

The revelation effect in frequency judgment

BRIAN H. BORNSTEIN and CRAIG B. NEELY
Louisiana State University, Baton Rouge, Louisiana

In three experiments, we explored the revelation effect in a frequency judgment task. Participants estimated the frequency of words that had been presented one, two, four, or eight times. At test, half the words were revealed by completing word fragments, and half were presented intact. Estimated frequencies were reliably higher for revealed than for intact words, and in two of the three experiments, the revelation effect became larger as actual frequency increased. A revelation effect was obtained whether the revealed word was the same as (Experiment 1) or different from (Experiment 2) the word judged for frequency. Frequency estimates were higher for more distorted test items (Experiment 3).

When participants engage in an interpolated task prior to making recognition judgments, they demonstrate a positive response bias toward those items—whether or not the items have actually been presented previously—in comparison with items judged in the absence of the interpolated task. This phenomenon is known as the *revelation effect* (Watkins & Peynircioğlu, 1990), because in most manipulations, the test items are “revealed” in some way after they have been distorted initially, such as by their being rotated or presented as anagrams or word fragments (Hicks & Marsh, 1998).

Factors that have been examined concerning the revelation effect include the type of interpolated task (e.g., Watkins & Peynircioğlu, 1990; Westerman & Greene, 1996, 1998), the degree of initial distortion (Watkins & Peynircioğlu, 1990), the effect of effort and time required by participants to perform the interpolated task (Luo, 1993; Peynircioğlu & Tekcan, 1993), and the type of recognition judgment (Cameron & Hockley, 2000; Hicks & Marsh, 1998). For example, Watkins and Peynircioğlu obtained revelation effects for words that were unfolded letter by letter, rotated, or presented as anagrams, as well as for numbers that were presented as Roman numerals (e.g., VII, V, etc.) or as equations (e.g., $2 + 5$, $9 - 4$, etc.). Revelation effects occur even when no words are presented at study (Frigo, Reas, & LeCompte, 1999), the study list and test are in different modalities (Westerman & Greene, 1996), the interpolated task reveals an item other than the test item (Westerman & Greene, 1996), or the recognition judgment follows an unrelated task that does not reveal any item (e.g., a memory span task for letters; Westerman & Greene, 1998).

Nearly all of the extant research on the revelation effect has been done with recognition tasks, yes–no item recognition having been the most common (e.g., Frigo et al., 1999; LeCompte, 1995; Luo, 1993; Peynircioğlu & Tekcan, 1993; Watkins & Peynircioğlu, 1990; Westerman & Greene, 1996, 1998). Interestingly, when variant recognition tasks are used, revelation effects are less likely to occur. For example, either no revelation effect or an “antirevelation” effect occurs with forced-choice recognition (Hicks & Marsh, 1998) or list discrimination tasks (Westerman & Greene, 1996), and Westerman (2000) found no effect when the test involved recognizing words’ singular/plural status. Cameron and Hockley (2000) and Westerman (2000) also failed to obtain a revelation effect for associative recognition under some test conditions.

In only two studies has the revelation effect been examined in nonrecognition contexts such as semantic (Watkins & Peynircioğlu, 1990) or frequency judgments (Watkins & Peynircioğlu, 1990; Westerman & Greene, 1996). For example, Watkins and Peynircioğlu asked participants to make a number of nonrecognition judgments about words that had first been either revealed (by being rotated) or not. These judgments included category typicality ratings, lexical decisions, and frequency estimates. With regard to the latter, participants estimated how frequently a set of words occurred in everyday life (Experiment 10) or had been encountered in the previous week (Experiment 11). Prior to their making these judgments, half of the words were rotated. Revelation did not affect participants’ judgments, which Watkins and Peynircioğlu attributed to the tasks’ nonepisodic nature. However, Westerman and Greene (1996, Experiment 3) also explored the revelation effect in a frequency judgment task, obtaining contradictory results. Participants viewed a list of words presented either once or twice. At test, each word was either presented normally or revealed by being presented first in anagram form, and participants then indicated its frequency (i.e., whether it had been presented once or twice). They were more likely to judge both the once- and twice-presented words as having been seen twice when they were revealed at test, in comparison with words that

This research was conducted by the second author in partial fulfillment of the requirements for the doctoral degree, under supervision of the first author. We are grateful for the assistance of Richard Highsmith in data collection, and we thank Jason Hicks, William Hockley, Bruce Whittlesea, and Jeff Wilson for their helpful comments on an earlier version of the manuscript. Address correspondence to B. H. Bornstein, Department of Psychology, 238 Burnett Hall, University of Nebraska, Lincoln, NE 68588-0308 (bbornstein2@unl.edu).

were presented normally at test. Although limited to only two situational frequencies (i.e., one or two), this experiment provided evidence for a revelation effect in frequency judgment.

A possible explanation for the discrepancy between the findings of Watkins and Peynircioğlu (1990) and Westerman and Greene (1996) is that the former concerned estimates of background frequency, whereas the latter concerned estimates of situational frequency. Because background frequency estimation is likely to draw on general knowledge and involve a greater degree of metacognition, estimates of situational frequency may rely more heavily on retrieval processes. Thus, it is not surprising that factors affecting the retrieval process of recognition, such as the revelation effect, would have a greater effect on situational than on background frequency judgments.

The present experiments replicate and extend the previous situational frequency experiment done by Westerman and Greene (1996), who found a revelation effect with words that varied in frequency from one to two occurrences. Limiting frequencies to one and two may not represent the full range of frequency estimation. The present experiments included frequencies varying from one to eight. The inclusion of more frequencies than those used by Westerman and Greene (1996) offers insight into the possible interaction between frequency and revelation. It was hypothesized that revealing words in a frequency judgment task would increase participants' frequency estimates for those words relative to intact (i.e., nonrevealed) words.

EXPERIMENT 1

Method

Participants. The participants were 40 undergraduates who received extra course credit.

Materials and Design. The stimuli were words from the Gibson and Watkins (1988) word pool. All words were eight letters long, with linguistic frequency between one and eight per million. The words were randomly assigned to the conditions to avoid any effect of linguistic frequency. The participants saw a list of 100 words (20 words shown once, 20 words shown twice, 20 words shown four times, 20 words shown eight times, and 20 untested buffers), for a study list comprising 320 total word presentations. The order of the presentations was random, with the constraint that any two presentations of the same word were always separated by at least three other words. Also, the words were counterbalanced across frequency conditions so that a given word appeared at each presentation frequency equally often.

Half of the words were shown normally at test (*intact* condition), and the other half of the words were revealed by unfolding them as word fragments (e.g., the word *elephant* shown as e _ e _ _ _ _ _ , then as e _ e p _ _ n _ , then as e l e p h _ n _ , and finally as e l e p h a n t). The design, therefore, was a 4 (actual frequency) \times 2 (test format) within-participant design.

Words were randomly assigned to be shown intact or revealed, and the order in which intact and revealed items appeared was randomly determined. For counterbalancing purposes, if a word was intact for half of the participants, it was revealed for the other half of the participants, and vice versa. The counterbalancing of test format was crossed with the counterbalancing of frequency so that

each test item appeared at each actual presentation frequency, both intact and revealed, to create eight distinct counterbalanced conditions.

Procedure. The participants viewed each word in the study list in a black typeface presented on a white background. The presentation was projected onto a screen so that the actual size of the letters on each slide was approximately 1 ft in height. Each word was shown for 2 sec (Watkins & Peynircioğlu, 1990). The participants were asked to mouth each word as it was shown, to ensure that they were paying attention to the task. Following the study list, the participants viewed the 80 critical words one at a time on slides. All items, revealed and intact, at test were viewed for 10 sec. Participants were asked to write down the word that was presented (for intact items) or to solve the word fragment and then write the solution in a designated place on the answer sheet (for revealed items), followed by an estimation of the word's frequency. No limit was given regarding the estimation. The participants were simply told to write down the number of times that they remembered having seen the word, and if they did not remember having seen the word, then a reply of "zero" was acceptable.

Results and Discussion

A graph of the results can be seen in Figure 1. The dependent variable in this analysis was the mean numerical estimate that was given by each participant for each condition.¹ A 4 \times 2 repeated measures analysis of variance (ANOVA) was implemented to analyze the effects of frequency (one, two, four, or eight) and test format (intact vs. revealed).

Participants gave higher estimates to higher actual frequencies, resulting in a significant main effect of frequency [$F(3,39) = 56.4, p < .01$]. As predicted, participants also gave higher estimates to words that were revealed ($M = 4.08$) than to normal words ($M = 3.16$), resulting in a significant effect of test format [$F(1,39) = 44.37, p < .01$]. There was also a significant test format \times frequency interaction [$F(3,39) = 10.88, p < .01$]. Simple comparisons between revealed and intact items at each frequency level indicated statistically significant differences at every frequency [$F_s(1,39) < 6.10, p_s < .05$]. The interaction appears to have been due to the greater discrepancy between the revealed and intact estimates as frequency increased (see Figure 1). At the higher frequencies, participants had more latitude in estimation because of the larger numbers. In other words, at lower presentation frequency there was less room for the revelation effect to work.

EXPERIMENT 2

Westerman and Greene (1996, Experiment 3) obtained a revelation effect even when the word that was revealed at test was different from the word that was judged for recognition. In Experiment 2, we applied this methodology to a frequency judgment task.

Method

Participants. The participants were 40 undergraduates who received extra course credit.

Materials and Design. The materials and design were identical to those of Experiment 1, with one exception: An additional 80 words were used at test which appeared neither in the study list nor

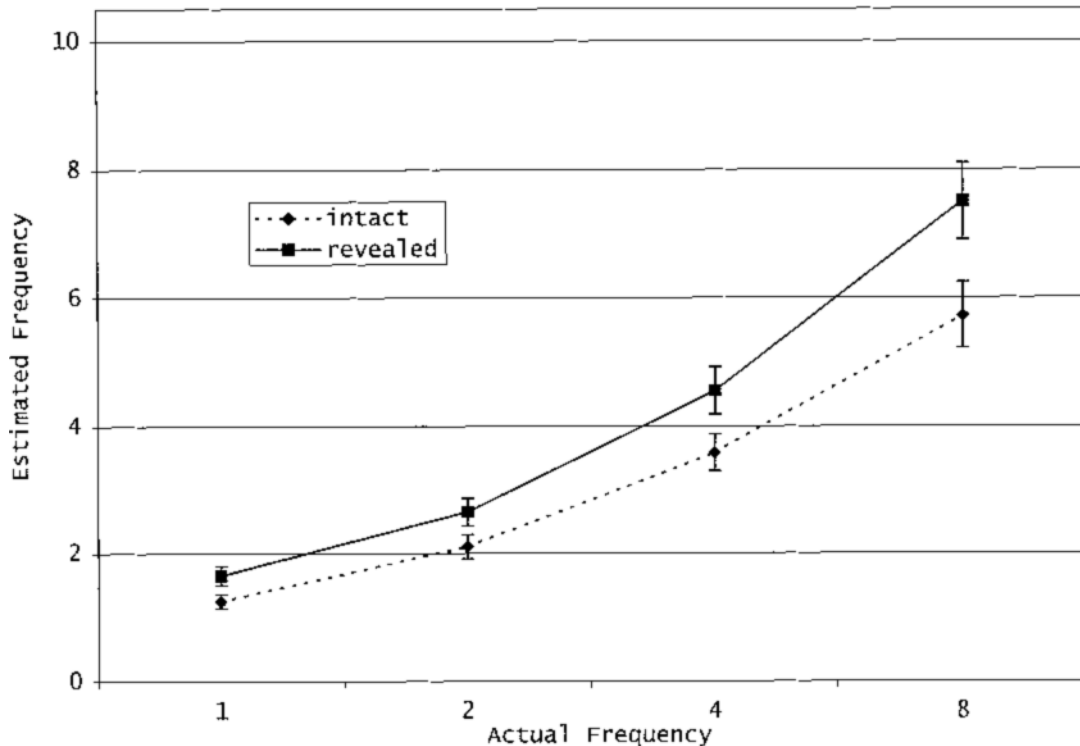


Figure 1. Mean frequency estimates (with standard errors) by actual frequency and test condition, Experiment 1.

as test items. They were simply presented, intact or revealed, immediately before each test item. These additional 80 words were also chosen from the Gibson and Watkins (1988) word pool.

Procedure. The procedure was identical to that in Experiment 1, with one exception. In the revealed condition, a word that was different from the word that was judged for frequency was revealed prior to the frequency estimation task. The participants were asked to solve the fragment of that word and write the solution in a designated place on the answer sheet. Next, the participants were shown a different word that was in fact presented on the study list and were asked to estimate the frequency of that word. For example, the word *vineyard* was unfolded; the participants were asked to write it down, and they were then asked how often the word *raindrop* had appeared on the earlier list. Thus, the word that was judged for frequency was not the word that was revealed just prior to the frequency estimate. The procedure in the intact condition was the same as that in Experiment 1.

Results and Discussion

A 4×2 ANOVA was implemented to analyze the effects of actual frequency (one, two, four, or eight) and the test format manipulation (revealed vs. intact). Items with higher actual frequencies were given higher estimates, resulting in a main effect of frequency [$F(3,39) = 68.01, p < .01$]. In addition, revealed items were given higher estimates ($M = 3.95$) than were intact items ($M = 3.25$), resulting in a significant main effect of test format [$F(1,39) = 53.01, p < .01$]. There was also a significant test format \times frequency interaction [$F(3,39) = 7.33, p < .01$]. Simple comparisons indicated statistically significant revelation effects at each frequency level [$F_s(1,39) > 5.6, p_s <$

.05]. The pattern of this interaction was the same as in Experiment 1, with larger differences between frequency estimates at higher actual frequencies (see Figure 2).

Since the revealed word was different from the word judged for recognition, obtaining a revelation effect in the present paradigm suggests that revelation does not enhance familiarity for a specific word (cf. LeCompte, 1995). However, it is consistent with Westerman and Greene's (1998) theory that revelation globally enhances familiarity for all similar stimuli. This global familiarity explanation has also been offered for frequency estimation (see, e.g., Hintzman, 1988). If revealing a word—even an unrelated word—increases the familiarity of words in general, then it would inflate participants' frequency estimates for revealed test stimuli.

EXPERIMENT 3

Experiments 1 and 2 provided evidence for the existence of a relationship between the revelation effect and frequency judgment. In Experiment 3, we examined one of the specific properties of the revelation effect in a frequency judgment paradigm. Watkins and Peynircioğlu (1990) found that the degree of distortion of the revealed word was positively correlated with the tendency to give positive responses to both target and lure words during a recognition test. In other words, presenting a starting fragment with only two letters prior to unfolding, compared with presenting a starting fragment with six letters, pro-

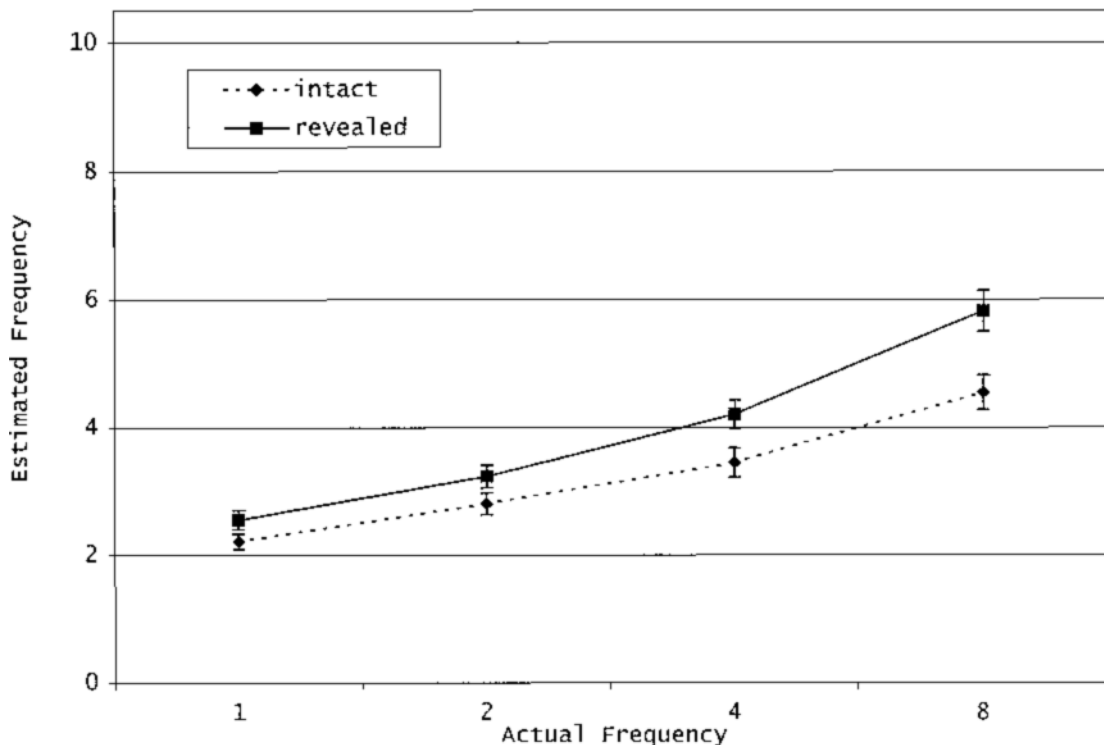


Figure 2. Mean frequency estimates (with standard errors) by actual frequency and test condition, Experiment 2.

duced a greater revelation effect. In Experiment 3, we replicated this methodology, using a frequency estimation paradigm. Because the findings from the first two experiments on revelation effects in frequency estimation were analogous to previous effects of revelation on recognition, we hypothesized that the effect of revelation on frequency judgments would be greater as the degree of distortion increased.

Method

Participants. The participants were 40 undergraduates who received extra course credit.

Materials and Design. The materials and design were identical to those used in Experiment 1, with one exception. During the test phase, the revealed words varied in the number of letters in the starting fragment, with two (e.g., e _ e _ _ _ _ _), four (e _ e _ h _ n _), or six (e l e _ h _ n t) letters present. The intact words were presented with all eight letters shown (e.g., e l e p h a n t). The position of the letters that were present in the fragments was based on the following algorithm: letter positions 3 and 5; then 1 and 7 added; then 2 and 8 added; and finally all 8 letters.

Each word was presented one, two, four, or eight times. At test each word was assigned randomly to the two-letter-fragment, four-letter-fragment, six-letter-fragment, or intact condition. The words were counterbalanced across conditions so that all words served in each frequency and test format condition equally often.

Procedure. The procedure was identical to that in Experiment 1.

Results and Discussion

The results are shown in Table 1. A 4×4 repeated measures ANOVA was implemented to analyze the ef-

fects of frequency (one, two, four, or eight) and test format: two-, four-, six-, or eight-letter-fragment (i.e., intact).

As with the previous experiments, higher actual frequencies resulted in higher frequency estimates, yielding a significant main effect of frequency [$F(3,39) = 638.58, p < .01$]. Results also indicated a significant main effect of the number of starting letters in the test fragment [$F(3,39) = 2.82, p < .05$]. Frequency estimates were lowest when the test items were not distorted at all (i.e., in the intact condition), greatest with maximum distortion (i.e., in the two-letter fragment condition), and intermediate with moderate degrees of distortion. A contrast comparing the three revealed conditions (two-, four-, and six-letter fragments) with the intact condition fell just short of conventional levels of statistical significance [$F(1,39) = 3.37, p = .07$]. Thus, just as Watkins and Peynircioğlu (1990) found a positive relationship between recognition judgments and test items' degree of distortion—that is, the more distortion, the more likely participants were to say that they had seen the word previously—there was a positive relationship between degree of distortion and frequency estimation as well.

The test format \times frequency interaction was only marginally significant [$F(9,39) = 0.97, p < .10$], indicating that the effect of degree of distortion was not reliably greater at higher than at lower actual frequencies. A possible explanation for this lack of an interaction is that the present experiment lacked sufficient power to detect an interaction. Because a larger number of test conditions

Table 1
Mean Frequency Estimates by Actual Frequency and Number of Initial Letters in Revealed Test Item Word Fragments

No. of Initial Letters in Word Fragment	Actual Frequency				Total
	1	2	4	8	
8 (intact)	2.18	3.16	6.17	9.30	5.20
6	2.21	3.99	5.69	9.88	5.44
4	2.35	3.80	5.66	9.54	5.34
2	2.35	4.05	6.80	10.05	5.81
Total	2.27	3.75	6.08	9.69	5.45

were included but the total number of words was not increased, there were only 5 words in each condition (20 of each test format/4 actual frequencies), compared with 10 words in each condition in Experiments 1 and 2.

GENERAL DISCUSSION

The purpose of the present studies was to examine the effect of revealing test items on frequency estimation. In three experiments, we replicated and extended, to higher situational frequencies, the research conducted by Westerman and Greene (1996) on frequency estimation and the revelation effect. Across four actual frequency levels (one, two, four, and eight), participants gave higher estimates to words that were revealed than to words that were intact at test, whether the word that was revealed at test matched the word that was judged for frequency (Experiments 1 and 3) or not (Experiment 2). In addition, the revelation effect was greater as objective frequency increased in two of the three experiments.

In Experiment 3, we examined the relationship between the amount of distortion of a revealed item and the magnitude of the revelation effect in frequency estimation. The variability of distortion has been shown to affect the magnitude of the revelation effect in standard recognition tasks (Watkins & Peynircioğlu, 1990). Although the frequency \times test format interaction was only marginally significant, the results of Experiment 3 did show a significant revelation effect across frequency conditions.

Watkins and Peynircioğlu (1990) failed to obtain a revelation effect in frequency judgment, but this failure may have been due to their use of a background frequency judgment task, whereas the present experiments, like that of Westerman and Greene (1996), concerned estimates of situational frequency. By its very nature, background frequency estimation is likely to draw on general semantic knowledge and involve a greater degree of metacognition, whereas the more constrained circumstances of situational frequency estimation draw comparatively more heavily on the retrieval of specific instances, thereby reflecting episodic knowledge. Obtaining a revelation effect with a relatively episodic (i.e., situational) frequency judgment task, but not with a relatively seman-

tic (i.e., background) frequency judgment task, is congruent with its characterization as an episodic memory phenomenon (Watkins & Peynircioğlu, 1990).

Although previous research has found scant evidence of the revelation effect in tasks that do not involve yes-no item recognition (Cameron & Hockley, 2000; Frigo et al., 1999; Watkins & Peynircioğlu, 1990; Westerman, 2000; Westerman & Greene, 1996), the present results provide strong evidence for an effect of a retrieval phenomenon—that is, the revelation effect—on the estimation of situational frequencies. Thus, the revelation effect appears to be a fairly broad and robust phenomenon that is not limited to a recognition test paradigm.

REFERENCES

- CAMERON, T. E., & HOCKLEY, W. E. (2000). The revelation effect for item and associative recognition: Familiarity versus recollection. *Memory & Cognition*, *28*, 176-183.
- FRIGO, L. C., REAS, D. L., & LECOMPTE, D. C. (1999). Revelation without presentation: Counterfeit study list yields robust revelation effect. *Memory & Cognition*, *27*, 339-343.
- GIBSON, J. M., & WATKINS, M. J. (1988). A pool of 1,086 words with unique two-letter fragments. *Behavior Research Methods, Instruments, & Computers*, *20*, 390-397.
- HICKS, J. L., & MARSH, R. L. (1998). A decrement-to-familiarity interpretation of the revelation effect from forced-choice tests of recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *24*, 1105-1120.
- HINTZMAN, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528-551.
- LECOMPTE, D. C. (1995). Recollective experience in the revelation effect: Separating the contributions of recollection and familiarity. *Memory & Cognition*, *23*, 324-334.
- LUO, C. R. (1993). Enhanced feeling of recognition: Effect of identifying and manipulating test items on recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 405-413.
- PEYNIRCIOĞLU, Z. F., & TEKCAN, A. I. (1993). Revelation effect: Effort or priming does not create the sense of familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 382-388.
- WATKINS, M. J., & PEYNIRCIOĞLU, Z. F. (1990). The revelation effect: When disguising test items induces recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 1012-1020.
- WESTERMAN, D. L. (2000). Recollection-based recognition eliminates the revelation effect in memory. *Memory & Cognition*, *28*, 167-175.
- WESTERMAN, D. L., & GREENE, R. L. (1996). On the generality of the revelation effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *22*, 1147-1153.
- 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.

NOTE

1. Variance increased with increasing presentation frequency in all experiments, with tests of sphericity revealing significant effects. Therefore, the *F* values reported in all experiments reflect the modified *F* values from a Greenhouse-Geisser adjustment to account for the variance issue.

(Manuscript received April 4, 2000;
 revision accepted for publication July 31, 2000.)