Do learners predict a shift from recency to primacy with delay?

Benjamin C. Storm[1] · Robert A. Bjork[2]

Abstract The shift from recency to primacy with delay reflects a fundamental observation in the study of memory. As time passes, the accessibility of earlier-learned representations tends to increase relative to the accessibility of later-learned representations. In three experiments involving participants' memory for text materials, we examined whether participants understood that there might be such a shift with retention interval. In marked contrast to their actual performance, participants predicted recency effects at both shorter and longer retention intervals. Our findings add to the evidence that the storage and retrieval dynamics of the human memory system, though adaptive overall from a statistics-of-use standpoint, are both complex and poorly understood by users of the system.

A question that has endured since the early days of research on human memory is why information and experiences, once encoded and remembered, become forgotten. One intuitive theory is that forgetting is the consequence of decay or disuse with the passage of time (Thorndike, 1914). As pointed out by McGeoch (1932), however, and as confirmed by many empirical observations, forgetting and disuse are not always correlated—and even when they are correlated, such correlations can often be attributed to experiences or contextual changes,

which McGeoch called "altered stimulating conditions," that are independent of the passage of time. What McGeoch emphasized as an alternative to decay is what he called "reproductive inhibition"—that is, we forget because access to some target memory is blocked by, or interfered with, by some related item in memory.

Intrinsic to the interference theory of forgetting is that there is competition among items or responses associated to a cue. When attempting to retrieve a particular item, other items associated with that cue can cause interference, thereby impairing access to the target item. The extent to which a target item suffers interference depends on a number of factors, such as the similarity, strength, and number of competing items in memory (Crowder, 1976).

The shift from recency to primacy with retention interval

Interference can occur both proactively and retroactively. Proactive interference (PI) is observed when earlier learning interferes with the remembering of later learning. Retroactive interference (RI), on the other hand, is observed when later learning interferes with the remembering of earlier learning. Interestingly, the effects of PI and RI have been shown to change with delay. Using an A-B A-D paired-associate paradigm, for example, research has shown that RI tends to decrease across a retention interval, whereas PI tends to increase (e.g., Briggs, 1954; Forrester, 1970; Koppenaal, 1963; Postman, Stark, & Fraser, 1968). This shift is sometimes strong enough to create a phenomenon known as *spontaneous recovery*. If participants learn A-B pairs before learning A-D pairs, for example, the ability to recall B given A can sometimes increase with delay such that performance improves over time, presumably because RI from the D response

✉ Benjamin C. Storm
bcstorm@ucsc.edu

[1] Department of Psychology, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

[2] University of California, Los Angeles, CA, USA

dissipates (Brown, 1976; Wheeler, 1995). Related findings have also been observed in the context of free recall (e.g., Bjork, 1975; Bjork & Whitten, 1974; Bower & Reitman, 1972). It has been shown, for example, that a shift from recency to primacy occurs with delay such that a recency advantage is observed after a short delay (later-presented items are remembered better than earlier-presented items), whereas a primacy advantage is observed after a long delay (earlier items are remembered better than later-presented items).

A shift from recency to primacy has been observed in many different contexts and has been considered by many researchers to be of significant theoretical importance (e.g., Bouton & Peck, 1992; Estes, 1955; Jost, 1897; Lang, Craske, & Bjork, 1999; Pavlov, 1927). Indeed, the tendency for memory to regress over time, with earlier representations gaining strength over later representations, has been argued to constitute a fundamental property of human memory (Bjork, 1978, 2001; Bjork, & Bjork, 1992). From the perspective of trying to most adaptively predict the information, prepotencies, and behaviors that are going to be most useful in the future (Anderson, 1990; Anderson & Milson, 1989), a memory system that regresses over time could be expected to provide certain mnemonic advantages. Memories with a longer history of use might, for example, be more likely from a statistical standpoint to become useful again in the future than memories with a relatively more recent frequency of use. That is, whereas information and procedures from the recent past are likely to be the most needed in the near future, the statistics of use may change after a period of disuse: As argued by Bjork (2001), the "fact that there has not been a need for recent information and procedures may signal changes that mean that older, typically better learned, information and procedures are once again relevant" (p. 229).

Goals of the present research

This research was conducted with several goals in mind. First, we sought to demonstrate that a shift from recency to primacy (and perhaps even absolute recovery) of earlier learned memory representations can be observed using the type of meaningful materials that one might encounter in educational situations. To date, the majority of research has employed cue–response associations or simple lists of stimuli, such as words and pictures. Memory regression effects are particularly important to consider in the context of education, however, because the ability to recall to-be-learned information might differ as a function of the delay after learning, with initial learning becoming increasingly more recallable relative to later learning, and knowledge or awareness of such dynamics could be useful in the design and implementation of educational situations. When students have misimpressions, for example, as in the case of naïve physics errors, they may at one point overcome such errors; but, owing to spontaneous recovery,

they may later find themselves affected by those errors once again. Moreover, there has been some disagreement about the extent to which memory representations really do regress over time (see, e.g., Wixted, 2004). A demonstration that regression effects can be observed with other, more complex, materials would provide additional support for the robustness and generality of the phenomenon.

A second goal of this research was to examine whether people consider memory regression effects when making metacognitive judgments about their learning. That is, assuming that evidence does emerge that earlier learning recovers relative to later learning, would participants take such dynamics into account when making predictions about their ability to recall such information? Might participants anticipate the recovery of old information relative to new information over time? From a metacognitive standpoint, there are good reasons to expect participants, when making such judgments, to be oblivious to factors such as recency and primacy, and most certainly to the possibility that the effects of such factors might differ across different delays. Research has shown that the metacognitive judgments people make can be influenced by a variety of cues or sources of information, many of which are largely indirect and inferential (e.g., Bjork, Dunlosky, & Kornell, 2013; Jacoby & Kelley, 1987; Koriat, 1997; Nelson & Narens, 1990; Schwartz, Benjamin, & Bjork, 1997). In some instances, people rely on knowledge-based beliefs or extrinsic cues when making judgments; in other instances, people are influenced by more salient experienced-based information or intrinsic cues, such as perceptual or retrieval fluency (e.g., Benjamin, Bjork, & Schwartz, 1998; Kelley & Jacoby, 1996; Koriat, Bjork, Sheffer, & Bar, 2004; Kornell & Bjork, 2009; Kornell, Rhodes, Castel, & Tauber, 2011). Koriat et al. (2004), for example, exposed participants to lists of cue–response pairs and then asked them to predict the likelihood of recalling the responses after varying delays. Amazingly, participants predicted similar levels of recall regardless of whether they were asked to predict performance after 10 minutes, 1week, or even 1 year. It was only when participants were prompted to consider their knowledge-based theories of forgetting and time (e.g., by having participants make predictions about performance at two different delays) that their predictions were then influenced by the length of the retention interval.

The work by Koriat et al. (2004) and others has suggested that even if participants do hold strong knowledge-based beliefs about a factor related to memory (such as the fact that forgetting occurs over time), such theories may nevertheless have little impact on judgments when such theories are outshined by more salient experienced-based information, such as the familiarity or fluency of the particular information being learned. Thus, in the context of this research, even if people do generally believe that memory regresses over time (an idea we will return to in Experiments 2 and 3), such a belief may not

necessarily influence the metacognitive judgments they make about their memory. Indeed, owing to the overwhelming saliency of intrinsic cues such as fluency and familiarity, we might expect participants to be relatively insensitive to factors such as primacy or recency when making judgments about their memory (Castel, 2008), a finding which would also be consistent with evidence that participants can be largely oblivious to the dynamics underlying the cue-specific nature of associative interference (Diaz & Benjamin, 2011).

Finally, at a more general level, it is important to note that memory cannot be adequately indexed by a single type or measure of strength (Bjork & Bjork, 1992; Estes, 1955; Hull, 1943; Tulving & Pearlstone, 1966). In the *New Theory of Disuse*, for example, Bjork and Bjork (1992) distinguish between the momentary accessibility of an item in memory (i.e., retrieval strength) and the extent to which an item is well learned or entrenched with other representations in memory (i.e., storage strength). The need for this distinction is particularly apparent in the context of the shift from recency to primacy because such demonstrations highlight the importance of considering both storage strength and retrieval strength (as well as their interaction) when predicting the future accessibility of an item. Specifically, primacy items are presumed to recover because their high degree of storage strength is initially masked by their low degree of retrieval strength (caused by the interference of recency items); when such interference dissipates, however, the storage strength of primacy items wins out, thereby leading to their increase in retrieval strength and thus their relative or even absolute recovery in recall performance. Although people may have metacognitive access to something akin to retrieval strength, in the sense that they can use cues and markers of current accessibility to inform their predictions, it seems unlikely that they would have access to something akin to storage strength. That is, people may not have access to the extent to which an item is embedded or well learned in memory independent of its momentary accessibility. Thus, barring the use of some more general knowledge-based assumption that memory regresses over time, it stands to reason that the metacognitive judgments people make about future memory performance would largely ignore, if not completely neglect, the possibility of a shift from recency to primacy with delay.

Experiment 1

Method

Participants Sixty undergraduate students (*M*age = 20.2 years) from the University of California, Los Angeles (UCLA), participated for credit in an introductory psychology course.

Materials Eight passages were created, each consisting of information about the geography, climate, people, and location of a particular region of the world. One passage served as the critical passage for which memory would be tested and metacognitive judgments would be made (Antarctica; see Appendix A). The other seven passages served as filler passages to create the primacy and recency conditions (Africa, Australia, Canada, Greenland, Hawaii, Norway, & Siberia). The average passage length was 267 words, and each passage was shown as four separate paragraphs.

The order of the paragraphs within a passage was the same for all eight regions. Participants first learned about the geography of a given region, then the climate, followed by the people, and finally the location. By having participants learn the same type of information about the different regions, we hoped to maximize interference between regions. The accessibility of climatic information related to one region should, for example, be affected by the accessibility of climatic information related to other regions. Previous work using these materials has shown that they are capable of inducing significant interpassage interference (Little, Storm, & Bjork, 2011; Storm, Bjork, & Storm, 2010).

Design Participants were randomly assigned to one of four between-subjects conditions using a 2 (position: first vs. last) × 2 (delay: 0 min vs. 30 min) factorial design. Thirty participants read the Antarctica passage and then seven filler passages (first position), whereas the other 30 participants read the Antarctica passage after reading the seven filler passages (last position). Within each position condition, half of the participants were tested immediately after reading the passages, whereas the other half was tested following a 30-min delay. All participants were asked to make a judgment of learning about the Antarctica passage immediately after reading all eight passages. Participants in the immediate condition were asked to predict their performance on a test that would be given immediately, whereas participants in the delayed condition were asked to predict their performance on a test that would be given after a 30-min delay.

Procedure The experiment began with participants being told that they were going to learn about various regions of the world. Each passage was presented as four separate paragraphs (geography, climate, people, and location) for 1 min (15 seconds per paragraph).

After studying all eight passages, participants in the immediate condition were given the following prompt:

> "You will now be asked to recall as many facts about *Antarctica* as you can. What percentage of the facts about *Antarctica* do you think you will be able to recall?

Please note that you will have 4 minutes to recall the facts, and that you will be taking the test immediately."

The word *Antarctica* was shown in italics to make it clear that they were only to be asked to recall information about Antarctica. Participants in the delayed condition were given the same prompt except they were asked to predict their performance after a 30-min delay:

"In thirty minutes you will be asked to recall as many facts about *Antarctica* as you can. What percentage of the facts about *Antarctica* do you think you will be able to recall? Please note that you will have 4 minutes to recall the facts, and that you will be taking the test in 30 minutes."

Participants were instructed to answer the prompt by telling the experimenter a number between 0 and 100.

Participants in the immediate condition were then given a blank sheet of paper and 4 minutes to recall as many facts about Antarctica as possible. Participants in the 30-min delay condition were given the same test after a 30-min interval, which was filled with a series of unrelated distractor tasks (e.g., learning and retrieving category-exemplar pairs). It is worth noting that participants were not informed of the nature of the distractor tasks before making their predictions, a factor which may have influenced the accuracy of such predictions. Given the purpose of the study, however, we felt it was more important to focus instructions on the particular passage that was to-be-tested and the particular retention interval that was to be employed, especially because there was little reason to predict that the unrelated distractor tasks would significantly influence recall performance for the passage.

Recall performance was scored by a research assistant blind to experimental condition. Fifteen critical facts were identified in the Antarctica passage, and performance was scored as the proportion of those facts that were included in each participant's response.

Results

Recall performance The proportion of facts about Antarctica that were recalled on the final test was analyzed using a 2 (position: first vs. last) × 2 (delay: 0 min vs. 30 min) analysis of variance (ANOVA). Significant main effects were not observed with regard to position, $F(1, 56) = 1.65$, $MSE = .02$, $p = .21$, $\eta_p^2 = .03$, or delay, $F(1, 56) = 0.55$, $MSE = .02$, $p = .46$, $\eta_p^2 = .01$, but a significant interaction was observed, $F(1, 56) = 18.09$, $MSE = .02$, $p < .001$, $\eta_p^2 = .24$. As shown in Fig. 1, when the test was immediate, participants recalled significantly fewer facts about Antarctica when they read it first than when they read it last, $t(28) = 3.75$, $p = .001$, $d = 1.37$. When the test was delayed, however, participants recalled significantly more facts about Antarctica when they read it first than when they read it last, $t(28) = 2.20$, $p = .04$, $d = 0.80$. These results suggest that a shift from recency to primacy was observed with delay. In fact, the shift was so strong that an effect of absolute recovery was observed such that participants who read the Antarctica passage first demonstrated significantly better performance after a 30-min delay than did participants who were tested after a 0-min delay, $t(28) = 2.49$, $p = .02$, $d = .91$. Participants who read the Antarctica passage last exhibited a more familiar effect of forgetting over time, performing significantly worse after a 30-min delay than did participants who were tested after a 0-min delay, $t(28) = 3.52$, $p = .002$, $d = 1.28$.

Metacognitive predictions We next examined the predictions participants made about their test performance using a 2 (position: first vs. last) × 2 (delay: 0 min vs. 30 min) ANOVA. Significant main effects were not observed with regard to condition, $F(1, 56) = 0.01$, $MSE = .03$, $p = .91$, $\eta_p^2 = .00$, or delay, $F(1, 56) = 0.07$, $MSE = .03$, $p = .79$, $\eta_p^2 = .00$, nor was a there a significant interaction, $F(1, 56) = 0.34$, $MSE = .03$, $p = .56$, $\eta_p^2 = .01$. As shown in the left panel of Fig. 1, participants predicted they would be able to recall approximately the same number of facts in all four conditions (*t* tests showed that all
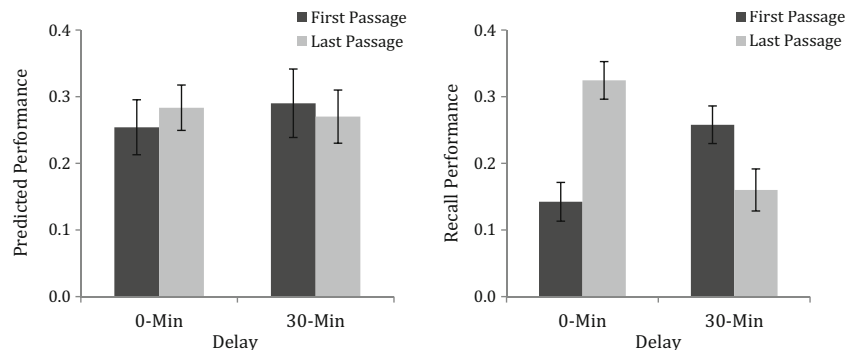
Fig. 1 The left panel shows the predicted performance of participants (displayed as the proportion of facts they thought they would recall at final test) as a function of condition. The right panel shows actual performance at final test as a function of condition. All participants made predictions about their performance on a specific passage after a specific delay and were then tested on that specific passage after that specific delay. Error bars reflect standard errors of the mean

p values were >.55). To further confirm that the pattern of results in recall and judgments differed, we conducted a 2 (measure: recall vs. judgments) × 2 (position: first vs. last) × 2 (delay: 0 min vs. 30 min) mixed-design ANOVA, which revealed a significant three-way interaction, $F(1, 56) = 7.77$, $MSE = .01$, $p = .007$, $\eta_p^2 = .12$. Taken together, these results suggest that despite substantial effects on actual recall performance, participants did not consider delay, condition, or their interaction when making predictions about performance.

Experiment 2

The results of the first experiment provide clear evidence of a memory regression effect. When given an immediate test, participants recalled more information about the critical passage when it was studied last than when it was studied first. When given a delayed test, however, this pattern reversed such that they recalled more information about the critical passage when it was studied first than when it was studied last. This shift from recency to primacy was strong enough to overcome the typical effects of forgetting that occur with the passage of time, thus resulting in not only a relative recovery in accessibility but an absolute recovery as well. That is, when given a delayed test, information about Antarctica was not only better recalled when it was studied first than when it was studied last but it was better recalled than it would have been had the test been given without a delay.

It is not surprising—given the discussion in the introduction about how people make metacognitive judgments—that participants failed to account for regression effects when predicting their performance. Indeed, participants predicted the same (nonsignificantly different) levels of performance in each of the between-subjects conditions. Not only did participants fail to consider the effects of PI and RI, but in a replication of Koriat et al. (2004), they also failed to consider the effects of retention interval. That is, participants did not predict worse performance after a 30-min delay than after a 0-min delay.

One could correctly point out that actual recall performance failed to differ—at least overall—between the delay manipulations, so the failure to consider retention interval in this context might in some sense be considered an accurate metacognitive judgment. We explore this issue further in Experiment 2 by having participants make predictions about their recall performance in two different delay conditions. As demonstrated by Koriat et al. (2004), participants do take retention interval into account when asked to make predictions about performance after two different delays, presumably because it is the comparison of the delay conditions that prompts them to consider the idea that forgetting occurs with time. Thus, although participants failed to predict forgetting across

a delay in Experiment 1, we conjectured that they would predict forgetting across the same delay in Experiment 2.

A second major goal of Experiment 2 was to provide a more sensitive assessment of what participants believe with regard to primacy and recency. For reasons specified above, it is not surprising that participants failed to account for these factors, as other more salient experience-based information (e.g., the familiarity of the information presented in the critical passage) might have served as a much more powerful cue for participants when making their metacognitive judgments than did, perhaps, any belief they might have that later-learned items should be easier to recall than earlier-learned items.

Given these considerations, participants in Experiment 2 were asked to make four predictions, thereby making both delay and position within-subjects manipulations. Specifically, participants were asked to predict recall performance for both the first and last passage presented on both an immediate and delayed final test. Perhaps under these conditions participants would exhibit some sensitivity to the primacy/recency manipulation. One prediction that seems particularly plausible, for example, is that participants will predict better performance on the last passage than on the first passage (recency advantage). Students have substantial experience with the fact that new learning can interfere with old learning, and there is also the consideration that the last passage was presented more recently than the first passage. Whether participants would predict that the recency advantage would become less across a delay, however, is much harder to say. Such a finding would suggest, however, that participants might be, at least to some degree, aware of the idea that earlier learning recovers over time relative to later learning.

Finally, to examine these issues further, an additional manipulation was added to the design. One group of participants (read condition) was asked to actually read the eight passages before making their judgments about subsequent recall performance. A second group of participants (baseline condition) did not read the passages. Instead, they were simply given a description of the paradigm and asked to make hypothetical predictions about recall performance as a function of position and delay. Because all item-based or experience-based information is removed in the baseline condition, the predictions that participants make should provide a much more sensitive assessment of their beliefs about the effects of primacy and recency and how those effects vary with delay.

Method

Participants A total of 96 UCLA undergraduates (Mage = 20.6 years) participated for course credit in an introductory psychology course.

Materials and procedure Participants were randomly assigned to one of two between-subjects conditions: read vs.

baseline. In this experiment, participants were asked to only make predictions about their performance on a future test. A test was not actually administered.

In the baseline condition ($n = 48$), participants were provided with the following prompt (bolding and underlining included):

> "Imagine that you read 8 brief passages, each describing a different region of the world. After reading all 8 passages you are given an immediate final test on only one of the passages. During this final test, you are asked to free recall as many facts about that region of the world as possible. What percentage of the facts do you think you will be able to recall if the region tested is the first region that you read about? What percentage of the facts will you be able to recall if the region tested is the last region that you read about?"

Eight boxes were then presented horizontally on a paper, each representing a different passage, with the first and last passages shown in bold. Within the first and last box was the word "test" followed by a blank line. Participants were instructed to report their predictions of performance on each of those lines. Participants then read a second prompt:

> "Now imagine the same situation, except that you are given a final test after a 30-minute delay on only one of the passages. What percentage of the facts do you think you will be able to recall if the region tested is the first region that you read about? What percentage of the facts will you be able to recall if the region tested is the last region that you read about?"

Once again, eight boxes were presented horizontally on the paper in the same way they were presented before, with participants instructed to report their predictions of performance on each of the lines. The order of the judgments in the immediate and 30-min delay conditions was not counterbalanced. That is, participants always made predictions about the immediate test prior to making predictions about the delayed test. This design feature was chosen to highlight the delay manipulation to participants, hopefully encouraging them to consider the factor of delay when making their metacognitive judgments (e.g., what happens to the accessibility of items across a subsequent delay?).

Participants in the read condition ($n = 48$) made the same predictions as participants in the baseline condition, except they did so after actually reading the eight passages. Specifically, participants in the read condition read eight passages, each describing a different region of the world (i.e., Norway, Siberia, Australia, Africa, Antarctica, Hawaii, Canada, & Greenland). The passages were the same as those used in Experiment 1, each consisting of information about location, geography, climate, and people. In this experiment, Norway and Greenland served as the critical passages, with each passage being presented either first or last. Half of the participants read the Norway passage first and the Greenland passage last, whereas the other half read the Greenland passage first and the Norway passage last. The passages were displayed in the same way as Experiment 1.

Immediately after reading the passages, participants were given the same prompts shown above, modified only slightly because they had actually read the passages. For the first prompt, participants read:

> "You just read 8 brief passages, each describing a different region of the world. You may now be given an immediate final test on only one of the passages. During this final test, you will be asked to free recall as many facts about that region of the world as possible. What percentage of the facts do you think you will be able to recall if the region tested is Greenland? What percentage of the facts will you be able to recall if the region tested is Norway?"

Eight boxes were then presented horizontally, each labeled with the name of a region and presented in the order they had been studied. As in the Baseline Condition, participants were instructed to report their predictions of performance for the first and last passages, which depending on counterbalancing condition, were listed as either Greenland or Norway. Participants then read a second prompt:

> "Now imagine that you are given a final test after a 30-minute delay on only one of the passages. What percentage of the facts do you think you will be able to recall if the region tested is Greenland? What percentage of the facts will you be able to recall if the region tested is Norway?"

Once again, eight boxes were presented, each labeled with the name of a region and presented in the order they had been studied, and participants were instructed to report their predictions for the first and last passages.

Results

The proportion of facts participants predicted they would be able to recall were analyzed using a 2 (condition: read vs. baseline) × 2 (position: first vs. last) × 2 (delay: 0 min vs. 30 min) mixed-design ANOVA, with condition serving as the only between-subjects manipulation. As shown in Table 1, all main effects and interactions were statistically significant. To focus subsequent analysis on comparisons of most interest, we examined data from the baseline and read conditions separately.

Table 1 Analysis of Variance for Predicted Recall in Experiment 2 as a Function of Delay (0 min vs. 30 min), Position (Primacy vs. Recency), and Condition (Read vs. Baseline)

Source	F	MSE	p	Partial η^2
Delay	157.15	.008	.001	.63
Position	18.52	.053	.002	.17
Condition	48.20	.104	.001	.34
Delay × Position	11.16	.005	.001	.11
Delay × Condition	5.47	.008	.022	.06
Position × Condition	11.55	.053	.001	.11
Delay × Position × Condition	10.57	.005	.002	.10

Degrees of freedom for all effects: (1, 94)

First, with regard to participants in the baseline condition, predictions of performance were analyzed using a 2 (position: first vs. last) × 2 (delay: 0 min vs. 30 min) ANOVA. Not surprisingly, participants predicted they would recall more facts immediately ($M = .58$, $SE = .02$) than after a 30-min delay ($M = .44$, $SE = .03$), $F(1, 47) = 114.73$, $p < .001$, $\eta_p^2 = .71$. Participants also predicted they would recall more facts about the last passage ($M = .60$, $SE = .03$) than the first passage ($M = .42$, $SE = .03$), $F(1, 47) = 20.28$, $p < .001$, $\eta_p^2 = .30$. As shown in Table 2, a significant interaction emerged such that participants predicted a larger difference in performance between the first and last passages in the 0-min condition than in the 30-min condition, $F(1, 47) = 12.53$, $p = .001$, $\eta_p^2 = .21$.

Although suggestive, the interaction in the baseline condition may be largely attributable to differences in overall performance. Note that participants predicted strong recency effects in both the 0-min, $t(47) = 5.16$, $p < .001$, $d = .74$, and 30-min conditions, $t(47) = 3.34$, $p = .002$, $d = .48$. Moreover, a significant difference was not observed when we calculated effects of predicted forgetting in the first and last conditions

Table 2 Predicted Recall Performance in Experiments 2 and 3 as a Function of Delay, Position, and Condition

Condition Delay	Position of Passage						
	First		Last		Recency Advantage		
	M	SE	M	SE	M	d	p
Experiment 2: Read (read passages)							
0 min	.32	.03	.34	.03	.02	.12	.392
30 min	.23	.03	.25	.03	.02	.12	.402
Experiment 2: Baseline (did not read passages)							
0 min	.47	.03	.70	.02	.23	.74	<.001
30 min	.37	.03	.51	.03	.14	.48	.002
Experiment 3: Baseline (Did not read passages)							
30 min	.41	.01	.45	.01	.05	.35	<.001
1 week	.28	.02	.33	.02	.05	.30	.002

relative to performance in the 0-min condition (by subtracting 30-min performance from 0-min performance and then dividing by performance in the 0-min condition). Specifically, for the first and last passages, respectively, participants predicted that 24% ($SE = 5\%$) and 28% ($SE = 3\%$) of the information that was recalled at the 0-min test would be forgotten by the time of the 30-min test, $t(47) = .79$, $p = .43$, $d = .11$.

A very different pattern was observed for participants who read the passages prior to making their predictions. The proportion of facts that these participants predicted they would recall was analyzed using a 2 (position: first vs. last) × 2 (delay: 0 min vs. 30 min) ANOVA. Once again, participants predicted that they would be able to recall more facts immediately ($M = .33$, $SE = .03$) than after a 30-min delay ($M = .24$, $SE = .02$), $F(1, 47) = 50.20$, $p < .001$, $\eta_p^2 = .52$. This time, however, participants did not predict that they would recall significantly more facts about the last passage ($M = .29$, $SE = .03$) than they would about the first passage ($M = .27$, $SE = .03$), $F(1, 47) = 0.76$, $p = .39$, $\eta_p^2 = .02$. Moreover, we failed to find any evidence of an interaction, $F(1, 47) = 0.02$, $p = .90$, $\eta_p^2 = .00$.

Experiment 3

The results of Experiment 2 provide a number of notable findings. First, as expected, when participants were asked to make predictions about performance at two different delays, they predicted better performance after a 0-min delay than after a 30-min delay. This suggests that the failure to consider delay in Experiment 1 reflected a failure to consider the effects of delay when making metacognitive judgments, and not to an accurate prediction that forgetting would not be observed with the passage of time. Interestingly, participants who actually read the passages (read condition) predicted a 9% forgetting effect across the 30-min delay, whereas participants who did not read the passages (baseline condition) predicted a 15% forgetting effect across the 30-min delay. This difference, which was statistically significant, $F(1, 94) = 5.47$, $MSE = .004$, $p = .02$, suggests that participants may have been more likely to rely on the knowledge-based factor that forgetting occurs with time in the baseline condition than in the read condition.

Second, with regard to the position manipulation, participants did predict better recall performance for the last passage than for the first passage (recency advantage), but only when they were asked to make hypothetical predictions in the baseline condition. No evidence of a predicted recency advantage was observed when participants were given the opportunity to actually read the passages. This finding suggests that although people may believe that later-learned information is going to be better recalled than earlier information, such a belief may not necessarily influence metacognitive judgments in the

presence of more salient experience-based information. It is important, however, to interpret this null effect with caution. Participants made judgments about the first passage after a much longer delay than they made judgments about the last passage. As a consequence of this difference, the extent to which participants may have relied on experienced-based information when making their judgments could have differed across the two conditions, thereby making comparisons between the conditions difficult to interpret.

Third, an interaction was observed in the baseline condition such that participants predicted a larger recency advantage in the 0-min condition (23%) than in the 30-min condition (14%). Although not a shift to primacy, which is what happened when participants were actually tested in Experiment 1, this finding does suggest that participants might be somewhat aware of the idea that recency effects become reduced across a delay. For reasons discussed in the Results section, however, this difference should be interpreted with caution, as the recency advantage may have been reduced in the 30-min condition compared to the 0-min condition not because participants believed that recency advantages become smaller with delay, but simply because performance for the last passage had more room to drop across the delay compared to performance for the first passage. This latter explanation was supported by a subsequent analysis showing that when predicted forgetting effects were calculated as a ratio relative to performance in the 0-min condition, participants did not predict differing amounts of forgetting for information in the first and last passage.

In Experiment 3, we sought to provide additional evidence that participants do not believe there is a shift to primacy over time. Before making their predictions, participants were provided with the starting assumption that recall performance would be the same on an immediate test for the first and last passages (i.e., 50%). Thus, if participants do believe that there is a shift to primacy with delay, then they should predict worse performance for the last passage than the first passage after a delay. Note that this condition also provides a more or less direct examination of people's metacognitive predictions regarding Jost's (1897) law: specifically, that "if two associations are now of equal strength but different ages, the older one will lose strength more slowly with the passage of time" (as translated by Woodworth & Schlossberg, 1954, p. 730).

A second goal was to ask participants to predict performance after a more lengthy delay. Although participants might not predict regression effects after a 30-min delay, they might be more likely to predict regression effects after a 1-week delay. The logic underlying this prediction is that a longer delay might be more likely than a shorter delay to prompt participants to consider knowledge-based factors related to changes in memory over time. Moreover, a longer delay could be more representative of the conditions under which regression effects are often observed in the natural use of memory.

That is, participants may have more experience with earlier learning recovering relative to later learning after a long delay than they do after a short delay. It is possible, therefore, that participants will predict a diminished recency effect (and even a primacy effect) in the 1-week condition compared to the 30-min condition.

Method

Participants A total of 119 UCLA undergraduates (*M*age = 20.6 years) participated for course credit in an introductory psychology course. The study was run in the context of another study. Specifically, participants completed the study described below before beginning a separate, unrelated experiment.

Materials and procedure The design was similar to the baseline condition of Experiment 2, except for a few important differences. First, participants were asked to predict performance on the first and last passages on hypothetical tests that were said to occur after 30-min and 1-week delays. Second, participants were told that if given an immediate test, people typically recall 50% of the facts about each of the regions. The prompt specifically read:

> "Imagine that you read 8 brief passages, each describing a different region of the world. After reading all 8 passages, and a 30-minute or 1-week delay, you are given a final test on only one of the passages. During this final test, you are asked to free recall as many facts about that region of the world as possible. What percentage of the facts do you think you will be able to recall if the region tested is the first region? What percentage of the facts will you be able to recall if the region tested is the last region? Please keep in mind that when the test is immediate, participants typically recall 50% of the facts about the first region and 50% of the facts about the last region (as shown below)."

Provided immediately below the prompt were three rows of boxes, with each row representing a different final test. The top row was labeled "Immediate" and had final recall performance already written inside the first and last boxes (50%). The middle row and bottom row were labeled "After a 30-Minute Delay" and "After a 1-Week Delay," respectively, and a blank space was provided in each of the first and last boxes for participants to provide their predictions.

Results

The proportion of facts that participants predicted they would be able to recall were analyzed using a 2 (position: first vs. last) × 2 (delay: 30 min vs. 1 week) repeated-measures

ANOVA. Participants predicted that they would recall more facts after a 30-min delay ($M = .43$, $SE = .01$) than after a 1-week delay ($M = .30$, $SE = .02$), $F(1, 118) = 89.94$, $p < .001$, $\eta_p^2 = .43$. Participants also predicted they would recall more facts about the last passage ($M = .39$, $SE = .03$) than the first passage ($M = .34$, $SE = .03$), $F(1, 118) = 16.12$, $p < .001$, $\eta_p^2 = .12$. As shown in Table 2, however, no evidence of a significant interaction was observed, $F(1, 118) = 0.00$, $p = .99$, $\eta_p^2 = .00$. These results suggest that participants do not expect a primacy advantage to emerge over time. Instead, they suggest that in contrast to Jost's (1897) law, participants expect a recency advantage to emerge such that—despite equivalent initial levels of recall—later-learned information will become relatively more accessible than earlier-learned information after a delay.

General discussion

The results of this research are unequivocal: The participants clearly did not appreciate what Bjork (1978) referred to as an absolutely fundamental property of human memory—namely, that access to items in memory tends to shift from recency to primacy with retention interval (or, said differently, that there is a "shift, with delay, from preferential access to newer memory representations to preferential access to older memory representations"; Bjork, 2001, p. 211).

This lack of appreciation can be juxtaposed with actual performance. Not only was relative recovery observed, but absolute recovery was observed such that earlier-learned information became more recallable after a 30-min delay than it was after a 0-min delay. Evidence of absolute recovery in this context with educationally relevant materials provides an important replication/extension of the classic memory phenomenon. It may be surprising to some readers that such a robust effect of absolute recovery was observed, but one possible explanation is that the particular materials and procedures we employed resulted in relatively little "extra-experimental" forgetting. Postman (1971) argued, for example, that recovery may be fairly common in most situations, but masked by sources of forgetting that act to decrease the recall of both recency and primacy information. In this study, participants recalled only 2% fewer facts after 30 min than after 0 min, and it may have been this lack of extra-experimental forgetting that allowed absolute recovery to be observed so robustly.

If it is the case that the shift from recency to primacy with delay is a fundamental and common characteristic of the memories of both human and nonhuman animals, as we argued earlier, why is it a feature of human memory that our participants did not appreciate—and we think, why is it unappreciated by people in general? After all, to the degree that the shift is common, people have multiple opportunities to experience older memories becoming more accessible than newer memories after a delay. Examples might be remembering,

years later, a friend's maiden name rather than their married name, or, after a period away from a sport such as tennis or golf, having one's old swing become more accessible than one's new/improved swing.

One consideration is that the shift to primacy, especially when it results in an absolute increase in the recall of such items across some retention interval, runs counter to a general characteristic of human memory the people *do* seem to understand—namely, that information and procedures become less accessible over time. That is, in general, we understand that we forget (even though, as demonstrated by Koriat et al., 2004, we are susceptible to believing that our subsequent ability to remember is governed by characteristics of the to-be-remembered materials, rather than length of a retention interval). From that standpoint, a shift from recency to primacy violates what people think of as a general rule of memory: Older memories, other things being equal, should be less recallable, not more recallable, than newer memories, because they are older. In that sense, the shift from recency to primacy is an "important peculiarity" of human memory (Bjork & Bjork, 1992). It is also peculiar because it is not characteristic of external storage systems, such as newspapers, books, or the memory in a computer, where things become gradually harder to find as time passes and more things are stored.

It remains to be seen whether there are conditions under which people do demonstrate metacognitive sensitivity to memory regression effects and the shift from recency to primacy with delay. It is possible, for example, that people are more likely to demonstrate such sensitivity after extensive relevant experience, particularly when making predictions in contexts that are highly similar to those that have been experienced. Although participants may have substantial experience with regression effects in some domains, they may have much less experience in other domains. It is entirely possible, therefore, that participants could be largely oblivious to regression effects in the context of predicting the retention of to-be-learned information from an educational passage while fully appreciating regression effects within other contexts, such as in making predictions about the memorability of another person's name or the effectiveness of their own golf swing.

Concluding comment

Across the relatively short history of research on metamemory and metacognitive processes, many findings have demonstrated that there can be striking discrepancies between how we learn and remember, versus how we think we learn and remember. These findings add to that list. As we sketched in the Introduction, and as one of us has argued in more detail in various ways at various times (Bjork, 1978, 2001, 2011; Bjork & Bjork, 1992), the shift, with delay, from preferential access to newer memory representations to preferential access to

older representations is adaptive from a statistics-of-use stand-point. This adaptive feature of the human memory system seems not, though, to be understood by users of the system.

Appendix A

Antarctica Passage

Geography Antarctica is the icy continent at the South Pole. It is covered by permafrost (permanently frozen ground), is surrounded by water, and is about 1.5 times larger than the United States. The world's largest desert is on Antarctica. Ninety-eight percent of the land is covered with a continental ice sheet; the remaining 2% of land is barren rock. Antarctica has about 87% of the world's ice.

Climate The South Pole is the coldest, windiest, and driest place on Earth. The coldest temperature ever recorded on Earth was at the South Pole; it reached -88.0°C! On average, most of Antarctica gets less than 2 inches of snowfall each year.

People Although scientific expeditions visit Antarctica, there are no permanent human residents (because of the extreme weather, which includes freezing temperatures, strong winds, and blizzards). There are about 4,000 seasonal visitors to Antarctica.

Location Antarctica hasn't always been located at the South Pole. It has drifted, like the other continents, and has ranged from the equator (during the Cambrian period) to the South Pole. During the time of the dinosaurs (the Mesozoic Era), Antarctica was more temperate and housed dinosaurs and many other life forms. Now, there is very little indigenous life.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review, 96,* 703–719.

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metacognitive index. *Journal of Experimental Psychology: General, 127,* 55–68.

Bjork, R. A. (1975). Retrieval as a memory modifier. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 123–144). Hillsdale, NJ: Erlbaum.

Bjork, R. A. (1978). The updating of human memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 12, pp. 235–259). New York, NY: Academic Press.

Bjork, R. A. (2001). Recency and recovery in human memory. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 211–232). Washington, DC: American Psychological Association Press.

Bjork, R. A. (2011). On the symbiosis of learning, remembering, and forgetting. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A Festschrift in honor of Robert A. Bjork* (pp. 1–22). London, UK: Psychology Press.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 35–67). Hillsdale, NJ: Erlbaum.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64,* 417–444.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6,* 173–189.

Bouton, M. E., & Peck, C. A. (1992). Spontaneous recovery in cross-motivational transfer (counterconditioning). *Animal Learning and Behavior, 20,* 313–321.

Bower, G. H., & Reitman, J. S. (1972). Mnemonic elaboration in multilist learning. *Journal of Verbal Learning and Verbal Behavior, 11,* 478–485.

Briggs, G. E. (1954). Acquisition, extinction, and recovery functions in retroactive inhibition. *Journal of Experimental Psychology, 47,* 285–293.

Brown, A. S. (1976). Spontaneous recovery in human learning. *Psychological Bulletin, 83,* 321–328.

Castel, A. D. (2008). Metacognition and learning about primacy and recency effects in free recall: The utilization of intrinsic and extrinsic cues when making judgments of learning. *Memory & Cognition, 36,* 429–437.

Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.

Diaz, M., & Benjamin, A. S. (2011). The effects of proactive interference (PI) and release from PI on judgments of learning. *Memory & Cognition, 39,* 196–203.

Estes, W. K. (1955). Statistical theory of spontaneous recover and regression. *Psychological Review, 62,* 145–154.

Forrester, W. E. (1970). Retroactive inhibition and spontaneous recover in the A-B, D-C paradigm. *Journal of Verbal Learning and Verbal Behavior, 9,* 525–528.

Hull, C. L. (1943). *The principles of behavior*. New York, NY: Appleton-Century-Crofts.

Jacoby, L. L., & Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin, 13,* 314–336.

Jost, A. (1897). Die assoziationfestigkeit in ihrer abhangigheit von der verteilung der wiederholungen [The strength of associations in their dependence on the distribution of repetitions]. *Zeitschrist für Psychologie, 14,* 436–472.

Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgments. *Journal of Memory and Language, 35,* 157–175.

Koppenaal, R. J. (1963). Time changes in the strengths of A-B, A-C lists: Spontaneous recovery? *Journal of Verbal Learning and Verbal Behavior, 2,* 310–319.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 31,* 187–194.

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 133,* 643–656.

Kornell, N., & Bjork, R. A. (2009). A stability bias in human memory: Overestimating remembering and underestimating learning. *Journal of Experimental Psychology: General, 138,* 449–468.

Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease-of-processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. *Psychological Science, 22,* 787–794.

Lang, A. J., Craske, M. G., & Bjork, R. A. (1999). Implications of a new theory of disuse for the treatment of emotional disorders. *Clinical Psychology: Science and Practice, 6,* 80–94.

Little, J. L., Storm, B. C., & Bjork, E. L. (2011). The cots and benefits of testing text materials. *Memory, 19,* 346–359.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39,* 352–370.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation, 26,* 125–173.

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex.* London, UK: Oxford University Press.

Postman, L. (1971). Transfer, interference and forgetting. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth and Schlosberg's experimental psychology* (3rd ed., pp. 1019–1132). New York, NY: Holt, Rinehart & Winston.

Postman, L., Stark, K., & Fraser, J. (1968). Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior, 7,* 672–694.

Schwartz, B. L., Benjamin, A. S., & Bjork, R. A. (1997). The inferential and experiential basis of metamemory. *Current Directions in Psychological Science, 6,* 132–137.

Storm, B. C., Bjork, R. A., & Storm, J. C. (2010). Optimizing retrieval as a learning event: When and why expanding retrieval practice enhances long-term retention. *Memory & Cognition, 38,* 244–253.

Thorndike, E. L. (1914). *The psychology of learning.* New York, NY: Teachers College.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5,* 381–391.

Wheeler, M. A. (1995). Improvement in recall over time without repeated testing: Spontaneous recovery revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 173–184.

Woodworth, R. S., & Schlossberg, H. (1954). *Experimental Psychology* (2nd ed.). New York, NY: Holt.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology, 55,* 235–269.