Journal of Experimental Psychology:
Learning, Memory, and Cognition
1982, Vol. 8, No. 4, 279-288

0278-7393/82/0804-0279$00.75

Overlearning and the Feeling of Knowing

Thomas O. Nelson, R. Jacob Leonesio,
Arthur P. Shimamura, and Robert F. Landwehr
University of Washington

Louis Narens
University of California, Irvine

Previous research on the feeling of knowing for nonrecalled items has focused on the validity of the feeling of knowing as a predictor for subsequent performance but generally has overlooked the question of what factors the feeling of knowing might be based upon. We explored the hypothesis that one factor underlying the feeling of knowing is the degree of prior learning for the sought-after item. Undergraduates learned a list in which various items were acquired to a criterion of either one correct recall, two correct recalls, or four correct recalls. Four weeks later, the subjects had a recall-retention test, then rank ordered the nonrecalled items in terms of the relative feeling of knowing for each of the nonrecalled items, and finally had a recognition test on the nonrecalled items. A major finding (in addition to those summarized at the end of the article) was that the magnitude of the feeling of knowing for nonrecalled items increased with the degree of prior learning.

Empirical research on the feeling of knowing—operationally defined in terms of the subject's predicted recognition for nonrecalled items—was initiated approximately 15 years ago by Hart (1965, 1967). Since that time, many investigators have examined the validity of the feeling of knowing for predicting subsequent performance (e.g., for predicting recognition, for predicting cued recall, for predicting performance on a second uncued recall test). By contrast, the present research explored the issue of what underlies a person's feeling-of-knowing judgments. Such research seems timely because, as Nickerson (1980) pointed out,

> Relatively little attention has been given to the question of what the feeling of knowing may be based on, and to date there is not a very satisfactory explanation of how it is that one can be sure that something is in memory when one is unable to retrieve it. (p. 362)

One factor that might underlie the feeling of knowing for a nonrecalled item is the degree of prior learning for that item. For instance, if a person does not remember ever having learned the item in the past (e.g., "What is Ronald Reagan's telephone number?"), then he or she might have a low feeling of knowing for the item.[1] Conversely, when a person remembers having learned the item in the past, then he might have a higher feeling of knowing for the item. Just as the degree of learning is a classic factor for subsequent memory of the actual to-be-retrieved item (e.g., Krueger, 1929; Postman, 1962), perhaps it also affects the memory of having encountered the to-be-retrieved item and provides a basis for the feeling of knowing.

A major goal of the present study was to test the hypothesis that the degree of prior learning is one factor that underlies the feeling of knowing. Another major goal was to determine some of the basic relationships between the degree of learning, the feeling of knowing, and various aspects of retention performance.

This research was supported by U.S. Public Health Service Grant MH-32205. We thank G. Grinnell, K. Schuler, and M. Vincent for assisting with the data collection.

Requests for reprints should be sent to Thomas O. Nelson, Department of Psychology, University of Washington, Seattle, Washington 98195.

[1] Hereinafter, "he or she" is referred to more briefly as "he."

Method

Use of a New Feeling-of-Knowing Methodology

The traditional methodology for investigating the feeling of knowing (Hart, 1965, 1967) has entailed YES–NO feeling-of-knowing ratings (either in terms of two degrees of fineness, such as YES and NO, or in terms of more than two degrees of fineness, but nevertheless always in terms of absolute judgments). Such absolute judgments are intrinsically confounded with the subject's placement of his YES–NO criterion. This potentially problematic source of variance in where the subject places his YES–NO criterion can be circumvented by using the feeling-of-knowing methodology developed recently by Nelson and Narens (1980a). An important aspect of their procedure is that the subject makes relative judgments rather than absolute judgments, thereby minimizing the effect of where he places his criterion of saying YES to a question of the absolute degree of feeling of knowing (cf. the use of N-alternative forced-choice recognition tests rather than YES–NO recognition tests—Shepard, 1967).

In the case of feeling-of-knowing judgments, Nelson and Narens (1980a) provided a procedure (now available as a computer program called FACTRETRIEVAL—Shimamura, Landwehr, & Nelson, 1981) in which the subject makes pairwise comparisons for a set of nonrecalled question cues and ranks them in terms of his feeling of knowing for recognizing the correct answer to each question. For instance, the subject might be asked whether he is more likely to recognize the correct answer to the question "What is the name of the man who assassinated Abraham Lincoln?" or to recognize the correct answer to the question "What city is the capital of Finland?" By ranking all of the various nonrecalled items in this way, the eventual result is an ordinal scale of the feeling of knowing for that set of items. Subsequently, the items can be tested for recognition, and recognition performance can be compared against the feeling-of-knowing ranking (i.e., items high in the feeling-of-knowing rank order should be more likely to be recognized than items lower in the feeling-of-knowing rank order).[2] The present study used the shorthand version (Nelson & Narens, 1980a, p. 78) in which, rather than presenting the subjects with two items at a time to rank, the subjects are presented with all of the nonrecalled items to rank at the same time.

Subjects and Design

The subjects were 50 University of Washington undergraduates whose participation partially fulfilled a course requirement. The design consisted of repeated measures on the within-subjects variable of degree of learning. For each subject, one third of the items occurred until each item was recalled exactly once, another third occurred until each item was recalled exactly twice, and the remaining third occurred until each item was recalled exactly four times.

Items

The items were 18 number–noun pairs of the form 48–DOLLAR. The 18 stimulus terms were drawn from a set of 20 two-digit numbers of low association value (range = 1.54–1.99) in the Battig and Spera (1962) norms. The 18 response terms were drawn from a set of 20 A and AA nouns from the Thorndike and Lorge (1944) norms. A list was derived by randomly drawing 18 stimulus terms from the stimulus pool, by randomly drawing 18 response terms from the response pool, and by randomly pairing stimulus terms with response terms. Eighteen such lists were derived and randomly assigned to subjects (with each list assigned approximately three times across the 50 subjects).

Apparatus

All stimuli were presented via an Apple II microcomputer and displayed on a television monitor. Subjects entered responses on the computer keyboard. Instructions, data collection, and statistical analyses were all computer controlled.

Procedure

The experiment can be subdivided into four discrete stages. First, the subject learned each paired associate to one of three criteria, which varied within subjects (acquisition). Four wk. later, the subject was given a recall test (retention). Then the subject provided a rank ordering of his feeling of knowing for those items not recalled in the retention test (feeling of knowing). Finally, a recognition test was given for each of the nonrecalled items (recognition). Each of these stages is discussed separately.

Acquisition. Each number–noun pair was randomly assigned (anew for every subject) to be learned to a criterion of one, two, or four correct responses, with the only restriction being that six pairs were assigned to each of the three criterion levels. The 18 pairs were randomly ordered and shown individually at a 16-sec rate during which the subject read aloud the number-word pair once and then remained silent until the next item. Following study, 26 sec of backward number counting occurred to minimize recall from short-term memory. Then the test phase began. Each number stimulus was presented, and the subject attempted to recall the correct noun response. Test trials were self-paced, with the subject typing each response onto the computer keyboard. A response was scored as correct if the first three letters of the subject's response matched the correct noun. (No two nouns used in this experiment had the same first three letters.) After each study–test sequence, all pairs that had reached their criterion of correct responses were removed. Thus, the next study–test sequence included only those items that had not reached criterion. This procedure continued until all items were

[2] This procedure can be used either with general-information questions (e.g., from the norms of Nelson & Narens, 1980b) or with traditional laboratory paired associates (e.g., number–word pairs). Although the former have their own advantages, the latter have the particular advantage of allowing control over the acquisition history for each item, which is especially important for manipulating the degree of learning; consequently, the latter were employed here.

learned to their preassigned criterion. The order of presentation for both the study and test phases was randomized. Prior to the main experiment, there was a warm-up task on the acquisition phase consisting of one study–test trial on three letter–letter pairs, using the same procedure as in the main experiment.

Retention. Without having been informed in advance of the impending retention test, the subject returned after 4 wk. and was given a recall test for each of the 18 previously learned pairs, using the same procedure as in the test phase of acquisition. The subject was asked to guess at the correct response when unsure. If the subject had absolutely no idea about the answer, he typed NEXT and proceeded to the next item. The subject was not yet informed about the subsequent feeling-of-knowing phase or the subsequent recognition phase. Latency of response was measured from the time the stimulus number was presented until the subject completed typing his response and signalled for the next item. Because the eventual four-alternative forced-choice recognition test would require four nonrecalled responses as the recognition alternatives (see Recognition section), subjects with fewer than four incorrect responses were eliminated from the experiment and replaced; this occurred for only 2 of the 50 subjects.

Feeling of knowing. Next, the subject was instructed to rank the stimulus terms of all pairs for which the correct response was not produced during the retention phase, with highest ranking given to items thought most likely to be recognized from among a set of four recognition alternatives (highest feeling of knowing). The subject was told that he had incorrectly recalled all of the to-be-ranked items and that subsequently he would receive a recognition test on them. All stimuli appeared simultaneously on the right side of the monitor in random order in a circular array. When the subject typed in a stimulus number and the desired feeling-of-knowing rank number for that item, the stimulus number disappeared from the circular array and reappeared on the left side of the monitor in the selected position of the feeling-of-knowing rank order. This process continued until all stimulus numbers had been moved from the circular array to the feeling-of-knowing rank order. The subject assigned ranks in any order desired and was given an opportunity to change responses. When satisfied with all rankings, the subject initiated the next stage.

After completing the feeling-of-knowing judgments on a relative basis (i.e., one item relative to another item), the subject made absolute feeling-of-knowing judgments in terms of whether or not he would recognize a given item (cf. Hart, 1965, 1967). The first absolute feeling-of-knowing judgment was made on the middlemost item from the relative feeling-of-knowing ranking. If this item was judged to be correctly recognizable, the second judgment was made on the middlemost item from the upper half of the relative ranking; if it was judged to be not correctly recognizable, then the second judgment was made on the middlemost item from the lower half of the relative ranking. This bisection process continued until a point occurred where one item in the relative ranking was judged to be correctly recognizable and the adjacent item was judged to be not correctly recognizable (or where the uppermost item was judged to be not recognizable or where the lowest item was judged to be recognizable—these extremes of absolute confidence over the entire ranking occurred only one time for the former and seven times for the latter). This method of obtaining absolute feeling-of-knowing judgments will be referred to as the Hart-like procedure.

Recognition. Finally, a four-alternative forced-choice recognition test was given for each of the nonrecalled pairs. Distractors were three other nouns randomly chosen from the nonrecalled items. Latency of recognition, as well as correctness of recognition, was recorded.

Results and Discussion

For each dependent variable of interest, we report the mean across subjects as the descriptive statistic to summarize group performance. However, it should be noted that for every reliable effect the ordering of conditions was the same regardless of whether the descriptive statistics summarizing the central tendency across subjects were means or medians. To maximize the generality of the inferential conclusions, we report (unless noted otherwise) only the results from distribution-free nonparametric tests (e.g., the Friedman test for comparing performance across within-subjects conditions).[3]

To help clarify the substantial number of dependent measures, Figure 1 shows the protocol obtained from one subject. After obtaining this kind of protocol from each of the 50 subjects, a series of analyses was conducted on the following topics: (a) the effects of the degree of learning, (b) within-subjects correlations of the feeling of knowing with other measures, (c) Hart-like analyses of the feeling of knowing, (d) search termination, and (e) across-subjects individual-differences correlations. Each topic is discussed in turn.

Effects of Degree of Learning

Table 1 shows the effect of degree of learning on various dependent variables from acquisition (first row), recall retention (second through fourth rows), feeling of knowing (fifth row), and recognition (sixth through eighth rows).

[3] Unless noted otherwise, every variable yielding significant results in the nonparametric test also yielded significant results in the corresponding parametric test (with significant defined as $p < .05$), and every variable yielding nonsignificant results in the nonparametric test also yielded nonsignificant results in the corresponding parametric test.

Acquisition. There was no reliable difference across the three conditions in the number of trials up to and including the first correct acquisition response, $\chi^2(2) = .67$. Also, once an item was correctly recalled during acquisiton, that item was very likely to be correctly recalled again on any subsequent tests during the acquisition phase;

Item		Dropout	Acquisition Trials					Test Recall (Second Session) Score	Response	Retention Latency
17	SCHOOL	2	C	C	–	–	–	C	SCHOOL	7.06
26	COTTON	1	C	–	–	–	–	C	COTTON	44.55
28	FAMILY	1	C	–	–	–	–	W	FRIEND	26.4
30	LAKE	1	W	C	–	–	–	W	MILE	36.1
32	MUSIC	2	W	C	C	–	–	C	MUSIC	6.83
35	DOLLAR	1	W	W	C	–	–	C	DOLLAR	7.74
40	HORSE	2	C	C	–	–	–	W	FLOWER	31.95
44	ENEMY	4	C	C	C	C	–	C	ENEMY	6.7
48	MILE	2	W	C	C	–	–	W	DAY	8.56
55	PARTY	2	C	C	–	–	–	C	PARTY	8.97
56	HOME	1	W	C	–	–	–	W	GUESS	18.82
70	DOCTOR	4	W	C	C	C	C	C	DOCTOR	5.48
76	BOOK	4	C	C	C	C	–	C	BOOK	9.12
77	VALLEY	2	C	C	–	–	–	C	VALLEY	7.85
80	UNCLE	4	C	C	C	C	–	W	NEXT	34.87
81	TABLE	1	C	–	–	–	–	W	FAMILY	48.88
95	CHURCH	4	C	C	C	C	–	C	CHURCH	10.36
97	FLOWER	4	C	C	C	C	–	C	FLOWER	15.62

Rank	Item	Dropout	Recognition	Recognition Latency
7 (high)	80 UNCLE	4	C	6.12
6	48 MILE	2	C	13.09
5	56 HOME	1	C	12.52
4	40 HORSE	2	W	14.31
3	30 LAKE	1	W	65.84
2	28 FAMILY	1	C	9.35
Subject's predicted cutoff - - - - - - - -				
1 (low)	81 TABLE	1	C	16.87

Dropout Criterion	Median Rank	Number Recalled	Proportion Recognized
1	2.5	2	.75
2	5	4	.5
4	7	5	1

Figure 1. Protocol obtained from one subject. (The upper-left portion shows the 18 number–word paired-associate items, the dropout criterion for each item [1, 2, or 4 correct recalls], and acquisition performance [W = wrong, C = correct]. The upper-right portion shows recall–retention performance [wrong vs. correct, subject's response, and response latency]. The middle portion shows feeling-of-knowing rank for the 7 nonrecalled items, recognition accuracy [wrong vs. correct], and recognition latency. The lower portion shows median feeling-of-knowing rank, number of items recalled on the retention test [out of 6 possible], and proportion of nonrecalled items that were recognized for each of the three dropout conditions.)

Table 1
Effect of Degree of Learning (Number Correct During Acquisition) on Various Dependent Variables

	Number of times correct during acquisition		
Dependent variable from each subject	1	2	4
Median trials to first correct acquisition response	2.1	2.2	2.1
Percentage correct on recall retention[a]	18.2	28.7	39.3
Median latency of correct recall retention	11.3	10.7	8.6
Median latency of incorrect recall retention	21.8	21.5	20.3
Median feeling of knowing rank[a]	5.8	6.8	8.4
Percentage correct on recognition retention[a]	59.4	65.4	76.3
Median latency of correct recognition retention[a]	12.6	10.9	9.1
Median latency of incorrect recognition retention	13.7	14.1	14.9

Note. Each entry is the mean across subjects of the dependent variable listed in the first column.
[a] Reliable effect at $p < .05$.

the mean probability of correct recall for an item after that item had already been correctly recalled was .96.

Recall retention. The number of times that an item had been recalled correctly during acquisition produced reliable differences in subsequent recall retention, $\chi^2(2) = 18.99$; a sign test showed that four-correct items were subsequently more likely to be recalled than two-correct items ($p < .01$), and two-correct items were more likely to be recalled than one-correct items ($p < .05$).

The latency of correct recall retention tended to decrease with the degree of learning, but this trend was not reliable, $\chi^2(2) = 3.27$. Of greater importance, the degree of learning had a negligible effect on the latency of incorrect recall retention, $\chi^2(2) = .01$; this lack of effect (and, in particular, the failure of four-correct items to have longer incorrect-recall latencies than two-correct items or one-correct items) is discussed in the section on search termination.

Feeling of knowing. Perhaps the most important result is that the degree of learning during acquisition produced reliable effects on the subsequent feeling of knowing for items not recalled during the retention test, $\chi^2(2) = 18.99$. A sign test showed a reliably greater feeling of knowing for four-correct items than for two-correct items ($p < .01$) and a reliably greater feeling of knowing for two-correct items than for one-correct items ($p < .05$). (Similar to the means shown in Table 1, the median across subjects of the individual-subject median feeling-of-knowing ranks was 5.5 for one-correct items, 6.25 for two-correct items, and 8.0 for four-correct items.) These results confirm the hypothesis that the feeling of knowing for nonrecalled items is based on the degree of prior learning of the sought-after information.

Recognition of nonrecalled items. The degree of learning affected the recognition of nonrecalled items such that nonrecalled items with a high degree of prior learning were more likely to be recognized than nonrecalled items with a lower degree of prior learning, $\chi^2(2) = 6.61$. This result occurred in spite of selection effects that would tend to produce the opposite pattern of results.

The latency of correct recognition decreased as the degree of prior learning increased, $\chi^2(2) = 17.78$. This result cannot be attributed to some general strategy during recognition testing, because the same decreasing trend did not occur for the latency of incorrect recognition, $\chi^2(2) = .32$.

Within-Subjects Correlations of the Feeling of Knowing with Other Measures

For a given subject, the Goodman-Kruskal gamma correlation between the feeling of knowing and each of a number of other measures was computed across that subject's nonrecalled items. Since the gamma correlation has a direct interpretation in terms of the scale of probability, one can meaning-

Table 2
Mean and 95% Confidence Interval for the Within-Subjects Gamma Correlations Between the Feeling-of-Knowing Rank and Selected Dependent Variables

Dependent variable correlated with feeling-of-knowing rank	Mean correlation	95% confidence interval
Trials to first correct acquisition response	.04	−.06 ↔ .13
Latency of incorrect recall retention	−.10*	−.18 ↔ −.02
Recognition	.17*	.06 ↔ .27
Latency of correct recognition	−.14*	−.24 ↔ −.04
Latency of incorrect recognition	.09	−.09 ↔ .27

* $p < .05$.

fully compute interval-scale statistics such as the mean (and standard error of the mean) of the gammas collapsed across individual subjects. For each of these correlations, the mean (and 95% confidence interval) across the individual-subject correlations is shown in Table 2.

There was no reliable relationship between the feeling of knowing and the number of trials to the first correct acquisition response (i.e., idiosyncratic item difficulty).[4] This conclusion is not modulated by the eventual degree of learning; the mean correlation for one-correct items considered separately is .10, for two-correct items is .05, and for four-correct items is −.04.

The relation between the latency of incorrect recall and the feeling of knowing was a reliable −.10, which is opposite in direction from what might be expected and is discussed in the section on search termination.

The third row of Table 2 gives a measure of the validity of the feeling of knowing as a predictor of subsequent recognition. The mean correlation was a reliable .17 ($p < .01$). Although in the expected direction, the magnitude of this mean correlation is not very high. (The median correlation between the feeling of knowing and recognition was .21, which, although slightly higher, is still not very great in magnitude.) However, further insight can be gleaned by examining this correlation separately for items having different degrees of prior learning. The mean correlation between the feeling of knowing and subsequent recognition was −.02 for one-correct items only, −.03 for two-correct items only, and .31 for four-correct items only. (The median correlation between the feeling of knowing and recognition was .00 for one-correct items only, .00 for two-correct items only, and .41 for four-correct items only.) Thus, the validity of feeling-of-knowing judgments as a predictor of subsequent recognition is modulated substantially by the degree of prior learning for the to-be-judged items. That is, subjects apparently have little or no skill at discriminating (in terms of the relative likelihood of recognition) between items that they have previously recalled only once (or twice). However, the validity of their predictions increases markedly for items that they have recalled more frequently. This finding may help to explain part of the low magnitude of the feeling-of-knowing validity during many prior studies in which the to-be-retrieved items had been acquired only to low degrees of learning. It may also help to explain why prior studies have found greater feeling-of-knowing validity for general-information items (Hart, 1965) than for laboratory paired associates (Hart, 1967); the laboratory paired associates received low degrees of learning, whereas the general-information questions may have had substantial degrees of prior learning (e.g., in previous high-school classes).

There is a small but reliable effect of the feeling of knowing as a predictor of the latency of correct recognition. As the feeling of knowing increases, the latency of correct

[4] One might consider attributing this result to too little spread between the various degrees of idiosyncratic item difficulty. However, the spread between the various degrees of idiosyncratic item difficulty was sufficient to produce differences in recall retention (i.e., the mean gamma between recall retention and the number of acquisition trials to the first correct response was a reliable −.20). Also, the effect of idiosyncratic item difficulty on recall retention increased with the degree of learning, such that the mean gamma was .02 for one-correct items considered separately, −.15 for two-correct items considered separately, and −.25 for four-correct items considered separately. (These data are discussed more fully in Leonesio and Nelson, in press.) Whether still greater spreads in idiosyncratic item difficulty would affect the feeling of knowing is, of course, an open question.

recognition decreases. This finding occurred in spite of the noise contributed both by differential reading time for various distractors and by the position of the target among the recognition alternatives (e.g., first vs. last recognition alternatives that the subject sees). Moreover, the correlation between the feeling of knowing and the latency of correct recognition cannot be attributed to some kind of general search strategy during recognition testing; if such a factor were operating, one would expect to find a similar correlation between the feeling of knowing and the latency of incorrect recognition, but the latter correlation was not reliably different from zero (sixth row of Table 2).

Hart-like Results

The first row of Table 3 shows that the subjects predicted they would recognize approximately two thirds (i.e., .67) of the items that they had not recalled; the 95% confidence interval around that mean was .61 ↔ .74, indicating that the subjects were quite similar in the overall proportion of their absolute feeling-of-knowing predictions. This proportion also increased reliably with the number of times that the item had been correctly recalled during acquisition, $\chi^2(2) = 16.41$.

The difference between the probability of correct recognition given an absolute feeling of knowing (see second row of Table 3) and the probability of correct recognition given a feeling of not knowing (see third row of Table 3) is reliable by a sign test ($p < .01$). Notice, however, that this effect is modulated by the degree of learning, as shown in the last three columns of Table 3.[5] Although the degree of learning did not reliably affect the *P*(recognition|FNK), the *P*(recognition|FK) was reliably affected.[6]

Search Termination

An ideal search-termination device would immediately terminate a search for an item not present in memory and would continue to search only for items that are present in memory (either until the sought-after item is found or until some criterion amount of search time has passed). This situation can be explored empirically by hypothesizing that the latency of incorrect recall retention should be shorter for items that subsequently are not recognized than for items that subsequently are recognized. Previous research initially confirmed this hypothesis (Laughery, Thompson, & Band, Note 1, Experiment 1), but subsequent research did not (Thompson, 1977, especially see p. 23; Laughery, Thompson, & Band, Note 1, Experiment 2). The present study also found no reliable correlation between the latency of incorrect recall retention and subsequent recognition performance (mean gamma = +.01 with the 95% confidence interval being −.10 ↔ .12).

Another possibility is that instead of search termination being determined by what a person knows (as defined by a recognition test), perhaps search termination is determined by what the person thinks he knows (as defined by the feeling of knowing). The hypothesis that the latency of incorrect recall retention is positively correlated with the magnitude of the feeling of knowing has been confirmed in previous research (Lachman, Lachman, & Thronesbery, 1979; Nel-

Table 3
Hart-like Results: Mean for Each Dependent Variable

		Number of times correct during acquisition		
Dependent variable	Overall	1	2	4
P (above threshold)	.67	.58	.67	.81
P (recognition|FK)	.71	.61	.67	.80
P (recognition|FNK)	.59	.59	.57	.59

[5] The difference between *P*(recognition|FK) and *P*(recognition|FNK) was not reliable by any statistical test for one-correct items (e.g., $t = .06$) or for two-correct items (e.g., $t = .28$) and was reliable for four-correct items only by a paired-score *t* test, $t(22) = 2.16$, $p < .05$.

[6] The omnibus effect of degree of learning on *P*(recognition|FK) was reliable by a repeated-measures analysis of variance, $F(2, 84) = 3.98$, $p < .05$, but not by a Friedman test, $\chi^2(2) = 4.5$, $p = .11$. However, one versus four correct during acquisition did reliably affect the ordinal aspects of *P*(recognition|FK); 76% of the subjects had higher *P*(recognition|FK) on four-correct items than on one-correct items (as opposed to vice versa), which is reliable by a sign test, $p < .01$.

son & Narens, 1980b; Thompson, 1977; Laughery, Thompson, & Band, Note 1) but not in the present study (Table 2). When this correlation is examined separately for items that were correct various numbers of times during acquisition, there still is no evidence of a positive correlation (mean gamma for one correct = −.13, for two correct = .03, and for four correct = −.12). Consequently, three additional tests of the search-termination/feeling-of-knowing hypothesis were conducted.

First, a sign test yielded no reliable difference between the feeling-of-knowing rank for the item with the longest incorrect-recall latency versus that for the item with the next-to-longest incorrect-recall latency; 54% of the subjects (27 out of 50) had a greater feeling-of-knowing rank for the former than for the latter. Second, a sign test yielded no reliable difference between the feeling-of-knowing rank for the item with the longest incorrect-recall latency versus that for the item with the median incorrect-recall latency; 49% of the nontie subjects (23 out of 47) had a feeling-of-knowing rank that was higher for the former than for the latter. Third, a sign test on the feeling-of-knowing rank for the item with the longest incorrect-recall latency versus that for the item with the shortest incorrect-recall latency yielded a reliable effect in the wrong direction; 30% of the subjects (15 out of 50) had a feeling-of-knowing rank that was higher for the item with the longest incorrect-recall latency than for the item with the shortest incorrect-recall latency.

In attempting to reconcile the present findings with those from previous research, we noticed that all of the aforementioned studies used general-information questions as the items, whereas the present study used laboratory-learned number–word paired associates. Other research in our laboratory also conforms to this pattern: Experiments using general-information items (e.g., Nelson et al., Note 2) yield a reliable positive relationship between the latency of incorrect recall retention and the magnitude of the feeling of knowing, whereas experiments using laboratory-learned paired-associate items do not. Why the hypothesized relationship between the latency of incorrect recall retention and the feeling of knowing should be qualitatively different for general-information items versus laboratory-learned paired-associate items is presently unknown.

Individual Differences

Table 4 shows a correlation matrix of Spearman rank-order correlations for every first-order relation between the eight major dependent variables.

Individuals with many (versus few) errors on the recall retention test did not have any reliable tendency to be slow or fast, although various latency measures were themselves reliably correlated. The only reliable correlations involving recall retention errors were with the probability of correct recognition and with the predicted probability of correct recognition. Subjects with a higher number of recall retention errors seem to have some insight into their poor retention because they

Table 4
Spearman Rank-Order Correlations for Individual Differences

	Variable							
Variable	1	2	3	4	5	6	7	8
1. Retention errors	—	−.14	−.04	−.05	−.39*	−.34*	−.21	−.21
2. Latency of incorrect retention		—	+.53*	+.07	+.19	+.05	+.17	+.49*
3. Latency of correct retention			—	+.08	+.06	+.23	+.15	+.61*
4. Feeling-of-knowing/recognition gamma				—	+.20	+.12	+.03	−.16
5. Predicted *P*(recognition)					—	+.22	+.03	+.21
6. *P*(recognition)						—	+.08	+.16
7. Latency of incorrect recognition							—	+.28*
8. Latency of correct recognition								—

Note. $48 \le N \le 50$ for all correlations.
* $p < .05$.

reliably tend to predict a lower probability of correct recognition ($\rho = -.39$); moreover, these predictions have some validity because the subjects with a higher number of recall retention errors do have a reliably lower probability of correct recognition ($\rho = -.34$).

Of major interest is that none of the correlations of the feeling-of-knowing/recognition gamma with other variables in Table 4 was reliably different from zero. This outcome is in accord with a heuristic distinction of two independent levels of memory, namely, metamemory and object-level memory (Flavell & Wellman, 1977). If metamemory and object-level memory are conceived as independent, then the aforementioned lack of relationship between the feeling-of-knowing/recognition gamma and measures of object-level memory would be expected.

Within the domain of metamemory, the Spearman rank-order correlation between the feeling-of-knowing/recognition gamma and the feeling-of-knowing/speed-of-correct-recognition gamma (not shown in Table 4) was a reliable .36. That is, subjects whose feeling-of-knowing rank order was highly correlated with subsequent recognition accuracy were also subjects whose feeling-of-knowing rank order was highly correlated with the speed of recognition across those items that were correctly recognized. Furthermore, this reliable Spearman rank-order correlation also suggests that these two kinds of gammas are, to some degree, tapping the same underlying structure. However, the magnitude of the correlation is not high, which suggests one or both of the following: (a) The population value of the Spearman rank-order correlation is actually much closer to unity (i.e., the two kinds of feeling-of-knowing gammas are tapping the same unidimensional structure) but is lower here because of noise; (b) the structure underlying the feeling of knowing is not unidimensional, and one of these feeling-of-knowing gammas taps one dimension of the underlying structure, whereas the other gamma taps another dimension of that multidimensional structure. Further tests of the latter alternative are possible by using other feeling-of-knowing procedures (e.g., examinations of intransitivities in pairwise feeling-of-knowing judgments—see Nelson & Narens, 1980a, p. 76).

Reliable Spearman rank-order correlations also occurred between the feeling-of-knowing/recognition gamma and the Hart-like measures of *P*(recognition|FK) and *P*(recognition|FNK). Subjects with a higher feeling-of-knowing/recognition gamma had a reliably higher *P*(recognition|FK), $\rho = +.43$, and a reliably lower *P*(recognition|FNK), $\rho = -.44$. The occurrence of these reliable correlations suggests that the feeling-of-knowing/recognition gamma and the two Hart-like statistics are at least partially tapping the same underlying structure (cf. previous paragraph). Also, the negative correlation between the feeling-of-knowing/recognition gamma and the *P*(recognition|FNK) is in accord with the findings of other research showing that people can make valid judgments about what they do not know (e.g., "knowing not" in Kolers & Palef, 1976).

Conclusions

The major findings of the present study are the following:

1. Overlearning affects both subsequent recall retention and the magnitude of the feeling of knowing for nonrecalled items. Thus, one factor that underlies feeling-of-knowing judgments is the degree of prior learning. Whether this overlearning effect is due to the number of overlearning study trials versus overlearning test trials, to massing versus spacing of the sequence of overlearning trials, to variable encoding during overlearning, and/or to other aspects of the overlearning situation is a topic for future research.

2. Idiosyncratic item difficulty, defined in terms of the number of acquisition trials before the first correct recall of an item, did not yield any reliable effect on subsequent feeling-of-knowing judgments.

3. Latencies of incorrect recall retention (cf. "search termination") were not positively correlated with the feeling of knowing. This discrepancy with some previous studies is speculated to be due to the use of laboratory-learned paired-associate items in the present study versus the use of general-information items in the previous studies.

4. The validity of the feeling of knowing for predicting subsequent recognition was highest within the set of items overlearned to a criterion of four correct acquisition responses per item (median correlation = .41) and lowest within the set of items learned to a criterion of one correct acquisition response per item (median correlation = .00).

5. The feeling of knowing was a valid predictor of the latency of correct recognition (i.e., items with a high feeling of knowing tend to have shorter correct-recognition latencies) but was not reliably related to the latency of incorrect recognition.

6. The validity of the feeling of knowing for predicting subsequent recognition varied across individuals and was reliably related to other metamemory measures but was not reliably related to any of the measures of object-level memory. This pattern of results is in accord with a heuristic distinction between metamemory and object-level memory.

Reference Notes

1. Laughery, K. R., Thompson, B., & Band, T. *How do we decide to terminate a memory search?* Paper presented at the annual meeting of the Psychonomic Society, Boston, November 1974.
2. Nelson, T. O., Landwehr, R. F., Leonesio, R. J., Raley, J. L., Weisman, A. M., & Narens, L. *Predictive validity of feeling-of-knowing judgments for the recognition of nonrecalled general-information facts.* Paper presented at the annual meeting of the Psychonomic Society, Philadelphia, November 1981.

References

Battig, W. F., & Spera, A. Rated association values of numbers 0–100. *Journal of Verbal Learning and Verbal Behavior*, 1962, *1*, 200–202.

Flavell, J. H., & Wellman, H. M. Metamemory. In R. V. Kail, Jr. & J. W. Hagen (Eds.), *Perspectives on the Development of Memory and Cognition.* Hillsdale, N.J.: Erlbaum, 1977.

Hart, J. T. Memory and the feeling of knowing experience. *Journal of Educational Psychology*, 1965, *56*, 208–216.

Hart, J. T. Memory and the memory-monitoring process. *Journal of Verbal Learning and Verbal Behavior*, 1967, *6*, 685–691.

Kolers, P. A., & Palef, F. R. Knowing not. *Memory & Cognition*, 1976, *4*, 553–558.

Krueger, W. C. F. The effect of overlearning on retention. *Journal of Experimental Psychology*, 1929, *12*, 71–78.

Lachman, J. L., Lachman, R., & Thronesbery, C. Metamemory through the adult life span. *Developmental Psychology*, 1979, *15*, 543–551.

Leonesio, R. J., & Nelson, T. O. Post-criterion overlearning reduces the effectiveness of the method of adjusted learning. *Behavior Research Methods & Instrumentation*, in press.

Nelson, T. O., & Narens, L. A new technique for investigating the feeling of knowing. *Acta Psychologica*, 1980, *46*, 69–80. (a)

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

Nickerson, R. S. Retrieval efficiency, knowledge assessment and age: Comments on some welcome findings. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, & L. W. Thompson (Eds.), *New Directions in Memory and Aging: Proceedings of the George Talland Memorial Conference.* Hillsdale, N.J.: Erlbaum, 1980.

Postman, L. Retention as a function of degree of overlearning. *Science*, 1962, *135*, 666–667.

Shepard, R. N. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, 1967, *6*, 156–163.

Shimamura, A. P., Landwehr, R. F., & Nelson, T. O. FACTRETRIEVAL: A program for assessing someone's recall of general-information facts, feeling-of-knowing judgments for nonrecalled facts, and recognition of nonrecalled facts. *Behavior Research Methods & Instrumentation*, 1981, *13*, 691–692.

Thompson, B. G. *The feeling of knowing: Decision to terminate the search.* Unpublished master's thesis, University of Houston, 1977.

Thorndike, E. L., & Lorge, I. *The Teacher's Wordbook of 30,000 Words.* New York: Columbia University Press, 1944.

Received June 29, 1981
Revision received November 23, 1981 ■