

# The revelation effect for autobiographical memory: A mixture-model analysis

**DANIEL M. BERNSTEIN**

*Kwantlen Polytechnic University, Surrey, British Columbia, Canada,  
and University of Washington, Seattle, Washington*

**MICHAEL E. RUDD**

*University of Washington, Seattle, Washington*

**EDGAR ERDFELDER**

*University of Mannheim, Mannheim, Germany*

**RYAN GODFREY**

*University of California, Riverside, California*

AND

**ELIZABETH F. LOFTUS**

*University of California, Irvine, California*

Participants provided information about their childhood by rating their confidence about whether they had experienced various events (e.g., “broke a window playing ball”). On some trials, participants unscrambled a key word from the event phrase (e.g., *widnwo*–*window*) or an unrelated word (e.g., *gnutge*–*nugget*) before seeing the event and giving their confidence ratings. The act of unscrambling led participants to increase their confidence that the event occurred in their childhood, but only when the confidence rating immediately followed the act of unscrambling. This increase in confidence mirrors the “revelation effect” observed in word recognition experiments. In the present article, we analyzed our data using a new signal detection mixture distribution model that does not require the researcher to know the veracity of memory judgments a priori. Our analysis reveals that unscrambling a key word or an unrelated word affects response bias and discriminability in autobiographical memory tests in ways that are very similar to those that have been previously found for word recognition tasks.

How do people decide whether a particular event occurred in the past? One important factor is the manner in which they process the event they are trying to remember. Unscrambling a word (anagram) just prior to making a recognition decision on that word or on an unrelated word increases the belief that the target word was seen before—a phenomenon called *the revelation effect* (Watkins & Peynircioğlu, 1990; Westerman & Greene, 1998). The revelation effect, which is mostly observed with verbal stimuli, has been extended to childhood autobiographical memory (Bernstein, Whittlesea, & Loftus, 2002). Participants express more confidence that events happened in their childhood if they unscramble a word embedded within descriptions of those events (e.g., “broke a nwidwo playing ball”) prior to making the confidence judgment.

Verde and Rotello (2004) have shown that revelation experiments in which the anagram is the same word as the target word (NWIDWO–WINDOW) yield different effects than do experiments in which the anagram is unrelated to the

target word (EBLNDRE [BLENDER]–WINDOW). Using signal detection theory, they demonstrated that the revelation effect for unrelated anagrams is due to increased response bias only (i.e., a general tendency to judge items as “old”), whereas the revelation effect for target-word anagrams is due to both increased response bias and impaired ability to discriminate old and new words as measured by the discriminability parameter  $d'$ .

The present work had two goals. First, we wished to test whether there would be a revelation effect in autobiographical memory when an anagram was presented immediately prior to (rather than simultaneously with) the rated life-event item. Doing this would conceptually replicate results found for word recognition tasks. Second, we wished to show that solving anagrams that were related versus those unrelated to life event-items would produce effects on discriminability and response bias similar to those previously found in standard recognition paradigms for anagrams that were identical to rather than unrelated

---

Daniel.Bernstein@kwantlen.ca

---

to target items (Verde & Rotello, 2004). However, unlike in Verde and Rotello, we could not use standard signal detection (SD) methods to achieve our second goal, because it is generally unknown which life events depicted in the test really happened to a participant (“true events”) and which did not (“false events”). We therefore developed a new SD mixture distribution model that helped us answer our research questions.

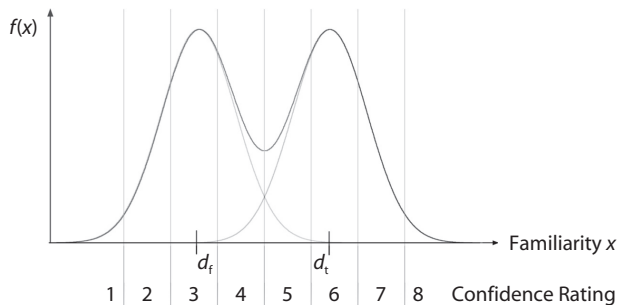
Assume that an unknown proportion,  $p$ , of items in the autobiographical memory test corresponds to true events from the participants’ past. By implication, a proportion  $(1 - p)$  of the test items must then describe false events. In keeping with the tenets of SD theory (see, e.g., Macmillan & Creelman, 1991), assume also that the familiarities of true and false events are independently normally distributed with the means  $d_t$  and  $d_f$  ( $d_t > d_f$ ), and the standard deviations  $\sigma_t$  and  $\sigma_f$ , respectively. Given these assumptions, the combined familiarity distribution aggregated across all true and false life events must be a two-components normal mixture. Roughly speaking, the left part of this mixture distribution corresponds to false events and the right part to true events, with an overlapping part in the middle (Figure 1).

SD theory is often applied to confidence ratings in “old” judgments (Macmillan & Creelman, 1991, chap. 3). Given a rating scale with  $X$  categories, a standard axiom is that the rating  $R$  exceeds rating category  $x$  ( $x = 1, \dots, X - 1$ ), if and only if the test item familiarity is larger than the response criterion,  $c_x$ . Applied to our mixture model, the probability of a confidence rating not exceeding  $x$  is

$$Pr(R \leq x) = p \cdot \Phi[(c_x - d_t) / \sigma_t] + (1 - p) \cdot \Phi[(c_x - d_f) / \sigma_f], \tag{1}$$

where  $\Phi(z)$  denotes the standard normal cumulative distribution function. It is useful for one to reconsider the model as a function of the discriminability parameter,  $d'$ , where  $d' = (d_t - d_f) / \sigma_f$ . Note that because  $d_t = d' \cdot \sigma_f + d_f$ ,

$$Pr(R \leq x) = p \cdot \Phi\{[c_x - (d' \cdot \sigma_f + d_f)] / \sigma_t\} + (1 - p) \cdot \Phi[(c_x - d_f) / \sigma_f]. \tag{2}$$



**Figure 1.** The signal detection mixture distribution model for false and true events combined (assuming  $X = 8$  rating categories, a true-events proportion of  $p = .50$ , equally spaced response criteria, and equal familiarity standard deviations  $\sigma_f = \sigma_t = 1$ ). Parameters  $d_f$  and  $d_t$  denote the mean familiarities of false and true events, respectively. The black line illustrates the familiarity mixture distribution. Gray lines illustrate the two component distributions for false and true events, respectively.

Assuming that response biases affect all response criteria equally (additive shift hypothesis), there are two equivalent options for measuring response bias effects: (1) assessing variations in response criteria for fixed  $d_f$  and (2) assessing variations in  $d_f$  for fixed response criteria (cf. Morey, Pratte, & Rouder, 2008, Figure 3). Although the first option is the most often used in SD applications, we chose the second option because it is more parsimonious in terms of the number of parameters required to capture response bias effects.

For fixed response criteria, an increase in response bias to rate all events “old” (i.e., a global familiarity illusion) would show up as an increase in  $d_f$ . In contrast, a reduced ability to discriminate between true and false life events would manifest as a decrease in  $d'$ . Past work on the revelation effect suggests that smaller  $d'$  parameters are likely to be obtained only when anagrams related to the target information are solved just before replying to the recognition probe. In contrast,  $d_f$  increases are expected both for anagrams that are related and anagrams that are unrelated to the recognition probe.

### EXPERIMENT 1

In Experiment 1, we presented anagrams of key words (NWIDWO; related anagram) or unrelated words (GNUTGE; unrelated anagram) prior to the putative life event (“broke a window playing ball”). Participants then rated the life event in terms of whether it happened in their own childhood. A control condition was included in which no anagram was presented prior to the life-event ratings (see Verde & Rotello, 2004, Experiment 4, for a similar procedure involving word recognition).

In many prior revelation studies, false-alarm rates increased in both the related and the unrelated anagram conditions relative to the no-anagram condition, but hit rates remained relatively stable (see Verde & Rotello, 2004). On the basis of such results, we expected that participants’ confidence that events occurred in their childhood would increase in both related and unrelated anagram conditions relative to the control condition. Moreover, we expected differential effects of related and unrelated anagrams on both discriminability and bias.

### Method

**Participants.** Seventy-two University of Washington undergraduates participated for course credit.

**Stimuli and Procedure.** Participants trained on 15 single-word anagrams and were given rules for unscrambling them. For example, participants attempted to unscramble the anagram ADRAM using the rules 2,3,1,5,4. The “2” and “3” refer to the second and third letter in the anagram (e.g., the “D” and “R” in ADRAM), respectively, which are also the first and second letters in the unscrambled version of the target word: DRAMA.

Stimuli were presented in capital letters on a flat-white wall using a 60-hz LCD projector interfaced to a Macintosh G4 running under MATLAB. Participants were run in groups of up to 8.

The three conditions—control, related anagram, and unrelated anagram—were run in separate blocks of 16 trials each, and were counterbalanced across participants. The item order was fixed. The test phase consisted of a 48-item life events inventory (LEI) taken from the study of Bernstein, Godfrey, Davison, and Loftus (2004).

On each test trial, participants either viewed a series of Xs (control condition) or unscrambled anagrams of key words (e.g., NWIDWO) or unrelated anagrams (KBAIGN–BAKING) before seeing and rating intact life events (BROKE A WINDOW PLAYING BALL). The number of Xs in the control condition and the number of letters in the unrelated anagram condition varied from 4 to 10, and matched the number of letters in the key word within the life event (see Appendix 1 in Bernstein et al., 2004). Participants used a 1–8 scale to rate their confidence that the events occurred in their childhood before the age of 10 (1 = *definitely did not occur*; 8 = *definitely did occur*).

Depending on the condition, participants either viewed Xs or unscrambled words for 10 sec prior to seeing and rating the life event. Pilot work indicated that participants could unscramble most words within 10 seconds. At the end of the 10-sec period, a beep sounded, followed 1 sec later by the life event. The participants were given 7 sec to rate the life event before the next trial began. Participants were instructed to stop working on unscrambling the anagram as soon as the beep sounded. Successive blocks of trials were separated by a 1-min break. The participants solved anagrams and rated life events on a sheet of paper provided by the experimenter.

**Results and Discussion**

Participants successfully unscrambled 79% of the anagrams.<sup>1</sup> Unscrambling success did not affect the pattern of results; therefore, all trials were included in the analyses. The data were analyzed in two ways. First, we calculated the mean confidence rating for each of the three conditions (see Table 1). As compared with the control condition in which the life event was preceded by Xs ( $M = 4.15$ ), unscrambling an unrelated anagram ( $M = 4.42$ ) significantly increased autobiographical memory confidence [ $t(71) = 2.39, p = .019$ ].<sup>2</sup> As compared with the control condition, unscrambling a related anagram ( $M = 4.37$ ) also significantly increased autobiographical memory confidence [ $t(71) = 2.20, p = .031$ ]. There was no difference between the unrelated and related anagram conditions ( $t < 1.0$ ). Thus, we obtained significant revelation effects in both anagram conditions.

Second, we fitted the SD mixture distribution model defined by Equations 1 and 2 to the  $3 \cdot 8 = 24$  rating category frequencies. Because Pearson  $\chi^2$  tests of association indicated heterogeneity of individual distributions for both experiments, we fitted the model to individual frequency data.<sup>3</sup> To ensure identifiability of our model, we placed three types of restrictions on the model parameters. First, as was outlined in our introduction, we assumed that participants utilized the same response criteria in each of the three blocks of trials corresponding to the three experimental conditions. Second, following a suggestion by Clarke,

Birdsall, and Tanner (1959), we assumed that participants used comparable confidence criteria on the negative side ( $x = 1, \dots, 4$ ) and on the positive side ( $x = 5, \dots, 8$ ) of the rating scale. The latter assumption results in the constraint that response criteria are spaced symmetrically around the boundary between categories 4 and 5 (i.e.,  $c_4$ ). To ensure identical zero points and units of measurement on the familiarity scale, we fixed  $c_1$  at 1.5 and  $c_4$  at 4.5 for each participant. Third, to reduce the number of parameters even further, we made use of the well-established result that the ratio of the standard deviations for new and old items tends to be fairly constant across studies (cf. Rotello, Macmillan, & Reeder, 2004, p. 590). In line with this result, we assumed that the ratio of the familiarity standard deviations of false and true events was constant across the three conditions [ $\sigma_f(j)/\sigma_t(j) = r$  for all conditions  $j$ , where  $r$  was allowed to differ between participants].

The restricted model is identifiable and includes 13 parameters for three conditions: six means ( $d_i$  and  $d'_i$ ), four parameters defining the six standard deviations ( $\sigma_i$  and  $\sigma'_i$ ), two distance parameters defining the response criteria ( $c_1, \dots, c_7$ ), and the proportion of true events ( $p$ ). Because each participant contributed  $3 \cdot (8 - 1) = 21$  independent rating frequencies to the combined data set, the goodness-of-fit test had  $21 - 13 = 8$  degrees of freedom.

We calculated minimum  $\chi^2$  parameter estimates (see Read & Cressie, 1988) for the observed raw frequencies of each participant by numerically minimizing the Pearson  $\chi^2$  statistic using Gegenfurtner's (1992) PRAXIS subroutine. To minimize the risk of convergence to local minima of the  $\chi^2$  function, the estimation process was repeated 10 times per participant, each time using new random start values that were independently drawn from uniform distributions across the permissible intervals of the parameters. The following permissible intervals were used for each participant and condition: [0.1, 0.9] for  $p$ , [0.5, 4.5] for  $d'_i$ , [4.5, 8.5] for  $d_i$ , [0.5, 4] for  $\sigma'_i$ , and [0.5, 1.5] for  $1/r$ —that is, the constant standard deviation ratio  $\sigma_i/\sigma'_i$ . Pearson's  $\chi^2$  was used as a goodness-of-fit statistic because it is defined even if there are empty cells in the raw data. Given a sample size of only  $N = 48$  per participant, the power of the  $\chi^2$  test is of course too low to perform formal goodness-of-fit tests at the individual level. However, the sum of the individual  $\chi^2$  statistics across all 72 participants provides a powerful goodness-of-fit test. This sum [ $\chi^2(576, N = 3,456) = 612.13, p > .05$ ] indicated an acceptable model fit.

The mean estimate of the  $p$  parameter was .54. Thus, on average, roughly half of the test items corresponded to true events according to our model-based analysis. The mean estimates (and standard errors) of  $d'_i, d_i, \sigma'_i, \sigma_i$ , and  $d'$  are summarized in the upper part of Table 2. As can be seen,  $d'_i$  tends to be larger in the unscrambling conditions, whereas the pattern is less clear for  $d_i$ . Importantly, as a consequence of the larger standard deviations, the discriminability parameter  $d'$ , which is measured in  $\sigma'_i$  units, is smallest in the related anagram condition.

Because the distributions of the parameter estimates were quite skewed, we performed Wilcoxon signed-rank

**Table 1**  
Mean Confidence Ratings and Standard Errors of the Mean (SEM) for Experiments 1 and 2

Condition	<i>M</i>	<i>SEM</i>
Experiment 1		
XXX	4.15	.10
Unrelated anagram	4.42	.11
Related anagram	4.37	.11
Experiment 2		
XXX	4.36	.10
Unrelated anagram	4.38	.13
Related anagram	4.48	.12

**Table 2**  
**Mean Parameter Estimates and Standard Errors of the Mean (SEM)**  
**for the Signal Detection Mixture Model (Experiments 1 and 2)**

Condition	Familiarity Means				Familiarity SDs				Discriminability	
	$d_f$		$d_t$		$\sigma_f$		$\sigma_t$		$d' = (d_t - d_f)/\sigma_f$	
	M	SEM	M	SEM	M	SEM	M	SEM	M	SEM
Experiment 1										
XXX	1.50	0.12	7.27	0.14	2.03	0.15	2.53	0.25	4.33	0.39
Unrelated	1.88	0.14	7.50	0.13	1.86	0.16	2.25	0.25	5.13	0.46
Related	1.79	0.14	7.38	0.13	2.18	0.16	2.69	0.25	4.10	0.41
Experiment 2										
XXX	2.20	0.17	7.48	0.14	1.77	0.16	1.87	0.25	4.52	0.47
Unrelated	2.16	0.17	7.35	0.17	1.80	0.13	1.88	0.26	4.88	0.56
Related	2.37	0.18	7.32	0.16	1.73	0.17	1.81	0.24	5.02	0.62

tests rather than *t* tests for the variables listed in Table 2. As compared with the control condition, the familiarity of false events,  $d_f$ , was significantly larger in the unrelated anagram condition ( $z = 2.35, p = .019$ ) and marginally significantly larger in the related anagram condition ( $z = 1.90, p = .057$ ). Note, however, that the latter effect would be significant at  $\alpha = .05$  with a one-tailed test in the predicted direction. The two anagram conditions did not differ significantly ( $z = 0.23$ ). These results closely resemble those previously found by Verde and Rotello (2003, 2004) for revelation effects on response bias in word-recognition tests. With respect to the familiarity of true events,  $d_t$ , none of the differences approached significance (all  $z$ s < 1.78).

Wilcoxon tests for estimates of the discriminability parameter showed that  $d'$  was significantly smaller in the related anagram condition than in the unrelated anagram condition ( $z = 2.14, p = .033$ ). None of the other differences in  $d'$  approached significance (all  $z$ s < 1.48). Because the differences in  $d' = (d_t - d_f)/\sigma_f$  might have been due to differences in  $\sigma_f$ , we also analyzed treatment effects on the familiarity standard deviations. As can be seen in Table 2,  $\sigma_f$  and  $\sigma_t$  are in fact larger in the related anagram condition compared to the unrelated anagram condition. This difference is marginally significant for  $\sigma_f$  ( $z = 1.90, p = .058$ ) and is significant for  $\sigma_t$  ( $z = 2.11, p = .035$ ). None of the other differences between standard deviations is significant (all  $z$ s < 1.47).

To summarize, Experiment 1 showed that unscrambling either related or unrelated anagrams prior to trying to remember a life event increases one's confidence that the event occurred in childhood. This increase in confidence mirrors that seen in recognition experiments involving the unscrambling of anagrams (Verde & Rotello, 2003, 2004; Watkins & Peynircioglu, 1990), prompting us to conclude that our results and those of Bernstein et al. (2004) and Bernstein et al. (2002) are indeed revelation effects. In addition, the results based on our new SD mixture model closely resemble those found by Verde and Rotello (2004) using standard signal detection theory in showing that solving unrelated anagrams produces familiarity illusions only, whereas solving related anagrams produces both familiarity illusions and a genuine decrease in memory accuracy.

## EXPERIMENT 2

Experiment 1 demonstrated that a revelation effect for childhood autobiographical memory is also obtained when participants unscramble either related or unrelated anagrams before trying to remember childhood events. In Experiment 1, we imposed a 10-sec deadline by which participants had to unscramble the anagram before immediately making their autobiographical memory ratings. If the revelation effect for autobiographical memory observed in Experiment 1 depends on shifts in familiarity associated with unscrambling, as we have argued elsewhere (see, e.g., Bernstein et al., 2004; Bernstein et al., 2002), then we reasoned that these shifts would be transient. In Experiment 2, we decreased the time pressure on our participants for switching between the two cognitive tasks by imposing a 20-sec delay between anagram unscrambling or viewing of Xs and the autobiographical memory rating task. We expected that the revelation effect obtained in Experiment 1 would disappear in Experiment 2.

### Method

**Participants.** Forty-eight University of Washington undergraduates participated for course credit.

**Stimuli and Procedure.** The stimuli and procedure were identical to those used in Experiment 1 except that after the initial 10-sec period in which participants either attempted to unscramble the anagram or viewed a series of Xs, they waited 20 sec before the putative life event appeared. Then they had 7 sec to rate the life event.

### Results and Discussion

Participants successfully unscrambled 82% of the anagrams. Unscrambling success had no effect on the pattern of results. Therefore, all trials were included in the analyses. Unlike in Experiment 1, no pairwise differences among confidence ratings were found between any of the three conditions in Experiment 2 (see Table 1).

Data analyses based on the SD mixture distribution model revealed that the sum of the individual Pearson  $\chi^2$  goodness-of-fit statistics [ $\chi^2(384, N = 2,304) = 426.88, p > .05$ ] again indicated an acceptable model fit. The mean  $p$  estimate for Experiment 2 was .50. The means and standard errors of the other parameter estimates are summarized in the lower part of Table 2. As can be seen from the table, differences between experimental condi-



tions are generally small. Neither the treatment effects on  $d_f$  and  $d_t$ , nor those on  $\sigma_f$  and  $\sigma_t$ , approach significance (all  $z$ s  $< .735$ ). The same holds for estimates of  $d'$  (all  $z$ s  $< .624$ ). Thus, when participants had to wait 20 sec between unscrambling an anagram and rating a life event, no revelation effect occurred (neither an effect on bias nor an effect on memory accuracy).

## GENERAL DISCUSSION

We have shown that unscrambling a word immediately before rating a life event increases people's confidence that the event occurred in their childhood (Experiment 1). We call this the *revelation effect for autobiographical memory* (Bernstein et al., 2004). We have also shown that this effect is short-lived: In Experiment 2, where participants unscrambled anagrams and then had to wait 20 sec before evaluating life events, no revelation effect emerged.

There is one very important difference between a standard recognition memory task—the task employed in most previous revelation experiments—and the autobiographical memory task studied in the present article. In recognition experiments, study and test lists are under the experimenters' control, enabling experimenters to estimate hits and false alarms. In the present work, we do not know which events did or did not occur in our participants' childhoods. Despite this limitation, we were able to profitably apply signal detection analyses to study childhood autobiographical memory using a mixture-model approach.

Although we cannot say definitively whether the effects observed in Experiment 1 were due to increases in hits or false alarms, or both, our findings are at least consistent with previous work demonstrating that there is a tendency to say that a word is more familiar (i.e., "old") after that word or an unrelated word is unscrambled. The revelation effect is typically stronger for false alarms than for hits (Hicks & Marsh, 1998). Consistent with this past finding are the results of our analyses based on the SD mixture distribution model, showing that unscrambling tends to increase the mean familiarities of false events more than those of true events. We therefore maintain that the increase in confidence observed in Experiment 1 likely reflects false memory to at least some degree. This effect, along with an increase in familiarity standard deviations in the related anagram condition as compared with the unrelated anagram condition, also decreases the discriminability of true and false events following the unscrambling of words related to the recognition probe. The same finding has previously been reported for word-recognition tasks (Mulligan & Lozito, 2006; Verde & Rotello, 2004).

Bernstein et al. (2002) have theorized that the revelation effect for autobiographical memory occurs as follows. Participants experience anagrams as dysfluent and establish an unconscious expectation that anagrams are hard to process. Upon either solving the anagram or not, they process the intact life event fluently. The mismatch between expectation (e.g., "this will be hard") and outcome (e.g., "wow, this life event is easy to read!") produces a sense of discrepancy. In the process of resolving

this discrepancy, participants identify particular events as being old because the unexpected fluency with which they process the intact events is mistaken for familiarity. Because familiarity is more likely to occur for events that are old than for events that are new, participants experience illusions of childhood autobiographical memory (see also Whittlesea & Williams, 2001). This fluency misattribution account can be reconciled with the results of our mixture-model analyses by assuming that the fluency induced by unscrambling unrelated anagrams affects familiarity additively, whereas unscrambling related anagrams affects familiarity multiplicatively. In the latter case, items that are initially high in familiarity are more strongly enhanced in their familiarity than are those that are initially low in familiarity, thus producing both a mean increase and an increase in the variability of familiarity. Both assumptions appear plausible, but further tests are needed.

In addition to showing how the revelation effect extends to autobiographical memory, we have demonstrated in the present article how mixture distribution models can be applied to autobiographical memory. Verde and Rotello (2004) developed their signal detection model of the revelation effect to account for data obtained in experiments in which the veracity of the participants' memories was known. We have shown how their methods can be extended and applied to autobiographical memory data in which the veracity of memory judgments is typically unknown. Although we cannot say which of our participants' particular memories were true and which were false, the fact that our manipulations influenced the hypothesized underlying familiarity distributions of true and false events in the same way that these manipulations have previously been shown to influence familiarities of actual true and false events leads us to conclude that our effects plausibly correspond to differential processing of true and false memories.

## AUTHOR NOTE

This work was supported by NIMH Individual National Research Service Award MH64264 to D.M.B. and by Grant SFB 504, project A12 from the Deutsche Forschungsgemeinschaft to E.E. The authors would like to thank Jeff Rouder for very helpful comments on a previous version of this manuscript. Correspondence concerning this article should be addressed to D. M. Bernstein, Department of Psychology, Kwantlen Polytechnic University, 12666-72nd Ave., Surrey, BC, V3W 2M8 Canada (e-mail: daniel.bernstein@kwantlen.ca) or E. Erdfelder, Lehrstuhl für Psychologie III, Universität Mannheim, Schloss EO 255, D-68131 Mannheim, Germany (e-mail: erdfelder@uni-mannheim.de).

## REFERENCES

- BERNSTEIN, D. M., GODFREY, R. D., DAVISON, A., & LOFTUS, E. F. (2004). Conditions affecting the revelation effect for autobiographical memory. *Memory & Cognition*, *32*, 455-462.
- BERNSTEIN, D. M., WHITTLESEA, B. W. A., & LOFTUS, E. F. (2002). Increasing confidence in remote autobiographical memory and general knowledge: Extensions of the revelation effect. *Memory & Cognition*, *30*, 432-438.
- CLARKE, F. R., BIRDSALL, T. G., & TANNER, W. P., JR. (1959). Two types of ROC curves and definitions of parameters. *Journal of the Acoustical Society of America*, *31*, 629-630.
- GEGENFURTNER, K. R. (1992). PRAXIS: Brent's algorithm for function minimization. *Behavior Research Methods, Instruments, & Computers*, *24*, 560-564.
- HICKS, J. L., & MARSH, R. L. (1998). A decrement-to-familiarity inter-

- pretation of the revelation effect from forced-choice tests of recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1105-1120. doi:10.1037/0278-7393.24.5.1105
- MACMILLAN, N. A., & CREELMAN, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.
- MOREY, R. D., PRATTE, M. S., & ROUDER, J. N. (2008). Problematic effects of aggregation in zROC analysis and a hierarchical modeling solution. *Journal of Mathematical Psychology*, **52**, 376-388. doi:10.1016/j.jmp.2008.02.001
- MULLIGAN, N. W., & LOZITO, J. P. (2006). An asymmetry between memory encoding and retrieval. Revelation, generation, and transfer-appropriate processing. *Psychological Science*, **17**, 7-11.
- READ, T. R. C., & CRESSIE, N. A. C. (1988). *Goodness-of-fit statistics for discrete multivariate data*. New York: Springer.
- ROTELLO, C. M., MACMILLAN, N. A., & REEDER, J. A. (2004). Sum-difference theory of remembering and knowing: A two-dimensional signal-detection model. *Psychological Review*, **111**, 588-616. doi:10.1037/0033-295X.111.3.588
- ROUDER, J. N., LU, J., SUN, D., SPECKMAN, P., MOREY, R., & NAVEH-BENJAMIN, M. (2007). Signal detection models with random participant and item effects. *Psychometrika*, **72**, 583-600. doi:10.1007/s11336-007-9013-4
- VERDE, M. F., & ROTELLO, C. M. (2003). Does familiarity change the revelation effect? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 739-746. doi:10.1037/0278-7393.29.5.739
- VERDE, M. F., & ROTELLO, C. M. (2004). ROC curves show that the revelation effect is not a single phenomenon. *Psychonomic Bulletin & Review*, **11**, 560-566.
- WATKINS, M. J., & PEYNIRCIOĞLU, Z. F. (1990). The revelation effect: When disguising items induces recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 1012-1020. doi:10.1037/0278-7393.16.6.1012
- WESTERMAN, D. L., & GREENE, R. L. (1998). The revelation that the revelation effect is not due to revelation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 377-386. doi:10.1037/0278-7393.24.2.377
- WHITTLESEA, B. W. A., & WILLIAMS, L. D. (2001). The discrepancy-attribution hypothesis: I. The heuristic basis of feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 3-13. doi:10.1037/0278-7393.27.1.3

#### NOTES

1. This rate of unscrambling success is lower than in most standard revelation experiments. Unlike in other revelation experiments in which the anagrams are a fixed length and participants have unlimited time to unscramble anagrams, in the present experiment, the anagrams ranged in length from 4 to 10 letters, and participants had a 10-sec deadline to unscramble anagrams.
2. The significance level  $\alpha = .05$  was used for all statistical analyses reported in the present article. All upper-tail probabilities ( $p$  values) refer to two-tailed tests.
3. Individuals are treated as fixed effects in our approach. For a random effects framework to handle variability between individuals and items in signal detection models, see Rouder et al. (2007) and Morey et al. (2008).

(Manuscript received February 27, 2007;  
revision accepted for publication December 10, 2008.)