

Easy comes, easy goes?

The link between learning and remembering and its exploitation in metacognition

ASHER KORIAT

University of Haifa, Haifa, Israel

The cue-utilization view in metacognition assumes that judgments of learning (JOLs) are based on inferences from mnemonic cues deriving from the online processing of items during learning. This view calls for a specification of the underlying heuristics, their validity in predicting memory performance, and the extent to which they are utilized. This study examines one such heuristic: *easily learned, easily remembered* (ELER). We first show that ease of learning, as indexed by self-paced study time and by the number of trials to acquisition, is indeed a valid cue for recall success. We then demonstrate that this correlation between learning and remembering underlies metacognitive predictions about the future recallability of different items. The results are discussed in terms of the idea that metacognitive judgments incorporate knowledge about the internal ecology of cognitive processes, much as the perception of the external world embodies knowledge about the ecological structure of the environment.

A friend said to me: Italian is a very easy language. Whenever I am in Italy, it takes me 2 to 3 days to be able to express myself in Italian. However, after I leave Italy, no more than a couple of days later, I forget everything that I have learned.

This is, perhaps, an example of the adage “easy come, easy go”: Information that is acquired faster is forgotten faster. In this article, we challenge this claim, showing instead that items that are easily learned are better remembered. We then examine the hypothesis that this correlation between learning and remembering is embodied in the heuristics that people use in making metacognitive judgments during learning about the future recallability of different items.

How do people monitor their knowledge during learning and remembering? One view that has been gaining impetus in recent years is the cue-utilization view, according to which metacognitive judgments are inferential in nature: They are based on a variety of heuristics and cues that have some degree of validity in predicting objective memory performance (e.g., Benjamin & Bjork, 1996; see Koriat, 1997, 2007; Schwartz, 1994). It has been assumed that reliance on such cues and heuristics is largely automatic and unconscious, giving rise to sheer subjective feelings of knowing.

The cue-utilization view of metacognition invites an analysis similar to that proposed by Brunswik (1956) for perception. Brunswik proposed that perception is centered on distal objects in the outside world but that these cannot be perceived directly and must be inferred from proximal cues that impinge on the senses. For example, various proximal cues, such as relative size, are diagnostic of the distance of an object, and are used by the perceptual system to infer that distance. The inference occurs unconsciously, giving rise to

a sheer perceptual experience. Therefore, in order to understand the factors that contribute to veridical perception, it is necessary to examine the correlation between proximal cues and distal variables (cue validity) and the extent to which these cues are used by the perceiver (cue utilization).

We propose that, just as the perception of the environment embodies knowledge about the ecological structure of the external world, the heuristics that underlie metacognitive judgments incorporate knowledge about the principles that govern cognition. Therefore, an essential step in the analysis of metacognitive judgments is the description of the “internal ecology” of human cognition; in particular, the correlations between proximal mnemonic cues that derive from learning and remembering operations, on the one hand, and actual memory, on the other. Once that correlational structure has been delineated, we can attempt to uncover the reasons for some of these correlations, examine the extent to which metacognitive judgments take advantage of these correlations, and find out how such cue utilization contributes to the accuracy and inaccuracy of metacognitive judgments.

Indeed, it has been observed that answers that come to mind faster are more likely to be correct than answers that are associated with longer latencies (cue validity; see Koriat, Ma’ayan, & Nussinson, 2006; Robinson, Johnson, & Herndon, 1997). In turn, the speed with which an answer comes to mind is one of the mnemonic cues underlying the subjective confidence in one’s answer (cue utilization; see Kelley & Lindsay, 1993). Similarly, the amount of partial information that comes to mind about an elusive memory target, and the ease with which it comes to mind, were

A. Koriat, akoriat@research.haifa.ac.il

shown to be diagnostic of the ability to recognize that target among distractors (cue validity; see Koriat, 1993). In parallel, the accessibility of partial information about the elusive target seems to be one of the bases of the feeling of knowing that is experienced when retrieval of a memory target fails (cue utilization; Koriat, 1995).

In this study, we examine a correlation between learning and remembering that characterizes the internal ecology of cognitive processes. In the first part of the article, we focus on cue validity: We test the idea that ease of learning is diagnostic of successful remembering, such that easily encoded information stands a better chance of being recalled later than does information that requires more effort to encode. This rule, which will be called *easily learned, easily remembered* (ELER), is just the opposite of what is implied by the adage “easy come, easy go.” Ease of learning was operationalized in two different ways; first, as the amount of time spent studying an item under self-paced instructions (Experiment 1), and second, as the number of trials needed to master that item during study (Experiment 2). The following hypothesis is tested: Recall is better for items to which relatively *less* study time is allocated, and for items that require relatively *few* trials to acquisition. This hypothesis has much in common with the idea that processing fluency is diagnostic of recall success (Koriat & Ma’ayan, 2005).

The second part of the study focuses on cue utilization, examining the idea that knowledge about the ELER correlation is incorporated into the heuristics that underlie metacognitive judgments. We test the hypotheses that judgments of learning (JOLs) made during study *decrease* with the amount of time spent studying an item under self-paced conditions (Experiment 3) and with the number of trials required to master an item (Experiments 4 and 5).

Although these hypotheses may seem counterintuitive, they are consistent with the idea that JOLs are based in part on the ease with which to-be-remembered items are processed during learning (Benjamin & Bjork, 1996; Koriat, 1997). Indeed, in Matvey, Dunlosky, and Guttentag’s (2001) study, learners were asked to generate some targets during the study of paired associates. For these items, JOLs were negatively correlated with the time taken to generate targets. Similarly, in a study by Hertzog, Dunlosky, Robinson, and Kidder (2003), JOLs increased with the success and speed of forming an interactive image between the cue and the target during paired-associate learning. Also, Benjamin, Bjork, and Schwartz (1998) had participants answer general-information questions and predict the likelihood of recalling the answers at a later free-recall test. The more rapidly participants retrieved an answer, the higher was their estimate that they would recall that answer at a later time. However, it should be stressed that in these latter studies, processing fluency was not always diagnostic of future recall; in fact, in the Benjamin et al. study, recall predictions were negatively correlated with actual recall. However, these are the exceptions rather than the rule, because ease of processing is generally predictive of recall success (Kelley & Lindsay, 1993; Koriat, Ma’ayan, & Nussinson, 2006).

EXPERIMENT 1

Experiment 1 examined the ELER principle, using the amount of time invested in studying an item during self-paced learning as an index of ease of learning. It tested the hypothesis that recall success is correlated negatively with study time. Evidence consistent with this hypothesis was reported recently by Koriat, Ma’ayan, and Nussinson (2006). That study focused on the basis of JOLs, exploring the general notion that metacognitive monitoring is sometimes based on the feedback from control operations. In a series of experiments, participants studied a list of paired associates under self-paced conditions: They were instructed to spend as much time as they wished studying each item. JOLs made at the end of each trial were found to *decrease* with the amount of time invested in studying the item. This finding was taken to support the idea that JOLs are based in part on study time under the memorizing effort heuristic that easily learned items are better remembered. It was proposed that the allocation of study time to each item is data driven—that is, determined by the item itself—and that the amount of effort invested in study is used by the learner as a cue for recall predictions.

Experiment 1 of the present study focused on cue validity rather than on cue utilization: It examined the actual validity of study time as a predictor of recall. The experiment differed in an important respect from the experiments reported by Koriat, Ma’ayan, and Nussinson (2006). In these experiments, the paired associates were deliberately sampled to represent different degrees of cue–target associations, so that some items were relatively easy to learn (e.g., *good-luck*), whereas others were more difficult (e.g., *mountain-butter*). This sampling had the advantage of producing a large variation in study time, JOLs and recall. The problem, however, is that the correlation between study time, on the one hand, and JOL and recall, on the other, could have been mediated by intrinsic item difficulty: Items with a strong cue–target association are faster to learn, elicit higher JOLs, and are also more likely to be recalled than those with weak association. In Experiment 1, in contrast, as well as in all of the experiments in this study, the paired associates were formed by words sampled from a list of unrelated words and paired randomly for each participant. In this manner, there was little systematic variation between the pairs in the strength of the a priori association. Nevertheless, we expect that differences between the pairs in ease of encoding, however small and idiosyncratic they may be, will be diagnostic of subsequent recall.

Method

Participants studied a list of paired associates under self-paced conditions. In order to ensure a reasonable level of recall, the list was presented three times for study before its recall was tested.

Participants. Twenty Hebrew-speaking undergraduates (8 women, 12 men) participated in the experiment for payment.

Materials. A list of 120 Hebrew words was compiled from Hebrew word-frequency norms (Frost & Plaut, 2005). Mean word frequency was 5.57 per million (range, 5.01–6.40). An attempt was made to avoid words with obvious associative links between them.

Apparatus. The experiment was conducted on a personal computer. The stimuli were displayed on the computer screen. Re-

sponses were spoken orally by participants and then entered by the experimenter on a keyboard.

Procedure. The experiment consisted of three study presentations of a list of 60 paired associates followed by a cued recall test. For each participant, a different random pairing of the words was used to create the 60 word pairs, and that pairing was maintained across the presentations. Participants were told that they would have to study 60 paired associates, and that their memories would be tested later by having to recall the second word in each pair in response to the first word. They were told that the entire list would be presented three times for study, and that they could spend as much time studying each pair as they needed. They were to press the left key of the mouse when they felt that they had studied the pair long enough. Further, they were told that their success in performing the task depended on their ability to recall as many words as possible during the test, while keeping the total time invested in studying the entire list as short as possible. Each study trial began with a 500-msec cross sign. The two words then appeared side by side and remained on the screen until the participant pressed the mouse to signal end of study. The next trial began after the keypress.

The test phase took place only after the three study presentations. The 60 stimulus words, each preceded by a 500-msec cross sign and a beep, were presented one after the other for up to 8 sec each. Participants had to say aloud the corresponding response word within the 8 sec allotted. The items were randomly presented for each participant for each of the study presentations and for the cued recall test.

Results and Discussion

Across participants, study time per item averaged 8.6, 5.7, and 4.0 sec for presentations 1, 2, and 3, respectively. The range of these means was very large across participants: 2.3–17.8 sec, 1.6–14.4 sec, and 1.1–10.5 sec for the first, second, and third presentations, respectively.

To examine the relationship between recall and study time, we calculated for each participant the gamma correlation across items between study time invested in the last presentation and recall (scored as 0 or 1 for each item). These correlations averaged $-.24$ across participants, significantly different from zero [$t(19) = 4.64, p < .001$]. A similar analysis using mean study time across the three presentations yielded $-.15$ [$t(19) = 3.42, p < .005$]. The former correlation was negative for 17 of the 20 participants, whereas the latter was negative for 15 of 19 participants (for one participant, it was zero). Both of these correlations were significant at the .01 level, according to a binomial test. Thus, the less effort that was invested in studying an item, the *more* likely it was to be recalled. These results suggest that the ELER rule holds true even when there are no systematic differences between the items in intrinsic item difficulty (see Experiment 6).

EXPERIMENT 2

In Experiment 2, we used a number of trials to acquisition as a measure of ease of learning. Participants studied a list of paired associates for several study–test cycles. A dropout procedure was used so that each item recalled on a particular block was deleted from the study–test phases of the subsequent blocks. A final recall took place 7 min after the last recall block. The hypothesis was that final recall should correlate negatively with number of trials to acquisition: The more often a pair is studied, the *less* likely it is to be recalled in the final testing.

Method

Participants. Thirty Hebrew-speaking University of Haifa undergraduates (20 women, 10 men) participated in the experiment for payment.

Materials and Apparatus. The apparatus was the same as in Experiment 1. The same list of 120 words as in Experiment 1 was used, except that 17 words were replaced to reduce further possible associations between words. As in Experiment 1, the words were paired randomly for each participant, to produce 60 word pairs.

Procedure. The procedure for the first presentation was similar to that used for the first presentation in Experiment 1, except that presentation rate was fixed rather than self-paced. Each study trial began with a 500-msec cross sign, followed by a 2.5 sec presentation of the pair, with a 1-sec interval between trials. Participants were instructed to study the pairs for a subsequent cued recall test. The procedure for the cued recall test, which took place immediately after the first presentation, was the same as that used in Experiment 1.

For the following presentations, only the pairs that the participant had failed to recall on the preceding test phase were presented. When the presentation was over, only those pairs were tested for cued recall. Participants were informed about the dropout procedure after the first recall test. The study–test cycles were terminated when participants achieved perfect recall. Following a filler task, involving mathematical and figural problems, that lasted 7 min, a final cued-recall test took place, in which memory for all the 60 pairs was tested. The order of presentation was random for each participant for each study and test phase.

Results and Discussion

Figure 1 presents mean final recall as a function of number of trials to acquisition. It can be seen that final recall was a negative function of number of trials to acquisition. Because of the uneven distribution of number of trials to acquisition, we compared mean final recall for items requiring one or two presentations (79.0%) with mean recall for items requiring three or more presentations (70.3%). The results yielded $t(29) = 3.11, p < .005$.

We also calculated a gamma correlation for each participant across the 60 items between number of trials to acquisition and final recall. The correlations ranged between $+.22$ and $-.70$, and averaged $-.29$ across participants, significantly different from zero [$t(29) = 6.45, p < .0001$]. The correlation was negative for 27 out of the 30 participants ($p < .0001$, by a binomial test). Thus, final recall decreased as the number of trials to acquisition increased.

An interesting feature of the results presented in Figure 1 should be noted. Final recall might have been expected to be worse for items dropped earlier from the study because they had a longer retention interval than items dropped later. In fact, however, the opposite was the case. Thus, we calculated for each participant and for each item the number of study trials lapsed between the last block in which that item was presented and the end of the study phase. The number of lapsed study trials averaged 51.1 across participants. We then divided the items at the median of lapsed study trials for each participant. Across participants, final recall for below-median and above-median items averaged 64.4% and 80.8%, respectively [$t(29) = 7.01, p < .0001$]. Thus, paradoxically, items with longer lapses exhibited better recall than those with shorter lapses. Hence, perhaps the relationship between number of study trials and final recall should be stronger than that depicted in Figure 1 under conditions in which

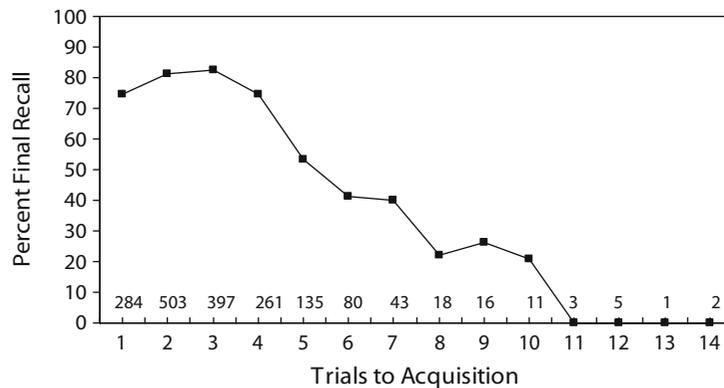


Figure 1. Percent final recall as a function of number of trials to acquisition. The bottom line indicates the number of observations on which each mean is based (Experiment 2).

the confounding is removed (or reduced) between number of trials to acquisition and number of trials until final testing (e.g., by assessing final recall after a longer interval).

EXPERIMENT 3

Experiments 1 and 2 focused on cue validity, supporting the reality of the ELER principle. Experiment 3 extended investigation to cue utilization, examining the hypothesis that the belief about the ELER principle is incorporated into the heuristics underlying learners' monitoring of their degree of learning. Participants studied a list of paired associates under self-paced instructions and made JOLs following the study of each pair.

In addition, Experiment 3 examined the changes that occur in both cue validity and cue utilization with repeated study of the items. Previous results suggested that with repeated practice, learners rely more heavily on internal mnemonic cues in monitoring their degree of mastery of each item, and furthermore, that the validity of these mnemonic cues in predicting recall improves with repeated study-test cycles (Koriat, 1997; Koriat, Ma'ayan, Sheffer, & Bjork, 2006). To examine the changes that occur with practice in both cue utilization and cue validity, the list of paired associates was presented for 4 study-test cycles. The pairs were formed by randomly pairing different words, as was done in Experiment 1.

Method

Participants. Twenty Hebrew-speaking University of Haifa undergraduates (10 women and 10 men) participated in the experiment for payment.

Materials, Apparatus, and Procedure. The apparatus and materials were the same as in Experiment 2. The words were paired randomly for each participant, and the same pairing was maintained throughout the 4 presentations of the list. The experiment included 4 study-test cycles. The study phases were self paced, as in Experiment 1, but after the left keypress signaling end of study, the question "Chances to recall (0%–100%)?" appeared. Participants reported JOLs orally on a 0%–100% scale, expressing the likelihood of recalling the target word in response to the cue. The next pair was presented as soon as the experimenter recorded the JOL on the keyboard. In addition, at the end of each study phase, participants were asked to make an aggregate estimate. They were told: "You were presented with 60 word pairs. How many of them do you think you will remember?"

The procedure for the test phases that followed each study phase was the same as in the previous experiments. The order of presentation of the pairs was randomly determined for each participant for each of the eight study and test phases of the experiment.

Results and Discussion

Mean study time per item dropped steadily with practice, averaging 11.3 sec, 8.0 sec, 4.4 sec, and 3.1 sec, for the first through fourth presentations, respectively. JOLs for those presentations averaged 50.7%, 47.8%, 63.8%, and 74.0%, respectively, whereas recall averaged 28.0%, 59.0%, 72.2%, and 80.3%, respectively. These results replicate the underconfidence-with-practice (UWP) effect reported by Koriat, Sheffer, and Ma'ayan (2002; see also Koriat, Ma'ayan, Sheffer, & Bjork, 2006; Serra & Dunlosky, 2005): JOLs overestimated recall in the first presentation, but underestimated recall from the second presentation on. A measure (recall vs. JOL) \times presentation ANOVA yielded $F(3,57) = 38.70$, $MS_e = 64.65$, $p < .0001$, for the interaction. A similar UWP pattern was observed for aggregate JOLs, which averaged (expressed as a percentage of the total number of items) 39.3%, 34.8%, 51.1%, and 65.2%, for the first through fourth presentations, respectively. A measure (recall vs. aggregate JOLs) \times presentation ANOVA also yielded a significant interaction [$F(3,57) = 29.37$, $MS_e = 89.38$, $p < .0001$].

We turn next to the examination of cue validity, as indexed by the gamma correlation between study time and recall (Figure 2A). The correlations were negative and significant for the second, third, and fourth presentations [$t(19) = 3.95$, $p < .001$; $t(19) = 9.32$, $p < .0001$; and $t(19) = 10.26$, $p < .0001$, respectively]. The correlation increased steadily with practice, so that by the fourth presentation it was $-.58$. A one-way ANOVA on these correlations yielded $F(3,57) = 37.72$, $MS_e = 0.036$, $p < .0001$. Thus, recall success is inversely related to study time, and with repeated practice, study time becomes an increasingly valid predictor of recall performance.

A similar pattern was observed for cue utilization as indexed by the gamma correlation between study time and JOL (Figure 2B). All correlations were negative and significant at the .01 level. In addition, the correlations

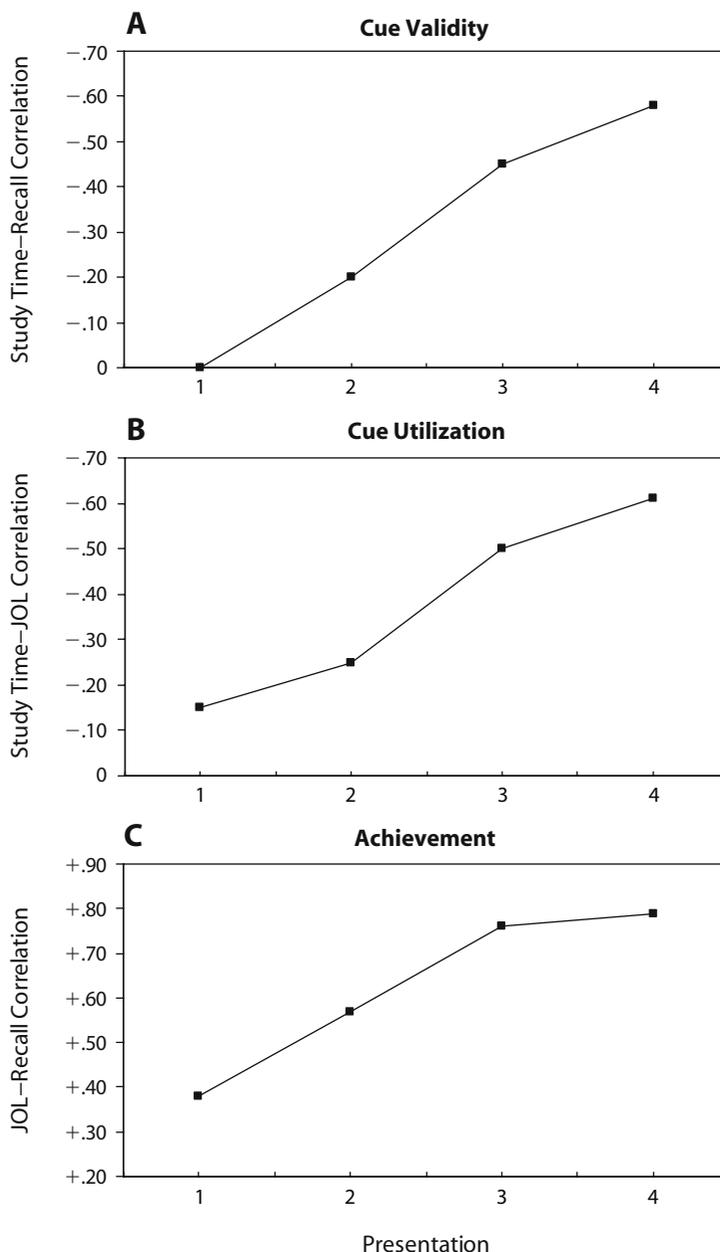


Figure 2. (A) Study time–recall correlation. (B) Study time–JOL correlation. (C) JOL–recall correlation, as a function of presentation (Experiment 3).

increased monotonically with presentation [$F(3,57) = 32.38$, $MS_e = 0.028$, $p < .0001$]. These results support the contention that JOLs incorporate knowledge about the predictive validity of memorizing effort: The more time one spends studying an item, the *less* confident one is that the item will be recalled at test. Furthermore, reliance on memorizing effort as a cue for degree of mastery increases with repeated presentations of the same list.

It is impressive that the changes in cue utilization (Figure 2B) mirror closely the changes in cue validity (Figure 2A). Possibly both of these changes contribute to the improvement in JOL accuracy with practice (Figure 2C): The JOL–recall gamma correlation increased monotonically

with study–test practice. This increase has been reported previously (e.g., King, Zechmeister, & Shaughnessy, 1980; Koriat, 1997; Koriat et al., 2002; Mazzoni, Cornoldi, & Marchitelli, 1990), but the results presented in Figure 2A and Figure 2B may offer an explanation of that systematic increase. A one-way ANOVA of the results presented in panel C yielded $F(3,57) = 31.37$, $MS_e = 0.026$, $p < .0001$.

Altogether, the results of Experiment 3 are consistent with the proposition (Koriat, Ma'ayan, & Nussinson, 2006) that unlike the common assumption in metacognition that monitoring drives control (e.g., Nelson & Narens, 1990), monitoring is sometimes actually based on the feedback from control operations. Thus, JOLs seem to

be based on study time or study effort. Assuming that the relative amount of time spent studying an item is largely data driven, dictated by the item itself in a bottom-up fashion, differences in self-paced study time provide a relatively valid cue for future recall, under the heuristic that easily learned items are better remembered.

EXPERIMENT 4

Experiment 4 also extended investigation to cue utilization but, like Experiment 2, it focused on number of trials to acquisition. The procedure was similar to that of Experiment 2, except that when the study phase was over participants were presented with each cue–target pair and asked to judge the likelihood that they will recall the target in response to the cue in a subsequent cued recall test. A final cued recall test was administered at the end, although performance on that test is possibly contaminated by the presentation of the entire list for JOL solicitation.

Method

Participants. Twenty Hebrew-speaking University of Haifa undergraduates (14 women, 6 men) participated in the experiment, 14 for course credit and 6 for payment.

Materials, Apparatus, and Procedure. These were the same as in Experiment 2, except that after achieving perfect recall and before the final cued-recall test, a JOL phase was added: Participants were presented again with the entire list of cue–target pairs in a random order, and were asked to indicate their JOLs for each pair reflecting the likelihood that they would recall the target to the cue in a subsequent cued-recall test. Each pair appeared for 2.5 sec and was replaced by the JOL probe, as in Experiment 3. Participants reported JOLs orally, and the next pair was presented as soon as the experimenter recorded the JOL on the keyboard.

Results and Discussion

Figure 3 presents mean JOL and mean final recall as a function of number of trials to acquisition. Presented also is the number of items on which each mean is based.

We focus first on the results for the final recall test. Although performance on this test may have been contami-

nated by the presentation of the cue–target pairs for JOL solicitation, the results replicated the negative function that was observed in Experiment 2. As in that experiment, mean final recall was higher for items requiring one or two presentations (93.1%) than for items requiring three or more presentations (83.3%) [$t(19) = 3.20, p < .005$]. The within-person gamma correlation between number of trials to acquisition and final recall averaged $-.33$ across participants, significantly different from zero [$t(19) = 3.36, p < .005$]. The correlation was negative for 15 out of the 20 participants, $p < .05$, by a binomial test. Thus, final recall decreased as the number of trials to acquisition increased.

Turning next to JOLs, it can be seen that participants' JOLs also decreased systematically with number of trials to acquisition. JOLs were higher for items requiring one or two presentations (65.8%) than for those requiring three or more presentations (56.4%) [$t(19) = 4.53, p < .001$]. The within-person gamma correlation across the 60 items between JOL and number of trials to acquisition averaged $-.25$, significantly different from zero [$t(19) = 4.90, p < .0001$]. The correlation was negative for 17 out of the 20 participants ($p < .005$, by a binomial test). Thus, JOLs also decreased as the number of trials to acquisition increased.

As can be seen in Figure 3, participants exhibited an underconfidence bias. They were significantly underconfident for items requiring one or two presentations [$t(19) = 11.38, p < .0001$], as well as for those requiring three or more presentations [$t(19) = 7.41, p < .0001$]. As far as resolution is concerned, however, mean gamma correlation between JOLs and final recall averaged $.41$ significantly different from zero [$t(19) = 5.16, p < .0001$]. This correlation was positive for 17 out of the 20 participants ($p < .005$, by a binomial test).

The results also illustrate the paradoxical pattern noted earlier, that final memory was relatively better for items for which there was a relatively long interval between their last presentation and their final recall testing. What is interesting is that this pattern was also found for JOLs. When items were divided according to the number of

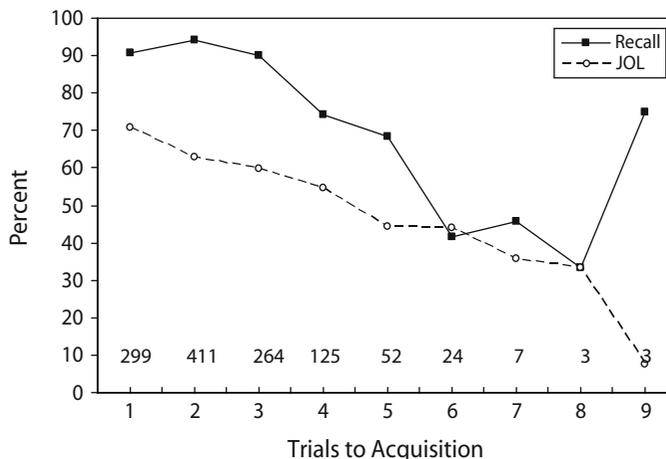


Figure 3. Percent final recall and JOL as a function of number of trials to acquisition. The bottom line indicates the number of observations on which each mean is based (Experiment 4).

lapsed study trials (between the last block in which the item was presented and the end of the study phase), final recall for below-median and above-median items averaged 81.9% and 93.1%, respectively [$t(19) = 3.60, p < .005$]. The respective JOL means were 55.2% and 66.1% [$t(19) = 5.11, p < .0001$]. Thus, paradoxically, items with longer lapses exhibited better recall and higher JOLs.

EXPERIMENT 5

The results of Experiment 4 are consistent with the idea that JOLs incorporate the belief that easily learned items are more likely to be remembered than those requiring more effort to learn. It is still possible, however, that the parallel between the results observed for JOLs and recall derives from independent effects of intrinsic item difficulty on both. The pairs used for each participant possibly differed in a priori judgments of item difficulty, and these differences may have served as the only basis of interitem differences in JOLs. To the extent that judgments of item difficulty have some validity in predicting recall (Koriat, 1997), this would explain the pattern observed in Figure 3. Evidence pertinent to this possibility was explored in Experiments 5 and 6.

Experiment 5 introduced a manipulation that was expected to affect ease of learning: Participants were exposed to half of the paired associates before they were required to study the entire list. If JOLs are sensitive to ease of learning regardless of source, an easy-to-learn item will elicit relatively higher JOLs, whether the item's ease derives from semantic aspects (semantic-associative relatedness that may affect its a priori difficulty) or from episodic aspects (preexposure).

Method

Participants. Twenty Hebrew-speaking University of Haifa undergraduates (12 women, 8 men) participated in the experiment, 12 for course credit and 8 for payment.

Materials, Apparatus, and Procedure. These were the same as in Experiment 4, except that a priming block preceded the first

presentation of the entire list. In the priming block, a randomly selected set of 30 paired associates was presented. Each pair appeared for 2.5 sec, with a 1-sec interval between pairs. Participants were asked to study the pairs, but actually no recall test followed. Instead, participants were told that before the cued recall test, they would be presented with another study list of pairs, some of which had already appeared in the previous list and some that were new. The assignment of pairs to the primed and unprimed sets was counterbalanced across each group of two participants.

Results and Discussion

As expected, recall for the first presentation was significantly higher for the primed items (43.4%) than for the unprimed items (22.0%) [$t(19) = 6.49, p < .0001$]. Also, mean number of trials to acquisition was significantly lower for the primed items (2.55) than for the unprimed items (3.01) [$t(19) = 7.73, p < .0001$]. The respective difference in mean final recall was not significant: 82.65% and 83.79%, respectively [$t(19) = 0.60$].

Figure 4 presents the means of final recall and JOL as a function of number of trials to acquisition, plotted separately for the primed and unprimed items. Presented also is the number of items on which each mean is based. It can be seen that for both recall and JOLs the functions for the primed and unprimed items are almost indistinguishable. The functions clearly indicate that JOLs and recall generally decrease with trials to acquisition.

To evaluate these results, we compared mean final recall for items requiring one or two presentations (few) with that for items requiring three or more presentations (many). A two-way ANOVA, trials to acquisition (few vs. many) \times priming (primed vs. unprimed) on final recall, yielded $F(1,19) = 19.19, MS_e = 184.17, p < .001$, for trials to acquisition, and $F < 1$ for both priming and the interaction. A similar ANOVA for JOLs yielded $F(1,19) = 23.22, MS_e = 213.86, p < .0001$, for trials to acquisition, and nonsignificant effects for both priming and the interaction [$F(1,19) = 3.04, MS_e = 91.82, p < .10$, and $F(1,19) = 2.01, MS_e = 48.64, p < .18$, respectively].

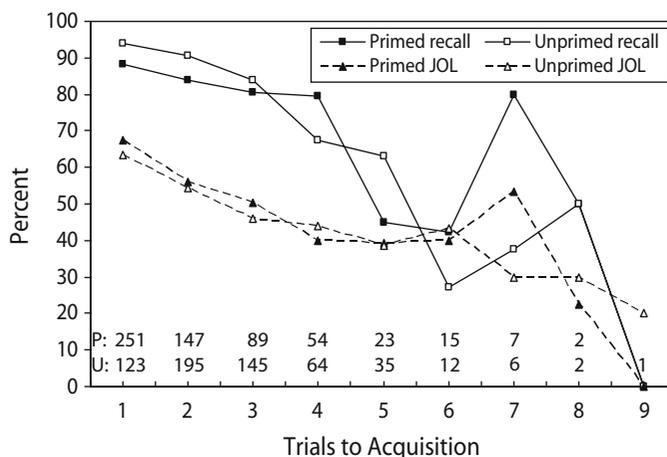


Figure 4. Percent final recall and JOL as a function of number of trials to acquisition. The results are presented separately for primed and unprimed items. The bottom lines indicate the number of observations on which the primed (P) and unprimed (U) means are based (Experiment 4).

In sum, Experiment 5 replicated the observations that both JOLs and final recall decrease as the number of trials to acquisition increase. In addition, the results show that JOLs are not merely a function of judged a priori item difficulty that can be gauged from semantic–associative characteristics of the items; rather, JOLs seem to be affected by differences in ease of learning, regardless of the source of such differences.

Individual Differences

Before presenting the final experiment, we would like to examine the results from the perspective of individual differences. All of the experiments reported so far established that items studied more easily are better recalled than those requiring more time or more trials to study. It is of interest to inquire whether or not this relationship holds true for differences between learners. Do individuals who study faster also recall better than individuals who study more slowly? Consider the following quotation from Swift (1918):

The opinion generally prevails that those who learn quickly forget easily, but experiments do not sustain this view. Practically all of the investigators have found that rapid workers remember more of what they learn than those who are slow. . . . Norsworthy tested eighty-three students, and after a month had passed the more rapid workers remembered more of what they had committed to memory than those who had learned the assignment with more effort. . . . The discovery that the statement “easy comes easy goes” is not true has far-reaching educational significance.

We tested Swift’s hypothesis on the data of Experiments 3 and 4. To anticipate: It appears that the results depend on whether the difference in the effort invested by different learners reflects differences in motivation or differences in ability.

In Experiment 3, we examined how individual differences in study time relate to recall and JOL. We calculated

for each participant the mean study time he or she allocated to each item across its four presentations and then divided participants at the median of that mean (6.0 sec/item) into fast learners (mean 4.5 sec/item) and slow learners (mean 8.8 sec/item). Figure 5 presents mean recall and JOL in the fourth presentation for slow and fast learners. It can be seen that recall was actually *higher* for slow learners than for fast learners [$t(18) = 3.22, p < .005$]: Participants who allocated more study time showed better recall. The results for JOLs were in the same direction but short of significance [$t(18) = 1.71, p < .11$].

Figure 5 also presents mean recall and JOL for items that received below-median study times (short) and for those that received above-median study times (long) for each participant. Focusing first on recall: It can be seen that for both slow and fast learners, recall was better for items that received relatively less study time. A two-way ANOVA, learners (fast vs. slow) \times items (short vs. long) yielded significant effects for items [$F(1,18) = 69.28, MS_e = 69.82, p < .0001$], and for learners [$F(1,18) = 4.72, MS_e = 321.15, p < .05$], but not for the interaction [$F(1,18) = 2.39, MS_e = 69.82, p < .15$]. Thus, the effects of between-participants variation were in the opposite direction to those of within-participants variation.

The results for JOLs paralleled those for recall, although the difference between slow and fast learners was short of significance. A two-way ANOVA similar to the one above yielded $F(1,18) = 95.41, MS_e = 50.70, p < .0001$, for items; $F(1,18) = 1.96, MS_e = 453.59, p < .19$, for learners; and $F < 1$ for the interaction. For both slow and fast learners, JOLs were higher for items that received less study time. It is noteworthy that the results for JOLs mirrored those for recall, except for the fact that JOLs were lower overall than recall, possibly because of the UWP effect. These results underscore the idea that JOLs incorporate the assumption that recall is better for easily learned items than for items that are studied with greater effort.

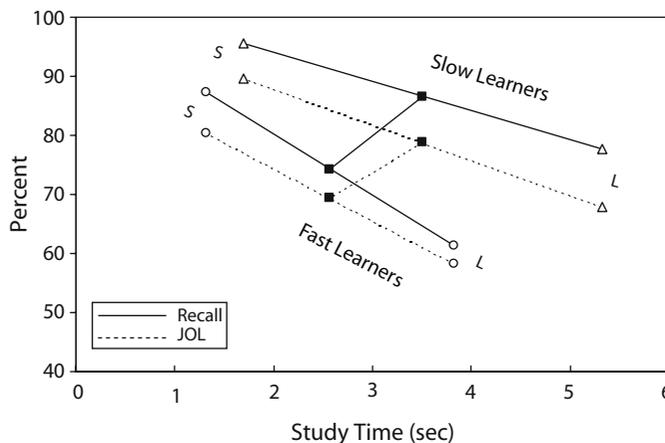


Figure 5. Percent recall and JOL as a function of mean self-paced study time per item. The results are presented separately for slow learners and fast learners and for each group they are plotted separately for items that received below median study times (S, short) and items that received above median study times (L, long) (Experiment 3).

A different pattern, one more consistent with Swift's (1918) hypothesis, emerged in the analysis of the results of Experiment 4, in which ease of learning was indexed by number of trials to acquisition. Participants were divided at the median of mean number of trials to acquisition (2.31) into fast learners and slow learners. In addition, for each participant, the items were divided at the median in terms of number of trials to acquisition. The means of final recall and JOL are presented in Figure 6 for slow and fast learners. It can be seen that recall was higher for fast learners than for slow learners [$t(18) = 2.98, p < .01$]. The results for JOLs were in the same direction—higher JOLs for fast learners (66.9%) than for slow learners (57.5%); but the difference was only near significant [$t(18) = 2.03, p < .07$].

Figure 6 also presents mean recall and JOL for items that required below-median number of trials to acquisition (few) and for those that required above-median number of trials to acquisition (many) for each participant. Recall performance was better for fast learners and for items that required relatively fewer trials to acquisition. Thus, a two-way ANOVA for items (few vs. many) \times learners (fast vs. slow) yielded significant effects for items [$F(1,18) = 15.35, MS_e = 35.60, p < .001$] and for learners [$F(1,18) = 8.62, MS_e = 42.11, p < .01$], but not for the interaction [$F(1,18) = 2.21, MS_e = 35.60, p < .16$].

A similar two-way ANOVA on JOLs yielded $F(1,18) = 19.53, MS_e = 66.44, p < .001$, for items; $F(1,18) = 2.91, MS_e = 202.26, p < .12$, for learners; and $F < 1$ for the interaction. Although the effect of learners was only nearly significant, the results for JOLs generally mirrored those for recall except for the fact that JOLs were overall lower than recall, possibly because of the UWP effect.

In sum, when ease of learning was defined in terms of average amount of study time per item, memory performance was better for the slow learners than for fast learners. In contrast, when it was defined in terms of trials to acquisition,

the opposite pattern was observed: Recall was better for fast learners. It should be noted that this differential pattern was also observed for the results of Experiments 1 and 2. When participants in Experiment 1 were divided at the median of their mean study time per item across the three presentations, recall performance was better for those who invested relatively more study time (72.3%) than for those who invested less study time (62.0%), although the difference was not significant [$t(18) = 1.22$]. In Experiment 2, in contrast, recall was better for fast learners (81.0%), who required an average of 2.5 study trials per item, than for slow learners (67.8%), who needed an average of 3.8 trials per item [$t(28) = 4.34, p < .001$].

It would seem that average self-paced study time reflects differences in motivation rather than differences in ability. Possibly, learners who are better motivated to excel spend more time studying each item and consequently exhibit better recall performance. In contrast, mean number of trials to acquisition reflects differences in learning ability, and these chronic differences appear to be correlated with memory ability. Thus, when ease of learning was defined in terms of self-paced study, the effects on recall were in opposite directions for between-persons variation than for within-persons variation. When they were defined in terms of trials to acquisition, the two effects were in the same direction. What is notable, however, is that regardless of how individual differences in ease of learning were defined, items that were learned more easily exhibited better recall, and this was true for both fast learners and slow learners. It is also noteworthy that the pattern of results observed for JOLs (in Experiments 3 and 4) mirrored closely the pattern observed for recall.

EXPERIMENT 6

Experiment 6 had two aims. The first was to examine whether a belief in the ELER rule would be evident when judges were asked to rate items on both ease of learning

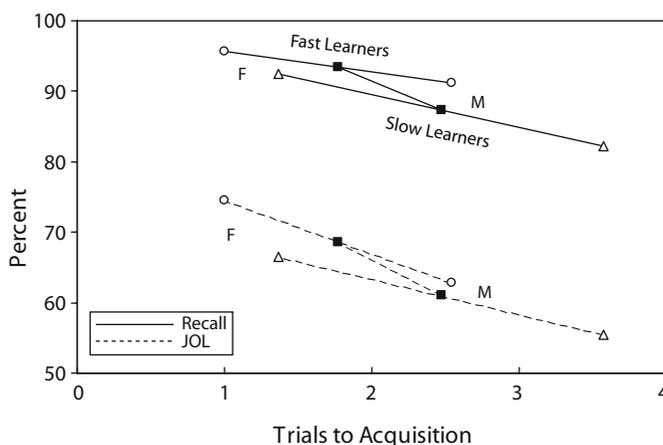


Figure 6. Percent final recall and JOL as a function of mean number of trials to acquisition. The results are presented separately for slow learners (requiring many study trials), and for fast learners (requiring few study trials), and for each group they are plotted separately for items that required below median number of trials (F, few) and those that required above median number of trials (M, many) (Experiment 4).

and ease of remembering. The second aim was to examine whether the correlations observed between trials to acquisition and recall mediated by commonly shared properties of the items, or whether they reflected idiosyncratic differences between items that affect ease of learning and remembering.

In this experiment, 10 participants (to be referred to as judges) were presented with lists of paired associates and asked to make ease-of-learning (EOL) and ease-of-recalling (EOR) judgments about each of the items in the list. Each judge was yoked to one of the first 10 participants in Experiment 2, receiving exactly the same list (i.e., the same word pairs in the same order of presentation) as that used for the yoked participant in the first presentation in Experiment 2. The instructions for the EOL judgments required the judge to estimate how many repetitions of the list he or she would need in order to master each specific item. The instructions for the EOR judgments required assessing the likelihood of recalling each target in response to its cue in a cued recall test. In this way, the EOL and EOR judgments made by each judge to each item could be compared, respectively, with the number of trials to acquisition and recall success for that item for the yoked participant. We examined whether EOL and EOR judgments predicted the actual study and test performance of the yoked participants in Experiment 2.

Method

Participants. Ten Hebrew-speaking University of Haifa undergraduates acted as judges; eight were paid for their participation and two received course credit. Each of them was yoked to one of the first 10 participants in Experiment 2.

Materials, Apparatus, and Procedure. Each judge received exactly the same list as that used for his or her yoked participant in Experiment 2, and in the same order as in the first presentation. The list was printed twice, once for EOL judgments and once for EOR judgments. The printed instructions for the EOL judgments indicated that new material often has to be rehearsed several times to be learned. Then, a description of the procedure of Experiment 2 was given (60 word pairs, a 3-sec rate presentation, cued recall at test), except that, instead of the dropout procedure that was used in Experiment 2, it was stated that the entire list was presented for 10 study-test cycles. For each of the pairs, the judge was asked to estimate how many presentations he or she would need in order to learn each pair. It was pointed out that different pairs might differ in their ease of learning and that the task of the judge was not to learn the pairs, but to try to estimate their relative ease of learning. Before making their estimates, judges were advised to inspect the first 5–10 pairs in order to get an impression of the relative ease of learning of the pairs. The list of pairs then followed, with the numbers 1 to 10 next to each pair. The judges indicated their estimate by circling one of these numbers.

There was a 5-min filler task, after which the EOR judgment task was administered. The instructions for this task began with a description of a hypothetical experiment in which the list was presented for one study-test cycle. For each word pair, judges were asked to assess the probability, on a 0%–100% scale, that in response to the cue they would recall the correct target word in the test phase. They were asked to write a number between 0 and 100 on the blank line next to each pair. The EOL and EOR forms were administered in that order for all participants.

Results and Discussion

For each pair of yoked participants, we calculated the correlations across the 60 paired associates between four variables, the EOL and EOR judgments of the judge, and the number of trials to acquisition and recall performance

(scored 0 or 1) for the Experiment 2 participant. The correlations are listed in Table 1.

First consider the EOL–EOR correlations. These correlations averaged $-.59$, significantly different from zero [$t(9) = 6.36, p < .0001$]. The correlation was negative for 9 of the judges ($p < .05$), by a binomial test. Thus, the EOL–EOR correlation discloses the ELER belief that items requiring fewer trials to master are more likely to be recalled than are those requiring more trials. It can be seen that a similar correlation exists for the 10 yoked participants from Experiment 2: The correlations between trials to acquisition and recall were negative for all participants ($p < .0001$, by a binomial test), and averaged $-.25$ [$t(9) = 8.25, p < .0001$].

In contrast to these within-participants correlations, the cross-participants correlations were not systematic. The correlation between EOL and trials to acquisition, which averaged $.03$, indicates that the judges failed to estimate the relative number of repetitions required to acquisition. Similarly, they failed to predict the relative recall of different items: The EOR–recall correlation averaged $-.03$.

These results suggest that the procedure used for constructing the lists of paired associates succeeded in removing systematic differences between the pairs (e.g., in actual and judged difficulty) that could give rise to cross-participants consistency. In fact, in previous studies, in which the paired associates varied in degree of cue–target association, judges' ratings of the difficulty of each pair were moderately predictive of recall success. For example, Nelson and Leonesio (1988) reported a mean within-persons correlation of $.48$ between ease-of-learning judgments and recall, and Koriat (1997; Experiment 2) reported a correlation of $.81$. Our results, then, suggest that the within-persons correlations observed between EOL judgments and EOR judgments in Experiment 4, and between trials to acquisition and recall in Experiment 2, are not due to commonly shared difficulty characteristics of items. Rather, they would seem to reflect the effects of item-specific cues that are idiosyncratic in nature. Indeed, Koriat (1997) also obtained evidence suggesting that the mnemonic cues that serve as the basis of JOLs are idiosyncratic, reflecting privileged access to personal aspects of encoding and remembering.

Table 1
Pearson Correlations Between EOL and EOR Judgments, TTA and Recall, TTA and EOL, and EOR and Recall

Yoked Participants	EOL–EOR	TTA–Recall	TTA–EOL	EOR–Recall
1	-.73	-.19	.02	.12
2	-.78	-.14	.25	-.10
3	.14	-.42	-.02	-.05
4	-.48	-.38	-.32	-.20
5	-.59	-.29	-.19	.06
6	-.88	-.18	-.11	-.10
7	-.50	-.32	.01	-.01
8	-.57	-.26	.25	.03
9	-.87	-.20	.24	.04
10	-.68	-.16	.15	-.11
Mean	-.59	-.25	.03	-.03

Note—EOL, ease of learning; EOR, ease of recalling; TTA, trials to acquisition.

GENERAL DISCUSSION

This study was predicated on the assumption that, in the same way that veridical perception benefits from regularities in the link between proximal and distal cues, metacognitive monitoring capitalizes on correlations in the “internal ecology” of cognitive processes between mnemonic cues and actual memory. We focused on one such correlation—that easily learned items are better remembered. In contrast to the adage *easy come, easy go*, this heuristic, labeled the ELER heuristic, assumes that ease of learning is positively correlated with future memory. The results provide evidence for the reality of this correlation (cue validity), and suggest that this correlation is indeed relied upon in monitoring one’s own knowledge (cue utilization).

Let us first examine cue validity. We found that ease of learning, as indexed both by self-paced study time and by number of trials to acquisition, was correlated across items with cued recall. Why are easily learned items more likely to be recalled than items that are more difficult to learn? In the real world there are many factors, such as simplicity, familiarity, and predictability, that may contribute both to the relative learnability of different items and to their relative recallability. These factors may mediate the link between ease of learning and ease of remembering. For example, in a study by Bahrick and Phelps (1987), participants studied the Spanish translations of English words using a dropout procedure similar to that used in Experiment 2. A cued recall test that took place eight years later indicated that recall decreased with number of trials to acquisition. In that study, the ELER correlation could be mediated by interitem differences in such factors as familiarity, and similarity between the Spanish and English terms, which could contribute to both ease of learning and successful retention. In the procedure used in this study, in contrast, an attempt was made to minimize systematic differences between the pairs in normative difficulty; and indeed, the results of Experiment 6 indicated that judges could not estimate the relative ease of learning and remembering the items that were used. Nevertheless, the results yielded support for the ELER rule. Of course, the items probably differed in idiosyncratic features that possibly affected both learning and remembering, and these differences may have contributed to the ELER correlation observed for each participant, but there is still the possibility of an inherent correlation between ease of learning and effective retention. The question then remains: Why do easily learned items stand a better chance of being recalled at a later time than do items requiring more trials or more time to learn? This correlation must be explained by principles of learning and memory. From the point of view of metamemory, however, it is a blessing, because it can help learners monitor the future recallability of studied items.

Let us turn to cue utilization. Whatever reason there is for the ELER rule, the results suggest that during learning learners do utilize that rule to monitor their degree of mastery of different items. Thus, in Experiment 3, JOLs were higher for items that received less study time than they were for those that received more study time. Furthermore, the results indicated that reliance on the ELER heuristic increased gradually with repeated study–test cycles, consistent with

the idea that with increased practice studying the same list of items, participants rely more heavily on mnemonic cues in making JOLs (see Koriat, 1997; Koriat & Bjork, 2006; Kori- at, Ma’ayan, & Nussinson, 2006). In Experiments 4 and 5, JOLs were found to decrease systematically with number of trials to acquisition, and the results for JOLs mirrored closely those obtained for recall.

Altogether, these results suggest that the heuristics underlying the monitoring of one’s own knowledge during learning incorporate knowledge about the empirical link between learning and remembering. In terms of Schwarz’s (2004) conceptualization, in making JOLs, learners apply a naive theory consistent with the ELER correlation. What is the process underlying the observation that JOLs decrease with increased ease of learning? Two possibilities exist that are not mutually exclusive. First, the correlation between JOLs and ease of learning may be mediated by specific features of items that affect both ease of learning and JOLs. Second, ease of learning affects JOLs directly: Easily learned items produce a stronger feeling of mastery than do items that require more effort to learn (Koriat, Ma’ayan, & Nussinson, 2006). With regard to the first possibility, the results of Experiment 6 suggest that, even if the effects of JOLs are mediated by item-specific properties, these properties are idiosyncratic rather than common to all participants. Kori- at (1997) also reported evidence that the mnemonic cues underlying JOLs, particularly those responsible for the improved accuracy of JOLs with practice (see Figure 2C), are idiosyncratic in nature. The results of Experiment 5, in turn, lean more heavily toward the second possibility. They suggest that ease of learning affects JOLs, whether the source of ease of learning derives from episodic effects (preexposure) or from item-specific properties (semantic–associative characteristics). Indeed, in the study mentioned earlier, Hertzog et al. (2003) found that when concrete and abstract word pairs were mixed in a list, the generation of an interactive image between the cue and the target was faster for concrete pairs than for abstract pairs. However, concreteness and generation latency influenced JOLs independently of each other, suggesting that processing fluency makes a unique contribution to JOLs. This suggestion is consistent with previous research emphasizing the contribution of encoding and retrieval fluency to subjective experience, in general, and to metacognitive judgments, in particular (Kelley & Rhodes, 2002; Kori- at & Ma’ayan, 2005; Schwarz, 2004). Both self-paced study time and trials to acquisition can be said to be indicative of the relative encoding fluency of the item, and perhaps encoding fluency affects JOLs directly.

The correspondence between cue utilization and cue validity has been emphasized in the analysis of percep- tion, and it is interesting that it holds true for metacogni- tion as well. Not only did the effects of ease of learning on cue utilization parallel those observed for cue validity, but the changes in the extent of cue utilization across re- peated presentations were found to closely mirror those that occur in the magnitude of cue validity. An analogous pattern was reported by Kori- at and Ma’ayan (2005): Par- ticipants made JOLs immediately after study, after a short delay or after a long delay. The effects of encoding fluency on JOLs were found to decrease with JOL delay, whereas

those of retrieval fluency were found to increase. These changes mirrored the changes that occurred in cue validity: Encoding fluency was a better predictor of recall than retrieval fluency immediately after study, but the reverse was true after some delay. The parallel between the effects of delay on cue utilization and cue validity was taken to suggest a flexible and adaptive utilization of different mnemonic cues according to their relative validity in predicting memory performance.

We should stress, however, that, although processing fluency is, by and large, correlated positively with memory performance, the correlation is not perfect; and, in fact, there are conditions in which reliance on fluency as a basis for recall predictions may be misleading (e.g., Benjamin et al., 1998). When fluency is nondiagnostic, it can impair metacognitive calibration (or absolute accuracy; see Koriati, 2007), leading to inflated memory predictions overall. In the present study, however, we focused on resolution (or relative accuracy); that is, the ability to judge the relative memorability of different items in a list. Resolution is likely to be impaired if list items differ in the extent to which they evoke unwarranted feelings of mastery. For example, Koriati and Bjork (2005) showed that some study items are particularly prone to instill illusions of competence, because in these items, the target (or answer), which is present at study, is likely to activate aspects of the cue that will not come forward at test, when only the cue is presented. The inclusion of these items in a study list not only impairs calibration, leading to inflated JOLs overall, but also impairs resolution—the ability to discriminate between items that are likely to be recalled and those that are not.

In concluding this article, we should comment on the counterintuitive nature of the results reported in this study. These results appear to run counter to the established wisdom in learning research—that memory improves with repeated presentation of an item (the famous “learning curve,” Ebbinghaus, 1885/1964) and with rate of presentation (e.g., the “total time hypothesis,” see Cooper & Pantle, 1967). Indeed, JOLs and recall have been shown to increase (although not at the same rate) with repeated presentation of the list (Koriati et al., 2002; Serra & Dunlosky, 2005), as was also found in Experiment 3 here; and both JOLs and recall increased with rate of presentation when rate of presentation was varied between items in a list (Koriati, 1997).

Furthermore, Bjork (1999; see also Simon & Bjork, 2001) cited extensive evidence in support of the claim that conditions of learning that create difficulties for learners, thus slowing their rate of learning, generally improve long-term retention. For example, varying the conditions of learning, or spacing the study sessions, not only improve long-term memory, but also reduce illusions of competence and improve the correspondence between metacognitive predictions and actual memory performance. In contrast, Koriati, Ma’ayan, and Nussinson (2006) obtained results consistent with the ELER principle, even when four months elapsed between study and test.

What distinguishes the conditions of the present study from the typical conditions in which metacognitive judgments and actual memory performance increase with amount of study time and with number of study opportu-

nities? This question was addressed by Koriati, Ma’ayan, and Nussinson (2006) in terms of the distinction between data-driven and goal-driven metacognitive regulation. In the experimental situations investigated in this study, the amount of effort invested in attempting to commit an item to memory is *data driven*, determined essentially by the item itself. The learner is, in a sense, a “participant observer”: He or she invests as much encoding effort as the item affords or calls for, and then uses the feedback from the experience of studying as a basis for recall predictions. Under such conditions, JOL and recall are expected to correlate negatively across items with the amount of encoding effort invested in each item.

In contrast, when self-regulation is *goal driven*, increased processing effort enhances both metacognitive feelings and actual memory performance. Thus, a student may strategically regulate study time in the service of various goals, spending more time studying topics that are associated with higher incentives. In that case, JOL and recall should correlate positively with study effort (see Koriati, Ma’ayan, & Nussinson, 2006).

Similarly, when presentation duration is experimentally determined, as is usually the case in most learning experiments, slower rates of presentation allow learners to engage in learning strategies that are hard to implement under faster rates of presentation. Also, manipulations that appear to create difficulties for the learners (Bjork, 1999) may benefit memory, because they induce learners to engage in effective processes that they would not have used otherwise.

Koriati, Ma’ayan, and Nussinson (2006) reported results suggesting the occurrence of data-driven and goal-driven effects within the same experimental task. Thus, when some study items were associated with higher incentives than others, the high-incentive items elicited longer study times and higher JOLs than did the low-incentive items, producing a positive correlation between study time and JOLs: the signature of goal-driven regulation. Within each level of incentive, however, longer study times were associated with lower JOLs, as would be expected for data-driven regulation.

In fact, the contrasting effects of data-driven and goal-driven regulation are nicely illustrated by the results depicted in Figure 5. As far as interitem differences in study time are concerned, JOLs decreased with study time, consistent with the pattern expected for data-driven regulation. In contrast, the results for individual differences are consistent with goal-driven regulation: Individuals who deliberately spend more time studying (presumably because of higher motivation or interest) exhibit higher JOLs as well as better memory performance. Thus, differences between individuals affected JOLs and memory performance in the opposite way from that observed for differences between items.

These results bring to the fore the importance of attribution processes (see Jacoby & Whitehouse, 1989; Schwarz, 2004; Strack, 1992). Although JOLs may be based on the feedback from one’s study experience, particularly the degree of fluency or effort during learning, such feedback may carry different implications depending on whether fluency and effort are attributed to properties of the task,

to one's own initiative, or to external factors. Thus, monitoring accuracy should depend not only on the general validity of the inferential heuristics underlying metacognitive judgments, but also on the attribution of fluency and effort to their proper sources.

AUTHOR NOTE

The research was conducted at the Institute of Information Processing and Decision Making, University of Haifa, Haifa, Israel. I gratefully acknowledge support for this research by the German Federal Ministry of Education and Research (BMBF) within the framework of German-Israeli Project Cooperation (DIP). I thank Hilit Ma'ayan for her help in the initial stages of this research, Rinat Gil for conducting the experiments, and Limor Sheffer for her help in the analyses of the data. Correspondence concerning this article should be addressed to A. Koriat, Department of Psychology, University of Haifa, Haifa 31905, Israel (e-mail: akoriat@research.haifa.ac.il).

REFERENCES

- BAHRICK, H. P., & PHELPS, E. (1987). Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 344-349.
- BENJAMIN, A. S., & BJORK, R. A. (1996). Retrieval fluency as a metacognitive index. In L. Reder (Ed.), *Implicit memory and metacognition* (pp. 309-338). Hillsdale, NJ: Erlbaum.
- BENJAMIN, A. S., BJORK, R. A., & SCHWARTZ, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, **127**, 55-68.
- BJORK, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance. Interaction of theory and application* (pp. 435-459). Cambridge, MA: MIT Press.
- BRUNSWIK, E. (1956). *Perception and representative design in psychological experiments*. Berkeley: University of California Press.
- COOPER, E. H., & PANTLE, A. J. (1967). The total-time hypothesis in verbal learning. *Psychological Bulletin*, **68**, 221-234.
- EBBINGHAUS, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Dover. (Originally published as *Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie* 1885)
- FROST, R., & PLAUT, D. (2005). *The word-frequency database for printed Hebrew*. Retrieved March 21, 2005, from word-freq.mscc.huji.ac.il/index.html.
- HERTZOG, C., DUNLOSKY, J., ROBINSON, A. E., & KIDDER, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 22-34.
- JACOBY, L. L., & WHITEHOUSE, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, **118**, 126-135.
- KELLEY, C. M., & LINDSAY, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory & Language*, **32**, 1-24.
- KELLEY, C. M., & RHODES, M. G. (2002). Making sense and nonsense of experience: Attributions in memory and judgment. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 41, pp. 293-320). San Diego: Academic Press.
- KING, J. F., ZECHMEISTER, E. B., & SHAUGHNESSY, J. J. (1980). Judgments of knowing: The influence of retrieval practice. *American Journal of Psychology*, **93**, 329-343.
- KORIAT, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, **100**, 609-639.
- KORIAT, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, **124**, 311-333.
- KORIAT, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, **126**, 349-370.
- KORIAT, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289-325). Cambridge: Cambridge University Press.
- KORIAT, A., & BJORK, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 187-194.
- KORIAT, A., & BJORK, R. A. (2006). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory & Cognition*, **34**, 959-972.
- KORIAT, A., & MA'AYAN, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory & Language*, **52**, 478-492.
- KORIAT, A., MA'AYAN, H., & NUSSINSON, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General*, **135**, 36-69.
- KORIAT, A., MA'AYAN, H., SHEFFER, L., & BJORK, R. A. (2006). Exploring a mnemonic debiasing account of the underconfidence-with-practice effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **32**, 595-608.
- KORIAT, A., SHEFFER, L., & MA'AYAN, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General*, **131**, 147-162.
- MATVEY, G., DUNLOSKY, J., & GUTTENTAG, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition*, **29**, 222-233.
- MAZZONI, G., CORNOLDI, C., & MARCHITELLI, G. (1990). Do memorability ratings affect study-time allocation? *Memory & Cognition*, **18**, 196-204.
- NELSON, T. O., & LEONESIO, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect." *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 676-686.
- NELSON, T. O., & NARENS, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 26, pp. 125-169). San Diego: Academic Press.
- ROBINSON, M. D., JOHNSON, J. T., & HERNDON, F. (1997). Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. *Journal of Applied Psychology*, **82**, 416-425.
- SCHWARTZ, B. L. (1994). Sources of information in metamemory: Judgments of learning and feeling of knowing. *Psychonomic Bulletin & Review*, **1**, 357-375.
- SCHWARZ, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, **14**, 332-348.
- SERRA, M. J., & DUNLOSKY, J. (2005). Does retrieval fluency contribute to the underconfidence-with-practice effect? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 1258-1266.
- SIMON, D. A., & BJORK, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 907-912.
- STRACK, F. (1992). The different routes to social judgments: Experiential versus informational strategies. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 249-276). Hillsdale, NJ: Erlbaum.
- SWIFT, E. J. (1918). *Psychology and the day's work: A study in the application of psychology to daily life*. London: Allen & Unwin.

(Manuscript received May 13, 2007;
revision accepted for publication July 31, 2007.)