had a negative impact on recognition of RP– items. There was no difference in the false alarm rates to new items from the practiced and unpracticed categories, $t(29) < 0.01$, $p = .999$. These similar false alarm rates suggest that response bias was consistent across categories.

Accuracy as measured by d' (Table 2) mirrored the differences observed in the hit rates. Accuracy was higher for the RP+ items than for both the RP– items, $t(29) = 13.21$, $p < .001$, $d = 2.24$, and the NP items, $t(29) = 10.09$, $p < .001$, $d = 1.84$, reflecting the facilitative effect of prior retrieval. Critically, there was evidence of RIF: Accuracy was lower for the RP– items than for the NP items, $t(29) = 2.56$, $p = .016$, $d = 0.30$.

Finally, we examined response latencies on trials in which studied items were accurately recognized (Table 3). Correct responses to RP– items were significantly slower than those for either RP+ items, $t(29) = 3.71$, $p = .001$, $d = 0.64$, or NP items, $t(29) = 2.77$, $p = .01$, $d = 0.17$. Correct responses were somewhat faster to RP+ items than to NP items, but the difference was not significant, $t(29) = 1.35$, $p = .19$.

Speeded recognition

Two of the participants were dropped from the analysis, one for using incorrect response keys and another for failing to meet the response deadline on the majority (78%) of trials. On

Table 1 Recognition hit and false alarm rates

Condition	Hits						False Alarms			
	RP+		RP–		NP		RP		NP	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Self-paced	.93	.01	.57	.03	.63	.03	.18	.03	.18	.03
Speeded	.76	.03	.49	.04	.47	.03	.26	.03	.29	.03

average, the remaining participants responded too slowly on 11% of trials. In order to avoid item selection effects, our analysis included data from all trials, regardless of response speed. (An analysis based only on trials in which the response deadline was met produced identical results.)

Our initial analysis of recognition performance consisted of pair-wise comparisons of hit and false alarm rates between the item types (Table 1). The hit rate was significantly higher for items given retrieval practice (RP+ items) than for either the unpracticed items from the same categories (RP– items), $t(27) = 5.47, p < .001, d = 1.42$, or the items from unpracticed categories (NP items), $t(27) = 7.65, p < .001, d = 1.75$. The hit rate advantage for RP+ items reflected the facilitative effect of earlier recall. Of critical interest, the hit rates for the RP– and the NP items did not differ significantly, $t(27) = 0.88, p = .385$. The false alarm rates for new items from practiced and unpracticed categories did not differ significantly, $t(27) = 1.53, p = .137$.

Accuracy (Table 2) mirrored the differences observed in the hit rates. Accuracy was higher for the RP+ items than for both the RP– items, $t(27) = 5.33, p < .001, d = 1.13$, and the NP items, $t(27) = 6.88, p < .001, d = 1.36$, reflecting the facilitative effect of prior retrieval. Critically, there was no evidence of RIF: Accuracy did not differ significantly for the RP– and the NP items, $t(27) = 1.43, p = .164$.

Finally, we examined response latencies on trials in which studied items were accurately recognized (Table 3). Hit latencies were nearly twice as fast as those in the self-paced condition. Unlike in the self-paced condition, item type had no effect on response latencies. Correct response latencies for RP– items did not differ significantly from those for either

RP+ items, $t(29) = 0.58, p = .57$, or NP items, $t(29) = 1.53, p = .14$. Correct response latencies for RP+ and NP items also did not differ significantly, $t(29) = 0.07, p = .95$.

RIF effect

The difference between NP and RP– d' values is an index of the RIF effect. Consistent with the dual-process prediction, RIF was larger in the self-paced (difference = 0.16) than in the speeded (difference = -0.14) condition, $t(56) = 2.61, p = .012$. When a manipulation decreases accuracy, any effect becomes difficult to detect when accuracy reaches floor. Mean accuracy in the speeded condition was not close to floor, but nevertheless one might ask whether the size of the RIF effect was simply a function of accuracy. If this were the case, in the speeded condition there should be a positive relationship between general accuracy (defined as d' collapsed over NP and RP– item types) and RIF. A median split on general accuracy showed that the size of RIF (difference = -0.14) was the same whether participants were less ($d' = 0.25$) or more ($d' = 0.87$) accurate. The fact that even high-performing participants showed no evidence of RIF (indeed, there was a trend in the opposite direction) offers little support for the notion that the difference between conditions was due solely to a decline in general accuracy.

It is worth noting that while there was no evidence of a positive RIF effect in the speeded condition (d' was not significantly lower for RP– than for NP items), a small effect is not necessarily incompatible with a dual-process interpretation. After all, recollection might occur on some trials despite the time pressure. This possibility in some samples is reflected in the 95% confidence interval for the

Table 2 Recognition accuracy

Condition	Recognition (Accuracy: d')					
	RP+		RP–		NP	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Self-paced	2.27	0.10	1.08	0.09	1.24	0.10
Speeded	1.39	0.15	0.63	0.09	0.49	0.09

RP+ and RP– d' is based on RP false alarms. NP d' is based on NP false alarms

Table 3 Recognition latency (hits)

Condition	Latency (ms)					
	RP+		RP–		NP	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Self-paced	1,099	77	1,400	95	1,171	59
Speeded	592	31	575	21	594	20

difference index $[-0.34$ to $0.05]$. It would fall on alternative accounts to explain the small size of such effects as compared to the robust ones observed in the self-paced condition and in other recognition studies, as well as the apparent lack of a relationship between general accuracy and RIF in the speeded condition.

Discussion

Category exemplars were recognized more slowly and less accurately following recall of other items from the same category. However, this was true only when recognition judgments were self-paced. There was no evidence of RIF in a speeded test, in which short response deadlines limited the ability to rely on recollection. Both the presence and absence of RIF following a retrieval manipulation mirrored the pattern observed by Verde (2004), who showed that manipulating the availability of recollection with an encoding manipulation can also result in the presence or absence of RIF in recognition. Establishing the viability of a dual-process account in which RIF is specific to recollection does not necessarily rule out alternative, single-process accounts (although the less-than-straightforward relationship between the size of the RIF effect and recognition accuracy may pose a challenge for simple models). Rather, the importance of doing so lies in the implications for the way that recognition data are interpreted with respect to inhibition theory.

According to the cue independence hypothesis, RIF is the product of an inhibitory mechanism that makes a memory representation generally inaccessible. RIF should therefore be observed regardless of the task used to access the representation. The present findings are inconsistent with this prediction. An inhibition account might accommodate the findings by taking a more nuanced approach to representation. For example, one can distinguish between an item's semantic and perceptual representation and its episodic and contextual representation. If retrieval practice selectively inhibits the latter, this could adversely affect recollection but have little effect on context-free familiarity. This is similar to an account offered by Racsmany and Conway (2006), who observed RIF in an implicit memory task only under certain conditions. Participants studied category–exemplar pairs (e.g., *fruit–plum*, *fruit–orange*) and practiced retrieving some of the exemplars (*fruit–pl*). Lexical decisions for nonpracticed exemplars were slowed when primed with a studied category (*fruit–orange*) but not when primed with a novel category (*food–orange*). RIF appeared only with reinstatement of the original context, contrary to cue independence, but consistent with context-specific forgetting resulting from inhibition of the contextual representation or from inhibitory connections between the contextual representation and the item's semantic and perceptual representation.

The case for context-independent inhibition rests on demonstrations that RIF persists even when retrieval cues in the final test differ from those present during retrieval practice. RIF is often observed when a final test uses novel recall cues or a different retrieval task altogether. However, the assumption is that people only attend to explicitly provided cues. There is reason to believe that people may often spontaneously reinstate the original cues in order to aid retrieval during the final test (Camp, Pecher, Schmidt & Zeelenberg 2009). This possibility makes it difficult to rule out a role for the inhibitory context in producing RIF.

Moreover, several findings are consistent only with a context-specific view of RIF. Camp, Pecher and Schmidt (2005) observed RIF in an implicit memory task only in the subset of participants who were conscious of seeing the test items earlier in the experiment. Racsmany and Conway (2006) and Camp, Pecher and Schmidt (2007) observed RIF only when the final test cues matched those used earlier during study and retrieval practice. Perfect et al. (2004) presented items twice, once in a context where retrieval practice took place and once in a context where it did not. They observed RIF when retrieval cues were taken from the first context, but not when they were taken from the second. If recollection is a form of contextual memory, the present study is similar to these other studies in illustrating what Perfect et al. referred to as “transfer-appropriate forgetting”: inhibition only upon reexposure to the original context in which inhibition took place.

References

- Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language*, *49*, 415–445. doi:10.1016/j.jml.2003.08.006
- Anderson, M.C., & Levy, B.J. (2007). Theoretical issues in inhibition: Insights from research on human memory. In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in cognition* (pp.81–102). Washington, DC: American Psychological Association. doi:10.1037/11587-005
- Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, *80* (3, Pt.2), 1–46. doi:10.1037/h0027577
- Camp, G., Pecher, D., & Schmidt, H. G. (2005). Retrieval-induced forgetting in implicit memory tests: The role of test awareness. *Psychonomic Bulletin & Review*, *12*, 490–494. doi:10.3758/BF03193793
- Camp, G., Pecher, D., & Schmidt, H. G. (2007). No retrieval-induced forgetting using item-specific independent cues: Evidence against a general inhibitory account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 950–958. doi:10.1037/0278-7393.33.5.950
- Camp, G., Pecher, D., Schmidt, H. G., & Zeelenberg, R. (2009). Are independent probes truly independent? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 934–942. doi:10.1037/a0015536

- Gómez-Ariza, C. J., Lechuga, M. T., Pelegrina, S., & Bajo, M. T. (2005). Retrieval-induced forgetting in recall and recognition of thematically related and unrelated sentences. *Memory & Cognition*, 33, 1431–1441. doi:10.3758/BF03193376
- Hicks, J. L., & Starns, J. J. (2004). Retrieval-induced forgetting occurs in tests of item recognition. *Psychonomic Bulletin & Review*, 11, 125–130. doi:10.3758/BF03206471
- Jacoby, L. L., Jones, T. C., & Dolan, P. O. (1998). Two effects of repetition: Support for a dual-process model of know judgments and exclusion errors. *Psychonomic Bulletin & Review*, 5, 705–709. doi:10.3758/BF03208849
- Koutstaal, W., Schacter, D. L., Johnson, M. K., & Galluccio, L. (1999). Facilitation and impairment of event memory produced by photograph review. *Memory & Cognition*, 27, 478–493. doi:10.3758/BF03211542
- Perfect, T. J., Stark, L.-J., Tree, J. J., Moulin, C. J. A., Ahmed, L., & Hutter, R. (2004). Transfer appropriate forgetting: The cue-dependent nature of retrieval-induced forgetting. *Journal of Memory and Language*, 51, 399–417. doi:10.1016/j.jml.2004.06.003
- Racsmány, M., & Conway, M. A. (2006). Episodic inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 44–57. doi:10.1037/0278-7393.32.1.44
- Racsmány, M., Conway, M. A., Garab, E. A., & Nagymáté, G. (2008). Memory awareness following episodic inhibition. *Quarterly Journal of Experimental Psychology*, 61, 525–534. doi:10.1080/17470210701728750
- Rotello, C. M., Macmillan, N. A., & Reeder, J. A. (2004). Sum-difference theory of remembering and knowing: A two-dimensional signal-detection model. *Psychological Review*, 111, 588–616. doi:10.1037/0033-295X.111.3.588
- Saunders, J., & MacLeod, M. D. (2002). New evidence on the suggestibility of memory: The role of retrieval-induced forgetting in misinformation effects. *Journal of Experimental Psychology: Applied*, 8, 127–142. doi:10.1037/1076-898X.8.2.127
- Spitzer, B., & Bäuml, K.-H. (2007). Retrieval-induced forgetting in item recognition: Evidence for a reduction in general memory strength. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 863–875. doi:10.1037/0278-7393.33.5.863
- Spitzer, B., & Bäuml, K.-H. (2009). Retrieval-induced forgetting in a category recognition task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 286–291. doi:10.1037/a0014363
- Swets, J. A., Tanner, W. P., Jr., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review*, 68, 301–340. doi:10.1037/h0040547
- Veling, H., & van Knippenberg, A. (2004). Remembering can cause inhibition: Retrieval-induced inhibition as cue independent process. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 315–318. doi:10.1037/0278-7393.30.2.315
- Verde, M. F. (2004). The retrieval practice effect in associative recognition. *Memory & Cognition*, 32, 1265–1272. doi:10.3758/BF03206317
- Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, 114, 152–176. doi:10.1037/0033-295X.114.1.203
- Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, 117, 1025–1054. doi:10.1037/a0020874
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46, 441–517. doi:10.1006/jmla.2002.2864
- Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1341–1354.