

The role of unconscious memory errors in judgments of confidence for sentence recognition

CRISTINA SAMPAIO

Western Washington University, Bellingham, Washington

AND

WILLIAM F. BREWER

University of Illinois at Urbana-Champaign, Urbana, Illinois

The present experiment tested the hypothesis that unconscious reconstructive memory processing can lead to the breakdown of the relationship between memory confidence and memory accuracy. Participants heard deceptive schema-inference sentences and nondeceptive sentences and were tested with either simple or forced-choice recognition. The nondeceptive items showed a positive relation between confidence and accuracy in both simple and forced-choice recognition. However, the deceptive items showed a strong negative confidence/accuracy relationship in simple recognition and a low positive relationship in forced choice. The mean levels of confidence for erroneous responses for deceptive items were inappropriately high in simple recognition but lower in forced choice. These results suggest that unconscious reconstructive memory processes involved in memory for the deceptive schema-inference items led to inaccurate confidence judgments and that, when participants were made aware of the deceptive nature of the schema-inference items through the use of a forced-choice procedure, they adjusted their confidence accordingly.

There has been much disagreement in the literature on the issue of the degree to which memory confidence predicts memory accuracy (Bothwell, Deffenbacher, & Brigham, 1987; Perfect & Hollins, 1996; Read, Lindsay, & Nicholls, 1998; Sporer, Penrod, Read, & Cutler, 1995). One proposal that can account for some of the dramatic differences found among experiments on memory confidence is the hypothesis that unconscious memory errors are responsible for some of the inaccuracies of confidence judgments (Koriat, 1995; Leippe, 1980; Zechmeister & Nyberg, 1982, pp. 248–249). For example, Leippe argued that

as reconstructive processes become more extensive, the accuracy–confidence relationship should become correspondingly smaller. The reason for this is that, while people may have a veridical feeling about how accessible or “strong” their memorial representations of objects are, they are not likely to be *conscious* of the transformations that these representations may have gone through during encoding, storage, and retrieval. (p. 264)

Although the hypothesis of the role of unconscious memory processes influencing memory confidence has considerable theoretical appeal, it has not been explicitly tested.

We have proposed a metamemory approach to the study of confidence (Brewer & Sampaio, 2006, 2008; Brewer, Sampaio, & Barlow, 2005) in which we view confidence

as an inferential process based on cues that are available at the time of the judgment and on metamemory beliefs about how these cues relate to accuracy. More specifically, we think that the main sources of information on which memory confidence is based are (1) the process used in a memory task (e.g., recall vs. familiarity judgments) along with the products of the memory task (e.g., an image) and (2) the metamemory beliefs about how different memory processes and products relate to memory accuracy (e.g., the belief that memory information in recollective recall is likely to be accurate). In many memory situations, these metamemory beliefs are valid, and their use leads to a positive relationship of confidence and accuracy. However, there are situations in which the relationship breaks down (Busey, Tunnicliff, Loftus, & Loftus, 2000; Kelley & Lindsay, 1993; Koriat, 1995; Koriat & Goldsmith, 1996; Tomes & Katz, 2000). For example, in a previous recognition experiment (Brewer & Sampaio, 2006) we showed that individuals assign appropriate levels of confidence for standard nondeceptive experimental materials but assign inappropriately high levels of confidence for deceptive materials. In that experiment, when participants were presented with a sentence such as *The hungry python caught the mouse* and tested with a deceptive foil such as *The hungry python ate the mouse* in a simple recognition test, they frequently judged the foil to be *old* and showed high confidence in their false judgment. We assumed that

C. Sampaio, cristina.sampaio@wwu.edu

in this experimental condition, participants were not aware of the operation of the gist-based memory processes and that their lack of awareness about their incorrect responses led to the breakdown of the usual positive relationship between memory confidence and accuracy. It seems to us that the hypothesis of unconscious memory errors provides a very plausible account of these data. However, in that experiment we provided no *direct* empirical evidence to support the crucial hypothesis.

Our goal in the present experiment was to provide the first empirical test of the hypothesis that unconscious memory errors are responsible for some of the inaccuracy of confidence judgments that have been reported in the literature. In this experiment, we included a test procedure that should lead participants to be more aware of the deceptive nature of schema-inference items. Specifically, we contrasted the data from a group of participants tested with a simple recognition procedure against the data from a group of participants tested with a forced-choice procedure. Our reasoning was that a participant presented with a sentence such as *The hungry python caught the mouse* and later asked to choose between *The hungry python caught the mouse* and the deceptive foil *The hungry python ate the mouse* would be more likely to become aware of the deceptive nature of this type of item than would a participant who was asked to carry out a simple recognition memory task. If this was found to be the case, then on the basis of the metamemory theory of confidence we predicted that individuals tested with a forced-choice procedure would continue to make memory errors but would drop their confidence levels for these items. This would lead to a stronger confidence/accuracy relationship for the schema-inference items in the forced-choice condition than in the simple recognition condition.

METHOD

Participants

The participants were 36 undergraduate students from the University of Illinois at Urbana-Champaign who carried out the experiment as part of a course requirement.

Materials

Acquisition Items

Schema-inference deceptive sentences. The schema-inference sentences were designed to lead the listener to make schema-based inferences (e.g., *The narcotics officer pushed the doorbell* leads to the inference that the narcotics officer rang the doorbell). The input sentence designed to lead to a schema-based inference we call the *original sentence*, and the sentence that expresses the expected inference we call the *inferred sentence*. The original sentences have been shown to produce large numbers of memory errors in recall (Brewer, 1977), and these errors were apparently not noticed by the participants. Thus, we hypothesized that an inferred sentence would produce a high rate of false alarms in recognition memory without participants being aware of the errors. The 24 schema-inference pairs used in this experiment were taken from Brewer (1977) and Birch and Brewer (1995), and most are given in the Appendix of McDermott and Chan (2006).

Nondeceptive base sentences. The base sentences were 72 *nondeceptive* sentences constructed to cover a wide range of difficulty in memory. The range of difficulty was developed by writing sen-

tences of varying length, grammatical structure, imagery, and level of abstraction. For example, *The zebra outran the lion* is an example of an easy sentence; *The Federal Reserve sets the discount interest rate* is an example of a moderately difficult sentence; and *Deductive logic depends on entailment among propositions* is an example of a difficult sentence.

Presentation Lists (Items Presented During Acquisition)

Two presentation lists were counterbalanced for schema-inference item type (i.e., the schema-inference items that were in original form on one list were in inferred form on the other, and vice versa). Each list contained a total of 102 sentences (12 original schema-inference sentences, 12 inferred schema-inference sentences, all 72 base sentences, and 6 filler sentences). The order of the sentences in the list was randomized.

New Test Items (Foil)

Deceptive schema-inference foil sentences. The *deceptive* items were the foil schema-inference sentences that expressed the expected schema-based inference (an inferred item) when the acquisition sentence was an original item (e.g., the presentation of *The hungry python ate the mouse* as a test item when *The hungry python caught the mouse* was the acquisition item). These items were considered deceptive because we hypothesized that participants presented with the original sentence of a schema-inference item would be likely to make a schema-based inference and thus would falsely recognize the sentence that expressed that inference.

Nondeceptive schema-inference foil sentences. The *nondeceptive* items for the schema-inference items were the schema-inference sentences that contained information that was designed to lead to a schema-based inference (an original item) when the acquisition sentence was an inferred item (e.g., the presentation of *The hungry python caught the mouse* as a test item when *The hungry python ate the mouse* was the acquisition item). Across the schema-inference foil sentences, half were deceptive and half were not.

Nondeceptive base foil sentences. We constructed 216 sentences to be foil items for the base sentences in the recognition tests. The base items had been written to vary in intrinsic difficulty, and their foils were developed to vary widely in similarity with respect to the base items in order to provide an even broader range of difficulty in the recognition memory tasks. A set of three sentences was written for each base presentation sentence. The first sentence of the set was written to be highly similar to the list sentence, the second was written to be moderately similar to the list sentence, and the third was written to be dissimilar to the list sentence (e.g., base list sentence, *The tornado picked up the elm tree*; highly similar sentence, *The tornado picked up the oak tree*; moderately similar sentence, *The tornado picked up the cow*; dissimilar sentence, *The mermaid swam to the sailor*). There were 72 highly similar foils, 72 moderately similar foils, and 72 dissimilar foils. To provide an overall wide range of item difficulty, we used base items that varied widely in intrinsic difficulty, along with foils that differed widely in similarity to them (see Lindsay, Read, & Sharma, 1998, on the importance of having a wide range of item difficulty in studies of memory confidence). Because our goal was simply to produce a wide range of item difficulty to study memory confidence rather than to show the well-known effects of intrinsic item difficulty and foil similarity in recognition memory, the data are not broken down to display such effects in the analysis of the results.

Old Test Items

The old items of the recognition test forms (base, schema-inference, and filler) were items that had been presented on the presentation lists.

Simple Recognition Procedure

Simple recognition item counterbalancing. The test booklet was organized so that only 1 member of a schema-inference pair

(original/inferred) was tested for a given participant, and only 1 of the 3 items from a similarity foil set or its corresponding base item was tested for a given participant. Each test booklet contained a total of 108 test items (24 schema-inference, 72 base/foil, and 12 filler). There were four forms of test booklets counterbalanced for the type of schema-inference item (original, inferred) and for the type of base sentence (old base, highly similar foil, moderately similar foil, dissimilar foil). Twelve filler items were included on all forms. Across the booklets, half of the schema-inference items were tested as *old* items and half were tested as *new* items. The order of the items was randomized, and each counterbalanced form followed the same random order.

Test booklet format for simple recognition. The sentences were arranged in the test booklet so that there were four sentences per page. Below each sentence was a *yes–no* recognition test. Participants were instructed to read the sentence and to circle *yes* if they thought the sentence had appeared in the list presented to them earlier in the experiment and *no* if they thought it had not. The participants were given explicit instructions that their responses should be based on literal (surface structure) memory. The instructions they were given stated that “if the experimenter read the sentence *The sailor picked the rock up* and you are given the sentence *The sailor picked up the rock* you should circle N [for No] because this sentence is not exactly the same as the one the experimenter read.”

Below the recognition test was a 7-point confidence scale, for which 1 = *totally uncertain (a guess)* and 7 = *absolutely certain*. Participants were instructed to circle 1 to indicate that they had absolutely no confidence and to circle 7 to indicate that they were totally confident that the sentence was exactly the same as the one that the experimenter read. When participants gave “no” responses (i.e., items they considered *not* to be old), they responded with their degree of confidence that the sentence was not the one that they had seen earlier. Participants were instructed to use intermediate numbers to indicate intermediate degrees of confidence. Below the confidence scale for each item was a multiple-choice introspective report questionnaire (taken from Brewer & Sampaio, 2006) that was designed to examine the bases for confidence judgments. The data from the introspective reports of the present study replicate the findings from the earlier study, but because they were not the focus of the present experiment they are not discussed in this article.

Forced-Choice Recognition Procedure

Forced-choice recognition item counterbalancing. The test booklet was organized so that each schema-inference item and each base item were tested only once for each participant. The test booklets contained 24 schema-inference pairs (original/inferred), 72 base pairs (base/foil), and 6 filler pairs, for a total of 102 pairs of items. There were three different forced-choice recognition test booklets. Each booklet included all 72 base items paired with 1 item from the similarity foil set; the booklets were counterbalanced so that each contained 24 base items paired with highly similar foils, 24 base items paired with moderately similar foils, and 24 base items paired with dissimilar foils. This set of counterbalancing procedures ensured that each base item was tested only once in each booklet (with one of the three different types of foils) and that each base item was tested once with each type of foil across the three booklets. The 24 schema-inference pairs and the 6 filler pairs were included in all three test booklets. The order of the item pairs was randomized, and the items in each counterbalanced booklet followed the same random order.

Test booklet format for forced choice. The sentences were arranged in the test booklet so that there were two pairs of sentences per page. The sentences of a given pair were presented side by side, and the pairs were counterbalanced for (right/left) order on the page. Below each sentence of each pair was a *yes–no* recognition test. Participants were instructed to read the pair of sentences, circle *yes* for the sentence that they thought had appeared in the list presented to them earlier in the experiment, and circle *no* for the sentence that they thought had not appeared in the list. Below the recognition test

for each of the two sentences in a recognition pair was the same 7-point confidence scale used in the simple recognition test condition, followed by the introspective report questionnaire.

Procedure

The data for the simple recognition condition and the data for the forced-choice condition were gathered at different times from the same participant pool. The participants used in the two different conditions were sampled randomly from the pool (i.e., a sign-up procedure was not used). Participants were tested in small groups. The instructions stated that participants would hear a list of sentences and later would be tested on information from the sentences, but they did not explicitly suggest either a rote-memory or a gist-memory strategy. The experimenter read the presentation list of sentences to the participants at the rate of approximately one sentence every 4 sec. Immediately after the last sentence was read, the participants received a test booklet containing the instructions and the test sentences (each sentence followed by the recognition task, the confidence scale, and the introspective report questionnaire). The instructions were read out loud by the experimenter, and the participants were allowed as much time as they needed to complete the recognition booklet. The time from the last presentation item to the first recognition item was about 12 min. The average time to complete the booklet was 22 min for the simple recognition test condition and 32 min for the forced-choice test condition. Participants were debriefed at the conclusion of the experiment.

RESULTS

For the purposes of the data analysis, the foil items containing a schema inference (i.e., the new inferred items) were considered to be deceptive. The old base items, the old schema-inference items, the filler items, the new base foils (moderate and low similarity), and the new original schema-inference items were considered to be nondeceptive. Examination of the data from the high-similarity base foils suggested that some of these items were functioning as deceptive items, so all high-similarity base foils were excluded from the analyses that involved comparing deceptive and nondeceptive items.

Memory Performance for Deceptive and Nondeceptive Items

The schema-inference items that were originally developed to show schema inferences in a recall task (Brewer, 1977) served as very effective deceptive items in the recognition memory tasks of the present experiment. In the simple recognition condition, the mean accuracy for the deceptive items ($M = .36$) was actually below chance ($M = .50$ for this *yes–no* recognition task, $p = .000$) and considerably below that for the nondeceptive items ($M = .70$) [$t(17) = 2.75$, $p = .01$]. In the forced-choice condition, the mean accuracy for the deceptive items was essentially at chance ($M = .51$, $p = .80$), contrasting strongly with the good performance for the nondeceptive items ($M = .82$) [$t(17) = 9.54$, $p = .000$].

Correlations of Confidence and Accuracy for Deceptive and Nondeceptive Items

The metamemory theory of confidence predicts that the nondeceptive items ought to show a positive relationship between confidence and accuracy in simple rec-

ognition and that there ought to be a breakdown of the confidence/accuracy relationship for deceptive items. We make an even stronger prediction about the nature of the confidence/accuracy relationship for deceptive items (Brewer & Sampaio, 2006): If the processes that lead to an error for deceptive items are similar to those involved in carrying out a successful response, we predict a negative correlation, because items that are most deceptive (e.g., most likely to be falsely recognized) are most likely to evoke the normally successful mental operations and thus lead to inappropriately high confidence.

The unconscious memory hypothesis predicts that for a forced-choice task the confidence/accuracy relationship for deceptive items ought to be less disrupted than the relationship for simple recognition; the changes should also be larger for deceptive than for nondeceptive items since we hypothesize that the forced-choice task will lead participants to become more aware of the deceptive nature of the deceptive items. Although the unconscious memory error hypothesis predicts an improvement in the confidence/accuracy relationship for the forced-choice condition, it does not predict the size of this correlation, because it depends on the degree to which the forced-choice condition makes each participant aware of the deceptive nature of each deceptive item.

A 2 (type of item: deceptive vs. nondeceptive) \times 2 (type of test: simple recognition vs. forced choice) mixed factorial ANOVA on the within-subjects gammas was conducted to test our hypotheses. The results revealed that there was a significant interaction of type of item \times type of test such that the forced-choice procedure led to a larger improvement in the confidence/accuracy relationship for the deceptive items than for the nondeceptive items [$F(1,31) = 14.76, p = .001$]. For the deceptive items, the gamma coefficient was $-.61$ in simple recognition and $+.18$ in forced choice. For the nondeceptive items, the gamma coefficient was $+.30$ in simple recognition and $+.61$ in forced choice.

In addition, we carried out a conceptual replication of the present experiment using gist-based synonym-substitution items as the deceptive materials (e.g., given *The St. Bernard was too big for the car* as an acquisition item, participants were tested with *The St. Bernard was too large for the car*). The data for simple recognition are reported in Brewer and Sampaio (2006), and the data for forced-choice recognition are reported in an unpublished manuscript (Sampaio & Brewer, 2007). In the experiment using gist-based deceptive items, the gamma was $-.44$ in simple recognition and 0 in forced choice. For the nondeceptive items, the gamma was $+.26$ for simple recognition and $+.54$ for forced choice. This pattern of data replicates that of the present experiment, supporting the hypothesis that unconscious memory processes reduce the confidence/accuracy relationship for deceptive items. However, the interaction of item type and type of test only showed a p value of $.17$ in the conceptual replication.

The unconscious memory error hypothesis can also be tested with an overall ANOVA on confidence scores for

memory errors. The unconscious memory error hypothesis predicts that, with a simple recognition technique, participants should make many errors on deceptive items and that these errors should be held with inappropriately high confidence. With a forced-choice technique, however, the participants should become more aware of the deceptive nature of the schema-based items, show reduced confidence for these erroneous responses, and thus exhibit an improved confidence/accuracy relationship. Hence, the unconscious memory error hypothesis predicts an interaction of type of item (deceptive vs. nondeceptive) with type of test (simple recognition vs. forced choice) on confidence for errors. A 2 (type of item: nondeceptive vs. deceptive) \times 2 (type of test: simple recognition vs. forced choice) mixed factorial ANOVA was carried out on the confidence data for memory errors and showed the predicted interaction [$F(1,34) = 4.45, MS_e = .349, p = .04$], such that the mean confidence for errors for deceptive items showed a greater drop when participants were tested with the forced-choice procedure (from 5.7 to 4.6) than did the mean confidence for errors for nondeceptive items (from 4.4 to 3.8), as the data in Table 1 show. These results support the view that the participants became aware of the deceptive nature of the schema-based items in the forced-choice condition and thus reduced their confidence for these errors accordingly.

We also can examine this same issue with data from our conceptual replication of the present experiment that used gist-based deceptive items, and indeed a 2 (type of item: nondeceptive vs. deceptive) \times 2 (type of test: simple recognition vs. forced choice) ANOVA also showed a significant interaction [$F(1,34) = 10.69, MS_e = .169, p = .002$]. The mean confidence for errors for deceptive items showed a greater drop (from 5.8 to 3.9) than did that for nondeceptive items (from 4.8 to 3.7). Thus, the data on memory confidence for errors gathered at a different time, with different participants, and with different materials replicated the findings of the present experiment.

The pattern of data in both experiments supports the unconscious memory error hypothesis. In particular, the data suggest that the forced-choice procedure caused the participants to realize that their normally successful metamemory beliefs did not apply for the deceptive items, and hence they adjusted their confidence downward. For

Table 1
Memory Confidence Judgments for Correct Responses and Errors in Simple Recognition and Forced-Choice Recognition for Deceptive and Nondeceptive Items

Item Type	Simple Recognition		Forced Choice	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Correct				
Nondeceptive	5.2	1.8	5.7	1.8
Deceptive	4.0	1.7	4.9	2.0
Error				
Nondeceptive	4.4	1.8	3.8	2.0
Deceptive	5.7	1.5	4.6	1.7

Note—Confidence judgments were made on a 1–7 scale.

example, a participant might have had a feeling of familiarity for both the original and the inferred schema-inference members of a schema-inference test pair. However, the instructions stated that only one was actually *old*, and thus the participant selected one of the two choices with reduced confidence.

DISCUSSION

Our results show that for nondeceptive items memory confidence is a moderately valid index of memory accuracy. This is consistent with the findings of Brewer and Sampaio (2006), Brewer et al. (2005), Mandler and Boeck (1974), Murdock (1965), Perfect and Hollins (1996), and Tulving and Thomson (1971). For deceptive items, we found that there was a disruption of the positive relationship between confidence and accuracy in both simple recognition and in forced-choice memory and that the size of the disruption was reduced for forced-choice tasks, which lead to more awareness of the deceptive nature of the deceptive memory items. The results of the present experiment also are consistent with research using the misinformation paradigm, which shows that providing postevent misleading information to participants leads to a breakdown in the confidence/accuracy relationship (Kelley & Lindsay, 1993; Tomes & Katz, 2000).

However, we believe that there is no absolute answer to the question of the general accuracy of memory confidence judgments. Gigerenzer, Hoffrage, and Kleinbölting (1991) have developed a theory of ecological rationality in which they argue that one's beliefs are valid reflections of past experiences as long as the current task is representative of the past experience. We think that this argument applies to memory confidence and that, in general, memory confidence is positively related to memory accuracy, but that the relationship is disrupted for materials that lead to unconscious memory errors (cf. Juslin, 1993, for a similar argument for semantic memory tasks). Generally the various products and processes associated with memory tasks are diagnostic of accuracy, and thus the metamemory beliefs about how they relate to accuracy are normally valid indicators of accuracy. However, individuals appear to assume that their metamemory beliefs are valid across all memory tasks, resulting in a miscalibration in situations for which their metamemory beliefs are unjustified (e.g., for deceptive items). Note that, from the participants' point of view, they have not made inappropriate confidence judgments. In our experiment, for example, the participants in the simple recognition condition appeared to believe that many of the deceptive inferred sentences *did* occur on the presentation list, so from their point of view they should call these items *old* with high confidence.

We believe that the accuracy of memory confidence judgments depends on the distribution of materials that has been experienced previously and on the makeup of the items being tested. Thus, with a list of nondeceptive items, one can have a strong positive relationship between confidence and accuracy. With a list including a mixture

of deceptive and nondeceptive items, one can have no relationship between confidence and accuracy. With a list of only deceptive items, one can have a strong negative relation between confidence and accuracy.

Our main goal in the present article was to test the hypothesis that unconscious memory errors, such as schema-based errors, are one basis for the inaccuracy of confidence judgments. We examined this hypothesis by contrasting data from a simple recognition procedure with data from a forced-choice procedure. Although researchers have speculated about the unconscious reconstructive processing hypothesis (e.g., Brewer & Sampaio, 2006; Brewer et al., 2005; Koriat, 1995; Leippe, 1980; Zechmeister & Nyberg, 1982), this article provides a direct empirical test of the hypothesis. Our data indicate that in a simple recognition task, participants applied their metamemory beliefs inappropriately, producing memory errors with high confidence. We hypothesized that with a forced-choice task many participants would become aware of the deceptive nature of the schema-inference items and thus would provide lower confidence for those items. The data from the present experiment and from a conceptual replication with gist-based materials support this prediction and are consistent with those obtained by Koriat and Goldsmith (1996) in the domain of semantic memory. The contrast between the data from the simple recognition task and the data from the forced-choice recognition task supports the view that unconscious reconstructive memory processes lead to a breakdown of the confidence/accuracy relationship.

Overall, the results of this experiment add support for the metamemory theory of confidence and provide the first direct evidence for the hypothesis that unconscious errors can cause a breakdown in the typically positive relationship between memory confidence and memory accuracy.

AUTHOR NOTE

This research was supported in part by a grant from the University of Illinois Research Board. Portions of this research were presented at the 2002 Annual Meeting of the Psychonomic Society, Kansas City, MO. We thank Ellen Brewer for comments on an earlier draft of this article. Correspondence concerning this article should be addressed to C. Sampaio, Department of Psychology, Western Washington University, 516 High Street, Bellingham, WA 98225 (e-mail: cristina.sampaio@wwu.edu).

REFERENCES

- BIRCH, S., & BREWER, W. F. (1995). The fate of originally presented surface information following recall errors in sentence memory tasks. *European Journal of Cognitive Psychology*, *7*, 145-167.
- BOTHWELL, R. K., DEFFENBACHER, K. A., & BRIGHAM, J. C. (1987). Correlation of eyewitness accuracy and confidence: Optimality hypothesis revisited. *Journal of Applied Psychology*, *72*, 691-695.
- BREWER, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition*, *5*, 673-678.
- BREWER, W. F., & SAMPAIO, C. (2006). Processes leading to confidence and accuracy in sentence recognition: A metamemory approach. *Memory*, *14*, 540-552.
- BREWER, W. F., & SAMPAIO, C. (2008). *The metamemory approach to confidence: A test using semantic memory*. Unpublished manuscript.
- BREWER, W. F., SAMPAIO, C., & BARLOW, M. R. (2005). Confidence and accuracy in the recall of deceptive and nondeceptive sentences. *Journal of Memory & Language*, *52*, 618-627.
- BUSEY, T. A., TUNNICLIFF, J., LOFTUS, G. R., & LOFTUS, E. F. (2000).

- Accounts of the confidence–accuracy relation in recognition memory. *Psychonomic Bulletin & Review*, *7*, 26-48.
- GIGERENZER, G., HOFFRAGE, U., & KLEINBÖLTING, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506-528.
- JUSLIN, P. (1993). An explanation of the hard–easy effect in studies of realism of confidence in one's general knowledge. *European Journal of Cognitive Psychology*, *5*, 55-71.
- KELLEY, C. M., & LINDSAY, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory & Language*, *32*, 1-24.
- KORIAT, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, *124*, 311-333.
- KORIAT, A., & GOLDSMITH, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, *103*, 490-517.
- LEIPPE, M. R. (1980). Effects of integrative memorial and cognitive processes on the correspondence of eyewitness accuracy and confidence. *Law & Human Behavior*, *4*, 261-274.
- LINDSAY, D. S., READ, J. D., & SHARMA, K. (1998). Accuracy and confidence in person identification: The relationship is strong when witnessing conditions vary widely. *Psychological Science*, *9*, 215-218.
- MANDLER, G., & BOECK, W. J. (1974). Retrieval processes in recognition. *Memory & Cognition*, *2*, 613-615.
- MCDERMOTT, K. B., & CHAN, J. C. K. (2006). Effects of repetition on memory for pragmatic inferences. *Memory & Cognition*, *34*, 1273-1284.
- MURDOCK, B. B., JR. (1965). Signal-detection theory and short-term memory. *Journal of Experimental Psychology*, *70*, 443-447.
- PERFECT, T. J., & HOLLINS, T. S. (1996). Predictive feeling of knowing judgements and postdictive confidence judgements in eyewitness memory and general knowledge. *Applied Cognitive Psychology*, *10*, 371-382.
- READ, J. D., LINDSAY, D. S., & NICHOLLS, T. (1998). The relation between confidence and accuracy in eyewitness identification studies: Is the conclusion changing? In C. P. Thompson, D. J. Herrmann, J. D. Read, D. Bruce, D. G. Payne, & M. P. Toglia (Eds.), *Eyewitness memory: Theoretical and applied perspectives* (pp. 107-130). Mahwah, NJ: Erlbaum.
- SAMPAIO, C., & BREWER, W. F. (2007). *A metamemory approach to memory confidence in sentence recognition*. Unpublished manuscript.
- SPORER, S. L., PENROD, S., READ, D., & CUTLER, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence–accuracy relation in eyewitness identification studies. *Psychological Bulletin*, *118*, 315-327.
- TOMES, J. L., & KATZ, A. N. (2000). Confidence–accuracy relations for real and suggested events. *Memory*, *8*, 273-283.
- TULVING, E., & THOMSON, D. M. (1971). Retrieval processes in recognition memory: Effects of associative context. *Journal of Experimental Psychology*, *87*, 116-124.
- ZECHMEISTER, E. B., & NYBERG, S. E. (1982). *Human memory: An introduction to research and theory*. Monterey, CA: Brooks/Cole.

(Manuscript received March 11, 2008;
revision accepted for publication October 22, 2008.)