JOURNAL OF MEMORY AND LANGUAGE 32, 1–24 (1993)

Remembering Mistaken for Knowing: Ease of Retrieval as a Basis for Confidence in Answers to General Knowledge Questions

COLLEEN M. KELLEY

Macalester College

AND

D. STEPHEN LINDSAY

University of Victoria, Victoria, British Columbia, Canada V8W 3P5

We propose that confidence in potential answers to general knowledge questions is based, in part, on the ease with which those answers come to mind. Consistent with this hypothesis, prior exposure to correct and to related but incorrect answers to general knowledge questions increased the speed, frequency, and confidence with which subjects gave those answers on a subsequent test of general knowledge. Similar effects were obtained even when subjects were warned that the list included incorrect answers (Experiment 2). The results of Experiment 3 indicated that the effects do not rely on deliberate search of memory for the list: Subjects who read a list with correct answers to half of the questions on a subsequent test gained full benefit of exposure to correct answers relative to subjects who read a list with correct answers to all of the questions, yet showed no cost on questions for which answers were not in the list relative to subjects who read a list of unrelated fillers. Finally, Experiments 4a and 4b demonstrated that prior exposure to incorrect answers can give rise to illusions of knowing even when subjects know that all of the answers on the study list were incorrect. In those studies, subjects were correctly informed that all of the answers on the list were incorrect, yet those who had studied the list with divided attention nonetheless tended to give the studied incorrect answers as responses to the knowledge questions. We discuss these findings in terms of Jacoby, Kelley, and Dywan's (1989) attributional approach to subjective experience.

What was Buffalo Bill's last name? If an answer to this question comes to mind, what leads you to be confident that it is the correct answer? The research presented here explores the hypothesis that the ease with which a potential answer comes to mind contributes to confidence in its accuracy. That is, we attempt to test the hypothesis that the easy generation of an answer during an attempt to respond to a general knowledge question contributes to feelings of confidence in that answer.

Most theories of question answering have focused on issues concerning the storage and retrieval of information rather than on the bases for confidence in the answers that are retrieved (e.g., Anderson, 1983; Graesser & Black, 1985; Reder, 1987). Although Lachman and Lachman (1980) pointed out that there must be metamemorial processes that evaluate the adequacy of retrieved and inferred answers, they offered no specific hypotheses about how such decisions are made. There have been

Portions of this work were supported by National Science Foundation Grant BNS-9109465, a Natural Science and Engineering Council of Canada postdoctoral fellowship to Steve Lindsay, and by Macalester College, the University of Victoria, and Williams College. Maureen Brand is acknowledged for her work on pilot versions of Experiment 1a, performed as part of an honors thesis at Williams College under the supervision of Colleen Kelley. Correspondence and reprint requests may be addressed to Colleen Kelley, Department of Psychology, Macalester College, St. Paul, MN 55105 (e-mail kelley@macalstr) or to Steve Lindsay, Department of Psychology, University of Victoria, P.O. Box 3050, Victoria, B.C., Canada, V8W 3P5 (e-mail lindsay@uvvm.uvic.ca).

0749-596X/93 $5.00

studies of confidence in comprehension (Glenberg & Epstein, 1985; Maki & Berry, 1984; Morris, 1990), judgments of knowing for recently learned information (Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Leonesio & Nelson, 1990), and feeling of knowing ratings when an answer cannot be produced (Brown & McNeill, 1966; Hart, 1965; Nelson & Narens, 1980), but relatively little research on confidence in answers given to general knowledge (fact retrieval) questions. Of the theoretical treatments of question answering that have focused on the issue of confidence in answers, most have emphasized the role of the amount of supporting information that can be generated: The more supporting evidence for an answer, and the less nonsupporting evidence, the higher the confidence in the accuracy of that answer (e.g., Koriat, Lichtenstein, & Fischhoff, 1980; Graesser & Hemphill, 1991).

Although the amount of evidence retrieved is surely an important basis for confidence, we propose that there is another important basis: the "fluency" or ease with which a potential answer comes to mind. We borrow the term "fluency" from Jacoby's work on bases for the subjective experience of remembering (e.g., Jacoby & Dallas, 1981; Jacoby, Kelley, & Dywan, 1989; Johnston, Dark, & Jacoby, 1985; Kelley & Jacoby, 1990). Jacoby and his colleagues suggested that the feeling of familiarity depends on an inference, analogous to Tversky and Kahneman's (1973) availability heuristic. Tversky and Kahneman proposed that people's estimates of probabilities are based on the ease with which they can think of examples. Similarly, Jacoby and his coworkers argued that the feeling of remembering arises from an unconscious inference about the ease with which recognition test items are perceived or the ease with which items come to mind during recall attempts (see also Baddeley, 1982). Consistent with this hypothesis that ease of processing provides a basis for the subjective experience of remembering, manipulations that facilitate processing of new distractor items on a memory test increase the likelihood that subjects will falsely "remember" those items (Jacoby & Whitehouse, 1989; Kelley, Jacoby, & Hollingshead, 1990; Lindsay & Kelley, 1991; Whittlesea, Jacoby, & Girard, 1990).

According to Jacoby, Kelley, and Dywan (1989), the way people interpret ease of processing is critically dependent upon their current goals and orientation. In the context of a memory test, people are likely to experience easy processing of test items as familiarity, because in that context memory is a salient source of easy processing. When the situation directs subjects to other judgments, however, the same ease of processing may be misattributed to other factors. For example, the familiarity of nonfamous names (e.g., "Sebastian Weisdorf") produced by reading them in an experiment can later be misinterpreted as fame (Jacoby, Kelley, Brown, & Jasechko, 1989; Jacoby, Woloshyn, & Kelley, 1989). In the first phase of those experiments, people read a list of names. In the second phase, those old names were mixed with new nonfamous and new famous names in a test of fame judgments. Relative to new nonfamous names, nonfamous names presented in the first phase were more often falsely identified as famous. That is, in the context of the fame judgment test, the familiarity due to having read the name on the list gave rise to a feeling that the name was a famous one. Other studies have shown that the effects of prior exposure on ease of processing can also be misattributed to other factors such as visual duration (Witherspoon & Allan, 1985), background noise level (Jacoby, Allan, Collins, & Larwill, 1988), or the difficulty of a problem (Kelley & Jacoby, in preparation).

The major purpose of the current experiments was to apply Jacoby, Kelley, and Dywan's (1989) attributional approach to the subjective experience of knowing or be-

lieving that an answer to a question is correct. We propose that one basis for confidence in an answer is the ease with which it is generated. Nelson and Narens (1990) also proposed that confidence in answers is partly determined by latency of retrieval, and they supported this with evidence that people have more confidence in answers that they retrieve quickly, whether those answers are correct or incorrect. Similarly, Morris (1990) found that the more quickly subjects retrieved information about a passage, the more confident they were that they could answer questions about that passage (cf. Begg *et al.*, 1989).

To date, there is only correlational evidence in support of the hypothesis that the fluency with which a potential answer comes to mind contributes to confidence in its accuracy. Unfortunately, such evidence is not very compelling, because by any theoretical account well-known answers should be produced more quickly than less-known answers (e.g., people will be both fast and confident when asked a well-known fact). Even if rapid retrieval were not a basis for confidence, but rather merely a by-product of the organization of semantic memory, one would still expect such correlations. To obtain evidence that ease of retrieval plays a causal role in creating confidence, one must experimentally manipulate ease of retrieval and observe its effect on confidence. That is what we have done in the experiments reported below.

One way to manipulate the speed of retrieving answers to questions is by recent prior exposure to those answers. For example, Kubovy (1977) found that subjects asked to report "the first one-digit number than comes to mind" were eight times more likely to report the number "one" than subjects asked to report "the first digit that comes to mind." That is, having just been exposed to the number "one" dramatically increased the likelihood of that number popping to mind during an attempt to generate a "random" number. More recently, Roediger and Blaxton (1987; Blaxton, 1989) and Hamann (1990) used general knowledge questions as an indirect test of memory for prior occurrence. In an indirect test, memory for a target event is inferred from its effects on performance of a task such as word fragment completion (for reviews, see Richardson-Klavehn & Bjork, 1988, and Schacter, 1987). Blaxton (1989) and Hamann (1990) found that prior exposure to correct answers improved performance on a later test of general knowledge. For example, having recently read the word "cheetah" on a list increased the likelihood that subjects would later produce the correct answer to the question, "What is the fastest animal on earth?"

These studies of the effects of prior exposure to answers to general knowledge questions did not address the issue of subjective experience, so it is not known whether the effect was experienced as knowing the answers, as remembering them from the study list, or as some combination. We propose that prior exposure increased the speed with which answers came to mind when questions were presented, which in turn led subjects to believe those answers to be correct. In the following experiments, we predicted that prior exposure to correct answers would increase the frequency and speed with which subjects gave those answers and, more importantly, that subjects would believe those answers to be correct. Even stronger evidence for our claim that the speed with which an answer comes to mind is a basis for confidence would be gained by changing what answers quickly come to mind. By our account, illusions of knowing should be produced by manipulations that cause incorrect answers quickly to come to mind. Specifically, in Experiments 1a and 1b we predicted that prior exposure to semantically related but incorrect answers would cause those answers to come readily to mind when a relevant question was read, such that subjects would confidently use

them as correct answers. The remaining experiments are designed to eliminate two rival hypotheses: that studied answers come to mind by way of intentional retrieval and that confidence in studied answers is based on recognition of answers as items from the list.

These studies go beyond previous research on knowledge tests as indirect measures of memory in that they provide evidence that the effect of prior exposure to a correct or incorrect answer is often experienced as simply knowing the answer. That is, we show that subjects not only give previously studied items as answers, but believe them to be correct. Further, Experiments 2, 3, 4a, and 4b provide converging evidence that the effects of prior exposure on question answering are truly automatic, in that they do not reflect intentional searches of memory for the study list or confidence in answers consciously recognized as study-list items.

Experiments 1a and 1b

Method

Overview. Subjects read a list of names of people, places, and things, and then took a general knowledge test. The acquisition list included the correct answers to one third of the test questions, related but incorrect answers to another third of the questions, and unrelated fillers. At test, subjects were instructed to give the correct answer to each question and rate their confidence in that answer. In Experiment 1a answer latency was measured by having the experimenter press the space bar when the subject began to say an answer. In Experiment 1b a voice key was used to measure answer latency.

Subjects. The subjects in Experiment 1a were 17 Williams College undergraduates who participated as part of a laboratory demonstration for a cognitive psychology course. Those in Experiment 1b were 24 University of Victoria undergraduates who participated for optional extra credit in an introductory psychology course. Subjects were randomly assigned to a study list according to a predetermined schedule.[1]

Materials. A set of 90 general knowledge questions (e.g., "What was Buffalo Bill's last name?") was constructed, drawing materials from Nelson and Narens (1980), encyclopedias, and trivia games. In addition to the correct answer (e.g., Cody), a related but incorrect answer (e.g., Hickock) and an unrelated filler (e.g., Melville) were assigned to each question. Three acquisition lists were constructed by randomly dividing the questions into three sets of 30: each acquisition list consisted of 30 correct answers, 30 related but incorrect answers, and 30 unrelated fillers, with assignment of particular items to these conditions counterbalanced across subjects. The acquisition list items and the general knowledge questions were presented in unique random orders for each subject. For Experiment 1b, some of the questions were reworded to standardize their length and some overly difficult questions were discarded, resulting in a set of 84 (28 in each condition).

Procedure. Experiment 1a was conducted on an Apple IIe using the A.P.T. II software package (Foltz & Poltrock, 1984) and Experiment 1b was conducted on an IBM ps/2 using Schneider's (1988) Micro-Experimental Laboratory. Subjects were tested individually. Subjects were told that a list of names of people, places, and things would appear one at a time in the center of the screen, and that they should read each item aloud and study it for a later memory test. One of the three acquisition lists was then presented, with each item appearing for 2 s and 1 s between items. Immediately thereafter, subjects were given the instructions for the general knowledge test. Subjects were told that questions would appear on the screen one at a time, and that they

[1] Of the 17 subjects in Experiment 1a, five were assigned to one list and six to each of the two remaining lists. There were no list effects on any of the dependent measures.

should answer each question and then rate their confidence in the accuracy of that answer. In Experiment 1a, each question trial was preceded by the work "READY?" which appeared in the middle of the screen for 1500 ms (this prompt was dispensed within Experiment 1b). Each question remained on the screen until the subject responded or 10 s elapsed. In Experiment 1a, as soon as the subject started to say an answer, the experimenter hit the space bar to stop the timer. In Experiment 1b, subjects were instructed to say the answer into a microphone which triggered a voice key. The confidence scale then appeared on the screen to prompt the subject to rate his or her confidence in the accuracy of the answer (1 = "not at all sure," 4 = "extremely confident"). After a 1-s pause, the next trial began.

Results and Discussion

Answers. The data are presented in Table 1. Separate one-way ANOVAs were performed on proportions of correct answers and related but incorrect answers as a function of item type (i.e., whether the correct, related, or unrelated answer had been presented in the first phase). An answer was coded as "related" only if it was the designated related item for that question (e.g., "Hickock," but not "Crocket"). The alpha level for all statistical tests was .05. For each analysis, we report the results from each of the two experiments.

As predicted, prior exposure had a significant effect on the proportion of questions answered correctly, $F(2,32) = 29.48$, $MS_e = .012$, and $F(2,46) = 39.88$, $MS_e = .012$. Newman–Keuls tests revealed that in both experiments correct answers were significantly more frequent on studied-correct items than on control items, and significantly less frequent on studied-related items than on control items. The absolute levels of performance were quite different in the two experiments, in that the Canadian students performed much more poorly in all condition. This is to be expected, because many of the questions concerned

TABLE 1
RESULTS OF EXPERIMENTS 1a AND 1b

Measure	Experiment	Item type: Correct	Filler	Related
Proportion correct	1a	.81[a]	.70[b]	.53[c]
	1b	.49[a]	.31[b]	.21[c]
Proportion related	1a	.05[b]	.09[b]	.31[a]
	1b	.06[b]	.09[b]	.29[a]
Confidence in correct	1a	3.54[a]	3.26[b]	3.39[b]
	1b	3.44[a]	3.12[b]	3.27[ab]
Confidence in related	1a	—	2.14[b]	2.77[a]
	1b	—	2.45[b]	3.02[a]
Median correct RT	1a	3071[a]	3674[b]	3507[b]
	1b	3242[a]	3857[a]	3853[a]
Median related RT	1a	—	4879[b]	3819[a]
	1b	—	5494[b]	3661[a]
Pearson *r*				
RT*Conf. in Correct	1a	−.43[a]	−.40[a]	−.44[a]
	1b	−.47[a]	−.50[a]	−.29[b]
RT*Conf. in Related	1a	—	—	−.44
	1b	—	—	−.37

Note. If two scores in a given row have different superscripts, they differ by Newman–Keuls at the .05 level. Confidence was rated on a 4-point scale.

United States trivia (e.g., capital cities, presidents, etc.) and Williams College is a much more selective school than the University of Victoria. The important finding for present purposes is that in both experiments prior exposure to correct answers facilitated performance on the general knowledge test, whereas prior exposure to related but incorrect answers impaired performance.

The nature of the impairment effect is made clear by an analysis of the proportion of questions for which the related but incorrect answers were given. As predicted, the proportion of such answers was affected by prior exposure, $F(2,32) = 41.15$, $MS_e = .008$, and $F(2,46) = 55.87$, $MS_e = .006$. Related but incorrect answers were significantly more frequent on studied-related items than on control items or studied-correct items. Thus recent exposure to incorrect answers led subjects to give them as correct answers on the general knowledge test.

Latency. In both experiments, there was a significant effect of prior exposure on time to correctly answer general knowledge questions, $F(2,32) = 14.86$, $MS_e = 148{,}542$, and $F(2,42) = 3.34$, $MS_e = 824{,}886$.[2] Newman–Keuls tests indicated that in Experiment 1a correct answers were produced significantly more quickly on studied-correct items than on control items or on studied-related items. These same effects were reliable in Experiment 1b as planned comparisons, although they were not significant by Newman–Keuls.

Because subjects rarely gave the related but incorrect answers in response to questions for which correct answers had been presented in the list, this level was excluded from the analysis of the effects of prior exposure on the latency of related but incorrect answers. Related but incorrect answers were given significantly more quickly on studied-related items than on control items, $F(1,14) = 5.30$, $MS_e = 1{,}695{,}713$, and $F(1,21) = 22.82$, $MS_e = 1{,}618{,}869$.[3]

Confidence. Prior exposure led answers to quickly pop to mind. By our account, the easy generation of answers is a basis for confidence, so prior exposure should increase confidence. In contrast, if confidence is not based on the ease with which ideas come to mind, but depends entirely on the organization of semantic memory, one might expect subjects to have lower confidence in answers that pop to mind due to prior exposure rather than due solely to their relation to stable knowledge. This is because correct answers that are accessible only with the boost in availability due to prior exposure may be less well known by the subject. The data accord with our prediction: In both experiments, prior exposure significantly increased confidence, $F(2,32) = 5.60$, $MS_e = .07$, and $F(2,42) = 4.67$, $MS_e = .117$. Newman–Keuls tests indicated that subjects were significantly more confident in their correct answers to questions for which correct answers had been presented on the list than on control items.

As in the analysis of the latency data, studied-correct items were excluded from the analysis of the effect of prior exposure on time to give related but incorrect answers. In both experiments, confidence in related but incorrect answers was significantly higher on studied-related items than on control items, $F(1,14) = 11.34$, $MS_e = 0.26$, and $F(1,21) = 9.73$, $MS_e = .369$. Even though subjects produced many more related answers in the studied-related condition than in the control condition, they were more, not less, confident in those responses.

Relation between confidence and latency. For each subject, Pearson product–moment correlations (r) were calculated for

[2] Two subjects in Experiment 1b were excluded from analyses of latency and confidence for correct responses: One gave no correct answers in the studied-related condition and one gave no correct answers in the studied-filler condition.

[3] In each experiment, two subjects were excluded from this analysis because they gave no related answers in the studied-filler condition.

the relationship between confidence ratings and latencies to produce correct answers for each type of item. These correlations between latency and confidence in correct answers were significantly negative in all three conditions in both experiments (by single-sample t tests). That is, as latency decreased confidence in the response increased. Importantly, the strength of the relationship was equivalent for studied-correct items and for studied-filler items. Thus, even though prior exposure altered what answers came to mind and how quickly, the relationship between latency and confidence remained the same. Our interpretation of this consistent relation between latency and confidence is that the ease with which an answer is generated is a basis for confidence in that answer. Note that in Experiment 1b the mean correlation between latency and confidence in correct answers was lower for studied-related items than for control items; this may be an artifact of the small number of correct answers given in that condition, but it might also reflect a genuine tendency for the latency–confidence relationship to be weaker on items for which subjects studied related answers but nonetheless give correct answers. In any case, the important finding is that although prior exposure to correct answers increased the frequency and speed with which those answers were later given, it did not alter the relationship between speed and confidence relative to controls. This is important because if the effect of prior exposure was mediated by recognition of answers as studied items (rather than by ease of generation), then the relationship between speed and confidence would be attenuated for studied items.

Finally, for each subject the correlation between latency and confidence was calculated for the studied-related items on which subjects gave the related answers. In both experiments, that correlation was significantly negative, $t(16) = 4.54$ and $t(20) = 4.02$.

Summary. The results of these experiments demonstrate that a single brief exposure to an item affects the way subjects later answer questions related to that item. Relative to control questions, prior exposure to correct answers led subjects to produce more correct answers, more quickly, and with greater confidence. Importantly, prior exposure to related but incorrect answers dramatically increased the frequency, speed, and confidence with which subjects produced those wrong answers. Finally, there were significant negative correlations between response latency and confidence, such that subjects were most confident in the answers that came to mind most quickly, even though prior exposure significantly affected the speed with which items came to mind.

Experiment 1b was conducted to ensure that the results of Experiment 1a were not an artifact of the technique used to measure response latencies (having the experimenter press the space bar as soon as the subject began to say an answer). The technique employed in Experiment 1a was used in the remaining experiments.

Experiment 2

We interpret the results of Experiments 1a and 1b as evidence that the easy production of an idea while trying to answer a question gives rise to confidence that the idea is the correct answer. That is, confidence in both correct and incorrect answers may be based on a metamemorial process that monitors the speed with which answers come to mind. A critic might argue, however, that the results reflect intentional recall of answers from the study list and a tendency to place confidence in answers recognized as study list items. Subjects in Experiments 1a and 1b were not informed about the relationship between the study list and the questions, but in debriefing they often reported that during the knowledge test they had recognized some of the answers as items from the study list. Thus it is possible that subjects deliberately searched memory of the study list during attempts to

answer questions and placed confidence in potential answers that they recognized as study list members.

Subjects in Experiment 2 were instructed to give the first answer that popped to mind when they read the question, regardless of its accuracy. That is, subjects were told to avoid laborious memory searches and decision processes, and to say whatever answer first popped to mind. Subjects were informed that the study list included equal numbers of incorrect but related answers and correct answers, and were warned that prior exposure to these answers might cause them to pop to mind when a question was read. Thus subjects were told that recognizing an answer as something from the study list was not a good basis for confidence in its accuracy. Because of the conflict engendered by asking subjects to report the first answer that came to mind even when that would sometimes be a wrong answer, subjects were given the option of rating an item as "definitely wrong" on a 7-point confidence scale (from "definitely wrong" to "definitely right"). We predicted that answers presented on the list would spontaneously come to mind in response to questions. Furthermore, we expected that despite being warned that the list contained incorrect answers subjects would be unable to escape the effects of prior exposure on the ease with which studied items came to mind, and would consequently believe them to be correct answers.

Method

Subjects. Subjects were 18 Williams College undergraduates who participated for optional extra credit in an introductory psychology course. Subjects were assigned to study list according to a predetermined schedule.

Materials and procedure. The materials were the same as those used in Experiment 1b, and the study phase procedure was the same as in Experiment 1b. At test, subjects were told that general knowledge questions would appear on the computer screen one at a time, and were instructed to give "the first answer that comes to mind." Subjects were told that the acquisition list included correct answers to one third of the questions and related but incorrect answers to another third: "Thus when an answer pops to mind you may sometimes recognize it as one that you saw on the list. However, that will tell you nothing about whether it is a correct answer or not, because we equally often presented correct answers and wrong answers." Subjects were instructed to rate their confidence in each answer on a scale from 1 (definitely wrong) to 7 (definitely right). Each question remained on the screen until the subject answered it (at which time the experimenter hit the space bar to record latency) or 10 s elapsed. The experimenter coded the subject's answer as correct, related, or other. Finally, the rating scale appeared on the screen, and the experimenter entered the subject's confidence rating. The next question was then presented.

Results and Discussion

Answers. The data are presented in Table 2. As predicted, there was a significant ef-

TABLE 2
RESULTS OF EXPERIMENT 2

Measure	Item type: Correct	Filler	Related
Proportion correct	.76[a]	.53[b]	.37[c]
Proportion related	.08[b]	.14[b]	.42[a]
Confidence in correct	6.20[a]	5.92[a]	5.83[a]
Confidence in related	4.68[a]	4.02[a]	4.84[a]
Median correct RT	2943[a]	3879[b]	3603[b]
Median related RT	4249[b]	4772[b]	3621[a]
Pearson *r*			
RT*Conf. in Correct	−.45[a]	−.39[a]	−.35[a]
RT*Conf. in Related	—	—	−.27

Note. If two scores in a given row have different superscripts, they differ by Newman–Keuls at the .05 level. Confidence was rated on a 4-point scale, with 1 = "not at all sure" and 4 = "extremely confident."

fect or prior exposure on the proportion of questions answered correctly, $F(2,34) = 78.14$, $MS_e = .009$. Newman–Keuls tests indicated that correct answers were significantly more often given on studied-correct items than on control items, and significantly less often given on studied-related items. Also as predicted, the proportion of related answers was increased by prior exposure to those incorrect answers, $F(2,34) = 73.88$, $MS_e = .008$. Newman–Keuls tests showed that the related but incorrect answers were significantly more often given on studied-related items than on control items or studied-correct items. Thus, despite the warning about the presence of incorrect answers on the study list, subjects often reported studied items as the first answer that came to mind.

Latency. There was a significant effect of prior exposure on time to produce correct answers, $F(2,34) = 14.81$, $MS_e = 280{,}997$. Newman–Keuls tests indicated that correct answers were given more quickly on studied-correct items than on control items and studied-related items, and that the latter two did not differ. Prior exposure also had a significant effect on time to produce related but incorrect answers, $F(2,34) = 5.64$, $MS_e = 1{,}060{,}673$. Newman–Keuls tests indicated that related but incorrect answers were given more quickly on studied-related items than on control items or studied-correct items, and that the latter two did not differ.

Confidence. Despite the instructions to give whatever answer first popped to mind and the warning that the list included incorrect answers, confidence ratings were nonsignificantly higher on studied-correct items than on control items or studied-related items, $F(2,34) = 2.72$, $MS_e = 0.25$, $p < .09$. Similarly, prior exposure to related answers led to a nonsignificant increase in confidence in those incorrect answers, $F(2,34) = 2.53$, $MS_e = 1.36$, $p < .10$. Despite the large increase in the frequency of related answers, confidence in those answers was somewhat higher on studied-related items than on control items or studied-correct items.

Comparing across experiments, the effects of prior exposure on confidence appear to be weaker in this experiment than in Experiments 1a and 1b. This makes sense, given that in this experiment subjects were specifically instructed to give the first answer that popped to mind, regardless of their confidence in it. Because of these instructions, the effects of prior exposure on answer production were even greater in this experiment than in Experiments 1a and 1b, but it appears that subjects were sometimes able to identify those answers as incorrect on some basis other than ease of generation (e.g., easy generation of incompatible evidence).

Relation between confidence and latency. As in Experiments 1a and 1b, as latency decreased confidence in the response increased (all mean *r*s significantly negative by single-sample *t* tests). Importantly, the strength of the correlation did not vary as a function of prior exposure, $F < 1$: Despite the large effects of prior exposure on what answers were given and the speed with which those answers were produced, the relation between latency and confidence did not change. Subjects also exhibited negative correlations between time to produce related but incorrect answers and confidence in those answers, $t(17) = 3.94$. We take this as further evidence that the speed with which an answer comes to mind is a basis for confidence in that answer.[4]

[4] Comparing across experiments, this correlation between latency and confidence in related answers was lower here than in Experiments 1a and 1b. Again, this makes sense, given that subjects were instructed to respond with the first answer that came to mind regardless of their confidence in its accuracy. The across-experiments comparison suggests that subjects sometimes realized that answers that came quickly to mind due to prior exposure were in fact not correct. This indicates that the speed with which a potential answer comes to mind is not the sole basis for confidence in its accuracy. Other factors, such as the amount of supporting or contradictory evidence subjects are able to retrieve, are also important in determining confidence. As we argue in the General Dis-

Summary. Prior exposure increased the probability that subjects would respond to questions by giving the studied items as answers and decreased response latencies for those answers, with no significant loss of confidence in the accuracy of those answers. These effects were obtained even though subjects had been instructed to respond with the first answer that popped to mind, and even though they had been warned that the list included related but incorrect answers to the test questions. This suggests that prior exposure automatically influences what ideas come to mind in response to general knowledge questions, and that subjects tend to believe that ideas that pop into mind are correct answers.

EXPERIMENT 3

Our claim is that having recently read an answer causes it to pop to mind automatically when a relevant question is later presented, and the easy production of an answer contributes to the subjective experience of confidence in its accuracy. Thus it is important to our account that the effects of prior exposure on question answering are not due to intentional search of memory of the study list. This has also been an important issue in recent research on indirect tests of memory; how can one be sure that effects of prior exposure on an indirect test do not reflect intentional, aware uses of memory (e.g., Jacoby & Kelley, 1991; Richardson-Klavehn & Bjork, 1988)?

Experiment 3 was designed to test whether the effect of prior exposure in our procedure is due to deliberate searches of memory for the list rather than to automatic facilitation of the ease with which those answers come to mind. The design for Experiment 3 was inspired by a study reported by Ross, Ryan, and Tenpenny (1989, Experiment 3) concerning the use of previously encountered clues in solving riddles. Subjects in that paradigm often report that during attempts to solve riddles the previously studied solutions pop to mind, rather than being deliberately retrieved. To test this claim, Ross *et al.* (1989) conducted a study in which subjects were informed that "some" answers to riddles had been presented in a previously studied set of sentences. In one condition the set of sentences included answers to all of the riddles, whereas in another condition the sentences included answers to all of the riddles, whereas in another condition the sentences included answers to only half of the riddles. Ross *et al.* argued that if subjects in the half-answer condition attempted to deliberately search memory for the sentences when trying to solve a riddle, then if they obtained full benefit on the riddles for which answers were presented (relative to the all-answer condition) they should also suffer some cost for riddles for which answers were not presented (relative to a no-answer control group). In contrast, if previously studied answers popped to mind automatically, then one might expect benefit on riddles for which answers had been presented without cost on riddles for which answers had not been presented. Consistent with this automatic process, Ross *et al.* found benefit but no cost for the group that received clues to half the riddles.

In the present study, subjects were tested in one of three conditions that varied in terms of the proportion of correct answers presented in the study list. Subjects in all three conditions read an 84-item list and then took an 84-item general knowledge test. In the All Correct condition, the list consisted of the correct answers to all 84 questions. In the Half-&-Half condition, the list contained the correct answers to 42 questions randomly mixed with 42 unrelated fillers. At test, subjects in both of these conditions were told that the correct answers to some of the questions had been presented in the list. In the All Filler condition, the study list consisted of 84 unrelated fillers, and at test subjects were in-

cussion, however, confidence in retrieved evidence may also be partly based on the ease or fluency with which that evidence is generated.

formed that there was no relationship between the study list and the knowledge test.

What pattern of results would be expected if the effects reported in the first three experiments relied on deliberate search of memory for items on the study list? One possibility is that subjects in the Half-&-Half condition would be less likely than those in the All Correct condition to use this strategy, in which case they would perform less well on questions for which the correct answers had been presented. Alternatively, subjects in the Half-&-Half condition might be equally likely as those in the All Correct condition to search memory of the study list for answers. If so, they would obtain equal benefit on questions for which the correct answer appeared on the study list, but they would do more poorly (e.g., answer more slowly) than subjects in the All Filler condition on questions for which the correct answer had not been presented. This is because they would be devoting resources to searching memory of the study list for an answer that was not presented, rather than simply trying to answer the question. In contrast, we propose that prior exposure to relevant answers causes those answers automatically to come readily to mind when an appropriate question is read, and so we expected that performance on the studied-correct items would be as good in the Half-&-Half condition as in the All Correct condition (i.e., equal benefits of prior exposure) and that performance on studied-filler items in the Half-&-Half condition would not be lower than performance in the All Filler condition (i.e., no cost of no prior exposure).

Method

Subjects. The subjects were 36 Williams College undergraduates who participated for optional extra credit in an introductory psychology course. Subjects were randomly assigned to condition according to a predetermined schedule.

Materials and procedure. The 84 items used in Experiment 1b were randomly divided into two sets of 42, and unrelated fillers were selected for each. The subjects in the Half-&-Half condition studied correct answers for one set and fillers for another, intermixed in a random order. Assignment of items to condition was counterbalanced across subjects in the Half-&-Half condition. The study list was presented as in Experiment 2. Immediately after presentation of the list, subjects were given the general knowledge test instructions. Subjects in the All Correct and Half-&-Half conditions were told that "The correct answers to some of the questions you are about to answer were presented in the study list." Subjects in the All Filler condition were told that there was no relationship between the study list and the questions. All subjects were instructed to answer each question and then rate their confidence in that answer (1 = "definitely wrong," 7 = "definitely right"). The questions were presented one at a time in the center of the screen. The experimenter pressed the space bar to record latency as soon as the subject began to say an answer, entered a code for that answer (correct, related, or unrelated), and then entered the subject's confidence rating.

Results and Discussion

Answers. The mean proportions of questions in each condition correctly answered within the 10-s limit are presented in Table 3. A one-way repeated-measures ANOVA revealed that, as predicted, subjects in the Half-&-Half condition more often answered questions for which the correct answer had been presented on the list than questions for which the correct answer had not been presented, $F(1,11) = 67.10$, $MS_e = .006$. More importantly, separate one-way between-subjects ANOVAs showed that subjects in the Half-&-Half condition benefitted from prior exposure to the correct answers as much as subjects in the All Correct condition, $F < 1$. Further, subjects in the Half-&-Half condition performed as well on studied-filler items as did subjects

TABLE 3
RESULTS OF EXPERIMENT 3

	Condition			
		Half-&-Half		
	All correct	Correct	Filler	All filler
Proportion correct	.73[a]	.70[a]	.44[b]	.48[b]
Confidence in correct	6.54[a]	6.51[a]	6.00[b]	6.06[b]
Median correct RT	2771[a]	2856[a]	3885[b]	3613[b]
Pearson *r*				
RT*Confidence in correct	−.42[a]	−.41[a]	−.53[a]	−.42[a]

Note. If two scores in a given row have different superscripts, they differ at the .05 level. Confidence was rated on a 7-point scale, with 1 = "not at all sure" and 7 = "extremely confident."

in the All Filler condition, $F < 1$. That is, subjects in the Half-&-Half condition gained full benefit of prior exposure (relative to the All Correct condition) without suffering any cost of non-prior exposure (relative to the All Filler condition). These data indicate that little if any of the effect of prior exposure in this procedure depends on intentional search of memory of the list.

Latency. In the Half-&-Half condition, median latencies for correct answers were significantly faster on studied-correct items than on control items, $F(1,11) = 10.46$, $MS_e = 607{,}419$. Median latencies for correct answers on questions for which correct answers had been presented in the Half-&-Half condition were equivalent to those in the All Correct condition, $F < 1$. Median latencies for correct answers on questions for which fillers were presented in the Half-&-Half condition were equivalent to those in the All Filler condition, $F < 1$.

Confidence. Consistent with our earlier findings, in the Half-&-Half condition mean confidence ratings on correct answers were higher on questions for which the correct answer had been presented in the study list than on questions for which the correct answer had not been presented, $F(1,11) = 6.40$, $MS_e = .25$. Confidence in correct answers on questions for which correct answers had been presented in the study list in the Half-&-Half condition was equivalent to that in the All Correct Condition, $F < 1$, and confidence ratings on correct answers for which fillers had been presented in the Half-&-Half condition were equivalent to those in the All Filler condition, $F < 1$.

Relation between confidence and latency. As in the earlier experiments, subjects tended to be more confident in answers that they produced quickly. More important, the strength of this relationship did not vary as a function of condition. Prior exposure to correct answers had dramatic effects on accuracy and latency, but did not alter the relationship between confidence and latency.

Summary. The results of Experiment 3 indicate that little if any of the effect of prior exposure in this procedure is due to subjects intentionally searching memory of the list during question answering. If the effects of prior exposure depended on intentionally searching memory of the list, one would expect either lack of full benefit on items for which correct answers were presented or cost on items for which answers were not presented in the Half-&-Half condition. Instead, prior exposure to correct answers led subjects to produce more correct answers, more quickly, and with greater confidence, and these benefits were gained without significant cost on items for which correct answers were not presented. Of course, deliberate search of memory for the list may have made a small contribution to the effects obtained here.

Consistent with this possibility, inspection of Table 3 reveals a pattern in which the Half-&-Half condition is slightly worse than the All Correct condition on studied-correct items, and slightly worse than the All Filler condition on studied-filler items. However, none of the differences approached statistical significance. It is likely, however, that deliberate searches of memory of the study list would have large and reliable effects under other conditions (e.g., if the questions were more difficult or if the list was shorter, as in Hamann, 1990). We think it likely that even with our procedure subjects do occasionally perform deliberate searches of memory of the study list when an answer does not quickly come to mind; the rarity of such searches is reflected in the results of Experiment 3, but if we increased the power of the experiment, we might detect reliable evidence of deliberate searches of memory of the list. It is clear, however, that such searches cannot account for the large and dramatic effects of prior exposure obtained in this and our other experiments.

EXPERIMENT 4a

As noted above, during debriefing subjects in Experiments 1a and 1b often reported that they had spontaneously recognized some answers as studied items after those answers came to mind. That is, the answer popped to mind in the course of reading and thinking about the general knowledge question, and then it was recognized as an item presented on the study list. This opens the possibility that effects of prior exposure on question answering are not mediated by speed of generation, as we have proposed, but by recognition of generated answers as study list items. That is, the automatic popping to mind of studied items would be due to prior exposure, but the confidence in those items might be due to a belief that items from the study list are likely to be correct. We have already presented evidence against this possibility in Experiment 2, in which subjects tended to give studied items as correct answers even though they were warned that the study list included incorrect as well as correct answers. Experiment 4a was designed to provide further evidence that the effects of prior exposure on confidence are not mediated by recognition of answers as study list members.

To eliminate the possibility that recognition of studied answers could contribute to confidence in the accuracy of those answers, we placed recognition memory in opposition to ease of generation as a basis for confidence (Jacoby, 1991; Jacoby, Woloshyn, & Kelley, 1989). At test all subjects were correctly informed that the study list had included only related but incorrect answers and unrelated fillers, and that they should not use anything from the study list as an answer. Thus subjects knew that any potential answer that they recognized as a list member was wrong and should not be given as an answer. Given this rule about the relation between recognition and status of the answer as wrong, can prior exposure nonetheless influence how questions are answered when recognition fails? Our prediction was that prior exposure would sometimes cause an item to pop to mind without subjects recognizing it as an item from the study list. To the extent that previously studied items pop to mind without being recognized as list members, we would expect subjects to give those items as correct answers.

To reduce later recognition, some subjects performed an attention-demanding task while studying the list, whereas others were allowed to devote full attention to the list. Divided attention greatly reduces list recognition, but leaves some effects of prior exposure intact (Jacoby, 1991). For example, in the false fame paradigm, Jacoby, Woloshyn, and Kelley (1989) found that nonfamous names gained as much familiarity from prior exposure under conditions of divided attention as under full at-

tention. Thus, provided that subjects processed the appropriate meanings of the items during study (e.g., identified "Hickock" as the name of an American frontiersman), we would expect that items studied under divided attention would tend to pop to mind when relevant questions were later presented. Moreover, we would expect that subjects in the divided attention condition would sometimes fail to recognize answers that came easily to mind due to prior exposure as items from the list.

We expected the effects of prior exposure to be much smaller in this experiment than in its predecessors. By an attributional account, one would expect that subjects given the opposition instructions would tend to interpret the easy production of an answer as potential evidence that it had been on the study list (because they are oriented toward the past). Thus the opposition instructions may make subjects very cautious about accepting ideas that pop to mind as correct answers, even when those ideas are not recognized as studied items. Note that the opposition instructions are much stronger than the warning given in Experiment 2; subjects are not merely warned that the study list included wrong answers, but rather are told that every single answer on the study list was wrong and should not be used. Much of the effects of prior exposure in the earlier experiments may have been accompanied by (without being caused by) recognition of answers as items from the study list. Under the opposition instructions, in contrast, subjects should not use an answer they recognize from the study list even if they think that answer is correct. That is, even if the ease with which an answer came to mind led a subject to feel that it was the correct answer, he or she would know that it was in fact wrong if it was recognized as a study list item. Thus results obtained using the opposition instructions likely underestimate the contribution of automatic influences of memory under conditions where subjects are not oriented toward the past as a misleading influence. Importantly, any effect of prior exposure obtained under these conditions can be taken as strong support for our hypothesis that the ease with which potential answers come to mind contributes to confidence in their accuracy.

The full attention condition serves as a check on subjects' ability to understand and follow the opposition instructions. Full attention at study would lead to superior memory for the studied answers. If subjects did not understand and follow the injunction against using answers recognized from the study list, then those in the full attention condition would use studied answers more often than subjects in the divided attention condition. In contrast, we predicted that subjects would attempt to follow the instructions but, because divided attention during study hampers subsequent recognition, that those in the divided attention condition would sometimes fail to recognize answers that popped easily to mind due to prior exposure as studied items, and so they would believe them to be correct answers.

Method

Subjects. The subjects were 36 Macalester College undergraduates who participated for optional extra credit in an introductory psychology course or for a nominal payment. Subjects were randomly assigned to a condition according to a predetermined schedule.

Materials and procedure. The 84 items used in Experiment 1b were randomly divided into two sets of 42. Each subject studied incorrect but related answers for one set and fillers for another, intermixed in a random order, with assignment of items to condition counterbalanced across subjects. To increase the likelihood that subjects in the divided attention condition would encode the appropriate meaning of the items, each item on the study list was paired with a category label or related term, e.g., "composer—Brahms" (Richardson-Klavehn & Bjork, 1987, p. 514). Subjects

were instructed to read the terms aloud and think about their meanings, and were informed that their memory for the words would be tested. Each item and its corresponding category term was presented in the center of the screen for 3 s, with 500 ms between items. Subjects in the divided attention condition studied the items in the first phase while listening for runs of three odd-numbered digits in a long string of tape-recorded numbers (Craik, 1982; Jacoby, Woloshyn, & Kelley, 1989). A total of 22 runs of three odd digits occurred within a list of 252 random numbers presented at a 1-s rate. Subjects were to tap the table each time they heard a sequence of three odd digits in a row.

Immediately after presentation of the list, subjects were given the general knowledge test instructions. Subjects in both conditions were correctly informed that the list contained only incorrect answers: "For one half of the questions, an incorrect but related word was presented on the list you just studied. So, when an answer pops to mind, you may sometimes recognize it as one that you saw on the list. If that is so, then you'll know that the answer is incorrect, so don't use it as your answer." These instructions were repeated and rephrased several times to ensure that subjects understood that any answer presented on the study list was wrong. Questions appeared one at a time in a random order in the center of the screen and remained there until the subject responded (at which time the experimenter hit a key) or 10 s elapsed. Subjects were then prompted to rate their confidence in the accuracy of their answer on a 7-point scale (1 = "definitely wrong" and 7 = "definitely right"). There was a 500-ms pause between questions.

Results and Discussion

The opposition instructions change the task from one of answering questions to a mix of answering questions and attempting to recognize items from the list so as to apply the rule that recognized items are not the correct answers. For example, the reaction time to produce an answer in this study is not a measure of the speed of generating an answer, but a mix of generation time and recognition time. Because of this change of task, and because the study list included only incorrect answers and fillers, our interest in this experiment focuses on the frequency with which subjects gave the related but incorrect answers. Those data are shown in Table 4. A mixed-models ANOVA, with full versus divided attention at study as the between-subjects factor and item type (studied-related vs. studied-filler) at the within-subjects factor, was performed on the proportion of questions for which subjects gave the related but incorrect answer. The most important finding is the reliable interaction between attention condition and item type, $F(1,34) = 6.20$, $MS_e = .003$. Simple effects analyses revealed that, as predicted, subjects in the divided attention condition more often gave

TABLE 4
RESULTS OF EXPERIMENT 4a

	Attention condition and item type			
	Full attention		Divided attention	
	Related	Filler	Related	Filler
Proportion related	.07[a]	.09[a]	.12[b]	.08[a]
Confidence in related	4.76[a]	4.91[a]	4.10[a]	4.69[a]
Median related RT	5824[a]	6120[a]	5344[a]	5125[a]

Note. If two scores in an attention condition in a given row have different superscripts, they differ at the .05 level by simple effects. Confidence was rated on a 7-point scale, with 1 = "not at all sure" and 7 = "extremely confident."

the related but incorrect answers presented on the study list than those that were not, $F(1,34) = 6.03$, $MS_e = .003$. In contrast, subjects who studied the list with full attention showed no tendency to more often give incorrect answers that were on the study list than incorrect answers that were not, $F < 1.2$.

There was a nonsignificant tendency for subjects in the divided attention condition to produce related answers more quickly than subjects in the full attention condition, $F(1,34) = 6.03$, $MS_e = .003$, $p < .08$. However, the increased production of related but incorrect answers on studied-related items in the divided attention condition does not appear to be due to a speed/accuracy trade-off. Importantly, item type had no effect on time to produce related answers, nor did item type and attention condition interact (both Fs < 1). Thus the nonsignificant tendency for subjects in the full attention condition to respond more slowly than those in the divided attention condition does not compromise interpretation of the primary analysis: They responded more slowly on questions for which no related information had been presented on the study list as well as on questions for which the related but incorrect answer had been presented.

Neither attention condition nor item type had a reliable effect on confidence in related but incorrect answers, and these two factors did not interact (all Fs < 1.88, all ps $> .18$). The small number of related but incorrect answers given (due to the opposition instructions) precludes analyses of the correlations between latency and confidence in those answers. As noted, latency in this experiment partially reflects time to deliberately search memory of the list, so we would expect that the relationship between latency and confidence would be attenuated.

Summary. As predicted, prior exposure to incorrect answers had a reliable effect even when subjects knew that all of the answers on the study list were wrong and should not be used as answers. The frequency of related but incorrect answers was much lower in this experiment than in the earlier studies, due to the opposition instructions. The finding that subjects in the full attention condition did not tend to give the related but incorrect answers from the list indicates that subjects understood that they were not to use items from the study list. Nonetheless, prior exposure led to a 50% increase in the frequency of these errors in the divided attention condition.

Experiment 4b

In view of the small number of related but incorrect answers in Experiment 4a, we thought it important to replicate the effect obtained in the divided attention condition. The procedure differed from Experiment 4a in several minor ways, but in the main it was a straightforward replication of the divided attention condition. The auditory digit-monitoring task was presented as the primary task during the acquisition phase rather than as a secondary task. To ensure that subjects nonetheless identified the appropriate meanings of the items during the acquisition phase, they made yes/no judgments concerning the relatedness of each item and the category term with which it was paired during the acquisition phase (e.g., frontiersman—Hickock). As in Experiment 4a, at test subjects were correctly informed that only wrong answers had been presented in the study list, and they were told not to give items from the study list as answers. We predicted that prior exposure under divided attention would sometimes cause those answers to pop to mind without subjects recognizing them as items from the list and so lead subjects to believe those answers to be correct.

Method

Subjects. The subjects were 18 undergraduate students at the University of Victoria, who participated for optional extra credit in an introductory psychology course.

Materials. A few of the questions used in Experiment 4a were dropped and replaced, and 10 new questions were added, producing a set of 94 knowledge questions. There were two 94-item study lists, each consisting of the related but incorrect answers to 47 of these questions intermixed with 47 unrelated fillers. Each related but incorrect answer on the study list was paired with a category term or close associate; filler items were paired with unrelated filler terms (e.g., clock—ape). In addition, seven related and eight unrelated pairs were used as practice items; none of these was related to any of the knowledge questions. The tape recording of random digits was the same as that used in Experiment 4a.

Procedure. Subjects were told that their primary task during the first phase of the experiment was to detect runs of three odd digits in a tape-recorded list of random digits (saying the word "now" each time they heard a run). Subjects were told that while doing the number-monitoring task, they would also be required to judge the relatedness of pairs of items that would appear on the computer screen. They were to press a key labelled "Yes" if there was a clear and obvious relationship between the items in a pair (e.g., "tree—oak") and press a key labelled "No" if there was not a clear and obvious relationship between the items (e.g., "clock—ape"). Subjects then practiced these tasks: they performed the number-monitoring task alone for approximately 30 s, after which the relatedness-judgments task was added and 15 pairs were presented while subjects continued to perform the number-monitoring task. Pairs were presented for 3 s, with 1 s between pairs. If subjects made an incorrect relatedness judgment or failed to respond within 3 s, the computer produced a quiet beep. After this practice phase, the tape was stopped and the instructions were reviewed. Thereafter subjects again began by performing the number task alone for a few seconds, then the relatedness-judgment task was added. After the 94 items had been presented in the context of the relatedness-judgment task, the number tape was stopped and subjects were given the knowledge test instructions. As in Experiment 4a, these instructions emphasized that only wrong answers to the test questions had appeared in the first phase and explicitly told subjects *not* to give items from the first phase as answers on the test.

Results

As in Experiment 4a, prior exposure to related but incorrect answers under conditions of divided attention increased the frequency with which subjects gave those incorrect answers as responses on the knowledge test. Despite the opposition instructions, related but incorrect answers were more often given when those answers had appeared on the list (mean proportion = .18) than when they had not (M = .12), $F(1,17) = 6.06$, $MS_e = .004$. Response latencies for related but incorrect answers were slightly but not significantly faster on studied-related items (mean of medians = 4746 ms) than on fillers (4965 ms), $F < 1$. Likewise, although prior exposure increased the proportion of questions for which subjects gave the related but incorrect answers, it did not lower confidence in those answers (M = 4.45) relative to controls (M = 4.32), $F < 1$. Thus the primary effect (a 50% increase in the frequency of related but incorrect answers) was not compromised by trade-offs with latency or confidence. As in Experiment 4a, the small number of observations per subject precluded an analysis of the correlation between latency and confidence in related answers.

Summary

As noted in the Introduction to these experiments, it is likely that the findings of Experiments 4a and 4b dramatically underestimate the effects of automatic, unaware influences of memory when the opposition instructions are not used. The opposition instructions promote an analytic and cau-

tious attitude toward answers that pop to mind (such that subjects may be less susceptible to unconscious influences of memory). Further, the opposition instructions lead subjects to withhold any answer recognized as a list member, even if that answer popped to mind when the question was read and even if it feels like the correct answer to them. Put differently, the current results show that even when people specifically try to avoid the effects of prior exposure on feelings of confidence, under some conditions they cannot do so.

The relatively small size of this effect of memory without awareness does not imply that the larger effects obtained in the earlier experiments were mediated by aware recognition of study list items. In the earlier experiments a subject might recognize an answer that came to mind as an item from the study list, yet have confidence in that answer because it came to mind rather than because they recognize it as a study list item (as suggested by the results of Experiment 2). That is, spontaneous recognition of study list items might occur without having any causal effect on confidence. Further, the likelihood of aware recognition of study list items was probably greater in Experiments 4a and 4b than in the earlier experiments, because subjects were specifically instructed to attempt to recognize ideas that came to mind as list members. The important point for present purposes is that even under these conditions prior exposure reliably contributed to misplaced confidence in related but incorrect answers; under other conditions we would expect that the magnitude of these effects would be considerably greater.

General Discussion

The results of the six experiments reported here provide strong support for the hypothesis that confidence in potential answers to general knowledge questions is based in part on the ease with which those answers come to mind. Retrieval latency for correct and for incorrect answers was negatively correlated with confidence in those answers, consistent with Nelson and Narens (1990). Our research goes beyond these correlational analyses by showing that manipulations that lead answers to pop quickly to mind also lead to confidence in those answers. Exposure to correct and to related but incorrect answers caused them to come readily to mind when relevant questions were read, which in turn led subjects confidently to report them as correct answers. This effect was obtained even when subjects were informed that the study list included incorrect answers, suggesting that the effect is not mediated by misplaced confidence in items recognized as study list members. Furthermore, the effect of prior exposure appeared to be relatively automatic, rather than mediated by intentional search of memory of the list: Subjects who read a list of fillers mixed with correct answers to half of the questions on a subsequent test gained full benefit of exposure to correct answers relative to subjects who read a list of correct answers to all of the questions, yet showed no cost on items for which the answers were not presented relative to subjects who read a list of unrelated fillers. Finally, prior exposure led subjects to give incorrect answers even when recognition of an item from the list was placed in direct opposition to ease of retrieval as a basis for confidence by informing subjects that all of the items on the list were incorrect. In the following we discuss the implications of these findings and their relationship to other research on memory and question answering.

The hypothesis that confidence is based on the ease with which potential answers come to mind should be viewed as a complement to models of answer retrieval (e.g., Anderson, 1983; Graesser & Black, 1985; Graesser & Bower, 1990; Reder, 1987) rather than as an alternative to them. Such models can be described as accounts of the fact that the ideas that pop into mind during attempts to answer questions very often are correct answers. Our results demonstrate,

however, that stable general knowledge is not the only factor that causes particular ideas to come to mind when people answer questions, and they suggest that confidence is based on the ease with which ideas come to mind rather than on knowledge structure per se. A single brief exposure to the word "Hickock" caused subjects who would have correctly said "Buffalo Bill's last name was Cody" to instead say, with considerable confidence, "Buffalo Bill's last name was Hickock." Clearly, this effect is not independent of stable general knowledge—it is partly by virtue of subjects' knowledge of American frontiersmen that the studied item "Hickock" comes to mind when the question about Buffalo Bill is encountered. Just as clearly, however, stable knowledge is not the only determinant of what comes to mind.

One might argue that prior exposure changes the "strength" of a representation of a fact, and it is that change in strength that is responsible for later giving the item in answer to a question. People do not have subjective experience of "knowledge structures" per se but rather of the products of processes involving such structures (e.g., Nisbett & Wilson, 1977). By our view, what it means to strengthen a representation is to increase the speed with which it is retrieved given appropriate cues.

Confidence Based on Retrieval of Supporting Information

Nelson and Narens (1990) argued that although rapid retrieval of an answer contributes to confidence in its accuracy, speed of retrieval is not the only basis for confidence. In support of this claim, they found that, at any given latency, confidence was higher for correct answers than for incorrect answers. It is likely that confidence is often influenced by the amount of evidence retrieved in support of an answer (e.g., Koriat *et al.*, 1980; Glucksberg & McCloskey, 1981; Graesser & Hemphill, 1991). We propose that the contribution of supporting evidence to confidence is also influenced by the ease with which that evidence comes to mind. For example, if "Hickock" comes to mind in response to a question about Buffalo Bill's last name, a subject might double-check that answer by using it as a retrieval cue—perhaps finding that ideas and images relevant to the Wild West quickly and easily come to mind and consequently feeling more confident that Hickock is the correct answer. Thus the ease and speed with which supporting evidence comes to mind may also contribute to confidence.

The notion that confidence in a potential answer can be bolstered by the easy production of related ideas suggests a link between our account of confidence in generated answers and discussions of the feeling of knowing in tip-of-the-tongue states (e.g., Brown & McNeil, 1966; Hart, 1965; Nelson & Narens, 1980). People in the tip-of-the-tongue state cannot retrieve the target information, but do retrieve other closely related information. Perhaps it is the easy production of related ideas and of information about syllabic structure, orthographic features, etc., that gives rise to the subjective feeling of knowing the to-be-recalled word, even though the word itself does not come to mind.

Confidence in Other Domains

Jacoby, Kelley, and Dywan (1989) proposed that easy processing in the context of an attempt to remember gives rise to a feeling of familiarity. Easy processing can also be misinterpreted as being due to characteristics of a current task. For example, in the context of an indirect memory test such as an anagram-solving task, the facilitating effects of having recently seen the solution word may be misinterpreted as the anagram being an easy one (Jacoby & Kelley, 1987). In the current experiments, effects of prior exposure to an answer were misinterpreted as evidence of the accuracy of that answer. Other psychologists have discussed related effects of availability on tests of

knowledge and judgment. For example, a number of researchers (Arkes, Hackett, & Boehm, 1989; Bacon, 1979; Begg, Armour, & Kerr, 1985; Begg & Armour, 1991; Hasher, Goldstein, & Toppino, 1977) have shown that repeated exposure to statements about obscure topics (e.g., the average body temperature of chickens) can increase people's belief in those statements. Begg, Anas, and Farinacci (1991) presented evidence that these effects occur because people misattribute the familiarity of facts due to prior exposure to those facts being true. Similarly, Pollard (1982) and Begg (1991) argued that atmosphere and content effects in syllogistic reasoning are also misattributions: conclusions that come readily to mind or are familiar are accepted as true (see also Koehler, 1991). Likewise, phenomena such as hindsight bias (Fischhoff, 1975), the knew-it-all-along effect (Wood, 1978), and reconstructive memory errors (e.g., Bartlett, 1932; Ross, 1989) may reflect memory attributions regarding the ease with which ideas come to mind. In all of these cases, subjects who are attempting to remember something may accept ideas that come readily to mind as the to-be-remembered ideas, even though other factors (e.g., an intervening learning experience, general knowledge, etc.) may be affecting what comes to mind. The Moses illusion (Erickson & Mattson, 1981) can also be discussed in terms of the easy production of answers. The Moses illusion occurs when people answer a question such as "How many animals of each kind did Moses take on the ark?" (two) without noticing that the question contains an error (Moses had the ark of the covenant, but did not invite animals into it). One's aim when answering a question is to produce an answer, and the non-name information in the question causes the answer to quickly and easily come to mind, despite the presence of the mismatching name. The quickly and easily generated response is therefore given as the correct answer.

Attributions and Direct versus Indirect Memory Tests

One of the starting points of our research in this area was interest in the relationship between direct tests of memory (e.g., recognition tests), on which subjects are directly asked to remember past events, and indirect tests of memory (e.g., fragment completion), on which subjects are not explicitly instructed to remember past events (Richardson-Klavehn & Bjork, 1988; Schacter, 1987). A number of memory researchers have expressed concern that the effects of prior exposure on indirect test performance may be mediated by intentional, aware remembering (e.g., Bowers & Schacter, 1990; Gardiner, 1988; Jacoby, 1991; Johnson & Hasher, 1987; Richardson-Klavehn & Bjork, 1988). For example, subjects might deliberately attempt to recall items from the study list while working on a putatively indirect fragment completion test. Indirect tests may be particularly likely to be contaminated by deliberate attempts to recall when they involve highly distinctive and memorable materials, a brief delay between study and test, and difficult test items, as in the present studies and those reported by Blaxton (1989) and (especially) Hamann (1990). Although deliberate searches of memory for studied items undoubtedly contribute to performance on some indirect tests, the results of Experiments 2 and 3 indicate that intentional retrieval played little if any role in the effects reported here. Large effects of prior exposure were obtained in Experiment 2, even though subjects were explicitly instructed to respond with the first answer that popped to mind and were informed that the study list was not a reliable source of correct answers. Experiment 3 provided more compelling evidence against deliberate searches of memory for the studied items: Our finding of full benefit with no cost in the Half-&-Half condition indicates that the effects of prior exposure to an-

swers on later question answering were automatic and unintentional rather than mediated by deliberate attempts to recall list items.

A related issue concerns the subjective experience that accompanies performance on an indirect test. Do subjects feel that they are remembering a list item, or do they simply feel that they are answering a question? This distinction between aware and unaware uses of memory is different from that between deliberate searches of memory and automatic popping to mind (Kelley & Jacoby, 1990; Schacter, 1989). For example, a studied item could pop to mind automatically and then be recognized as a studied item—an automatic and aware use of memory. To rule out aware remembering on indirect memory tests, researchers have attempted to hide the relation between study and test or have eliminated subjects who retrospectively report aware remembering on the indirect test (Bowers & Schacter, 1990). A much more powerful approach is to place aware and unaware uses of memory in opposition, as in Jacoby, Woloshyn, and Kelley's (1989) false fame studies and in our Experiments 4a and 4b. Even though subjects in those experiments were correctly informed before taking the test that the study list contained only related but incorrect answers, those who had studied that list under divided attention nonetheless tended to give the studied items as answers.

As noted above, the results of Experiments 4a and 4b very likely underestimate the frequency and magnitude of unconscious memory effects on question answering in everyday life. The opposition instructions used in those experiments oriented subjects toward the list as a source of easy generation of incorrect answers and gave them a rule for rejecting those items as answers. In everyday life, people are generally oriented toward answering the question rather than identifying the source of their answer (see Jacoby, Kelley, Brown, & Jasechko, 1989, and Johnson, Hashtroudi, & Lindsay, in press), and often there is no simple rule that allows us to identify answers that pop to mind as wrong. Therefore, we expect that unconscious effects of the past have pervasive influences on confidence in question answering.

Summary and Conclusion

This paper has illustrated how effects of prior exposure to answers can be used to investigate the basis for confidence in answers. Our evidence suggests that easy generation of an idea during an attempt to answer a question serves as a basis for confidence in that idea being the correct answer. Ease of generation is generally a reliable basis for confidence, but a variety of factors discussed above can lead to easy generation of incorrect answers and hence to misplaced confidence. Such attribution errors may underlie a broad range of psychological phenomena, including schema-based intrusions in recall and false recognition (e.g., Ross, 1989), poor calibration of comprehension (e.g., Glenberg & Epstein, 1985; Maki & Berry, 1984), underestimations of difficulty for others (Kelley & Jacoby, 1991), and unintentional plagiarism (Brown & Murphy, 1989; Jacoby, 1988, p. 167).

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Arkes, H. R., Hackett, C., & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, 2, 81–94.

Bacon, F. T. (1979). Credibility of repeated statements: Memory for trivia. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 241–252.

Baddeley, A. D. (1982). Amnesia: A minimal model and an interpretation. In L. S. Cermak (Ed.), *Human memory and amnesia* (pp. 305–336). Hillsdale, NJ: Erlbaum.

Bartlett, F. C. (1932). *Remembering*. Cambridge: Cambridge University Press.

Begg, I. (1991, Feb.). *Logic and intuition*. Paper pre-

sented at the annual meeting of the Lake Ontario Visionary Establishment, Niagara Falls, Ont.

BEGG, I., ANAS, A., & FARINACCI, S. (1991). *Repetition and the ring of truth: Implicit memory or unconscious influence?* Manuscript submitted for publication.

BEGG, I., & ARMOUR, T. (1991). Repetition and the ring of truth: Biasing comments. *Canadian Journal of Behavioral Science.*

BEGG, I., ARMOUR, V., & KERR, T. (1985). On believing what we remember. *Canadian Journal of Behavioral Science,* **17**, 199–214.

BEGG, I., DUFT, S., LALONDE, P., MELNICK, R., & SANVITO, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language,* **28**, 610–632.

BLAXTON, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **15**, 657–668.

BOWERS, J. S., & SCHACTER, D. L. (1990). Implicit memory and test awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **16**, 404–416.

BROWN, R., & MCNEILL, D. (1966). The "tip-of-the-tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior,* **4**, 325–337.

BROWN, A. S., & MURPHY, D. R. (1989). Cryptomnesia: Delineating inadvertent plagiarism. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **15**, 432–442.

CRAIK, F. I. M. (1982). Selective changes in encoding as a function of reduced processing capacity. In F. Klix, J. Hoffman, & E. van der Meer (Eds.), *Cognitive research in psychology* (pp. 152–161). Berlin: Deutscher Verlag der Wissenschaffen.

ERICKSON, T. D., & MATTSON, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior,* **20**, 107–123.

FISCHHOFF, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance.* **1**, 288–299.

FOLTZ, G., & POLTROCK, S. (1984). A.P.T. II Experiment Development System.

GARDINER, J. M. (1988). Functional aspects of recollective experience. *Memory and Cognition,* **16**, 309–313.

GLENBERG, A. M., & EPSTEIN, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **11**, 702–718.

GLUCKSBERG, S., & MCCLOSKEY, M. (1981). Decisions about ignorance: Knowing that you don't know. *Journal of Experimental Psychology: Human Learning and Memory,* **7**, 311–325.

GRAESSER, A. C., & BLACK, J. B. (Eds.). (1985). *The psychology of questions.* Hillsdale: Lawrence Erlbaum Associates.

GRAESSER, A. C., & BOWER, G. H. (Eds.). (1990). Inferences and text comprehension. *The Psychology of Learning and Motivation,* Vol. 25. San Diego: Academic Press.

GRAESSER, A. C., & HEMPHILL, D. (1991). Question answering in the context of scientific mechanisms. *Journal of Memory and Language,* **30**, 186–209.

HART, J. T. (1965). Memory and the feeling of knowing experience. *Journal of Experimental Psychology,* **38**, 208–216.

HASHER, L., GOLDSTEIN, D., & TOPPINO, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior,* **16**, 107–112.

HAMANN, S. B. (1990). Levels-of-processing effects in conceptually driven implicit tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **16**, 970–977.

JACOBY, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language,* **30**, 513–541.

JACOBY, L. L. (1988). Memory observed and memory unobserved. In U. Neisser and E. Winograd (Eds.), *Remembering reconsidered* (pp. 145–177). New York: Cambridge University Press.

JACOBY, L. L., ALLAN, L. G., COLLINS, J. C., & LARWILL, L. K. (1988). Memory influences subjective experience: Noise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **14**, 240–247.

JACOBY, L. L., & DALLAS, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General,* **3**, 306–340.

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

JACOBY, L. L., & KELLEY, C. M. (1991). Unconscious influences of memory: Dissociations and automaticity. In D. Milner & M. Rugg (Eds.), *The neuropsychology of consciousness.* London, UK: Academic Press.

JACOBY, L. L., & HOLLINGSHEAD, A. (1990). Toward a generate/recognize model of performance on direct and indirect tests of memory. *Journal of Memory and Language,* **29**, 433–454.

JACOBY, L. L., KELLEY, C. M., BROWN, J., & JASECHKO, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology,* **56**, 326–338.

JACOBY, L. L., KELLEY, C. M., & DYWAN, J. (1989).

Memory attributions. In H. L. Roediger III and F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Erlbaum.

JACOBY, L. L., & WHITEHOUSE, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, **118**, 126–135.

JACOBY, L. L., WOLOSHYN, V., & KELLEY, C. M. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, **118**, 115–125.

JOHNSON, M. K., & HASHER, L. (1987). Human learning and memory. *Annual Review of Psychology*, **38**, 631–638.

JOHNSON, M. K., HASHTROUDI, S., & LINDSAY, D. S. (in press). Source monitoring. *Psychological Bulletin*.

JOHNSTON, W. A., DARK, V., & JACOBY, L. L. (1985). Perceptual fluency and recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **11**, 3–11.

KELLEY, C. M., & JACOBY, L. L. (1990). The construction of subjective experience: Memory attributions. *Mind and Language*, **5**, 49–68.

KELLEY, C. M., & JACOBY, L. L. (1991). *Illusions of ease*. Manuscript in preparation.

KELLEY, C. M., JACOBY, L. L., & HOLLINGSHEAD, A. (1989). Direct versus indirect tests of memory for source: Judgments of modality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 1101–1108.

KOEHLER, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin*, **110**, 499–519.

KOLERS, P. A., & ROEDIGER, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, **23**, 425–449.

KORIAT, A., LICHTENSTEIN, S., & FISCHHOFF, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, **6**, 107–118.

KUBOVY, M. (1977). Response availability and the apparent spontaneity of numerical choices. *Journal of Experimental Psychology: Human Perception and Performance*, **3**, 359–364.

LACHMAN, J. L., & LACHMAN, R. (1980). Age and the actualization of world knowledge. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, & L. W. Thompson (Eds.), *New directions in memory and aging: Proceedings of George Talland memorial conference*. Hillsdale, NJ: Erlbaum.

LEONESIO, R. J. & NELSON, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 464–470.

LINDSAY, D. S., & KELLEY, C. M. (1991, June). *Ease of generation during recall induces a feeling of remembering*. Poster presented at the annual meeting of the American Psychological Society, Washington, DC.

MAKI, R. H., & BERRY, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **10**, 663–679.

MORRIS, C. C. (1990). Retrieval processes underlying confidence in comprehension judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **16**, 223–232.

NELSON, T. O., & NARENS, L. (1980). Norms for 300 general-information questions: Accuracy of recall, latency of recall, and feeling of knowing ratings. *Journal of Verbal Behavior and Verbal Learning*, **19**, 338–368.

NELSON, T. O., & NARENS, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol. 26* (pp. 125–173). New York: Academic Press.

NISBETT, R. E., & WILSON, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, **84**, 231–259.

POLLARD, P. (1982). Human reasoning: Some possible effects of availability. *Cognition*, **12**, 65–96.

REDER, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, **19**, 90–138.

RICHARDSON-KLAVEHN, A., & BJORK, R. A. (1988). Measures of memory. *Annual Review of Psychology*, **39**, 475–543.

ROEDIGER, H. L., III, & BLAXTON, T. A. (1987). Retrieval modes produce dissociations in memory for surface information. In D. S. Gorfein and R. R. Hoffmann (Eds.), *Memory and cognition processes: The Ebbinghaus centennial conference* (pp. 349–379). Hillsdale: NJ: Erlbaum.

ROSS, B., RYAN, W. J., & TENPENNY, P. L. (1989). The access of relevant information for solving problems. *Memory and Cognition*, **17**, 639–651.

ROSS, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, **96**, 341–357.

SCHACTER, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 501–518.

SCHACTER, D. L. (1989). On the relation between memory and consciousness: Dissociable interactions and conscious experience. In H. L. Roediger III and F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 355–389). Hillsdale, NJ: Erlbaum.

SCHNEIDER, W. (1988). Micro-Experimental Laboratory: An integrated system for IBM PC compati-

bles. *Behavior Research Methods, Instruments, and Computers*, **20**, 206–217.

SRINIVAS, K., & ROEDIGER, H. L, III (1990). Classifying implicit memory tests: Category association and anagram solution. *Journal of Memory and Language*, **29**, 389–412.

TVERSKY, A., & KAHNEMAN, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, **5**, 207–232.

WHITTLESEA, B. W. A., JACOBY, L. L., & GIRARD, K. A. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language*, **29**, 716–732.

WITHERSPOON, D., & ALLAN, L. G. (1985). The effects of a prior presentation on temporal judgments in a perceptual identification task. *Memory & Cognition*, **13**, 101–111.

WOOD, G. (1978). The knew-it-all-along effect. *Journal of Experimental Psychology: Human Perception and Performance*, **4**, 345–353.

(Received July 25, 1991)

(Revision received March 11, 1992)