Journal of Experimental Psychology:
Learning, Memory, and Cognition
2006, Vol. 32, No. 5, 1133–1145

0278-7393/06/$12.00 DOI: 10.1037/0278-7393.32.5.1133

Mending Metacognitive Illusions: A Comparison of Mnemonic-Based and Theory-Based Procedures

Asher Koriat
University of Haifa

Robert A. Bjork
University of California, Los Angeles

Previous research indicated that learners experience an illusion of competence during learning (termed *foresight bias*) because judgments of learning (JOLs) are made in the presence of information that will be absent at test. The authors examined the following 2 procedures for alleviating foresight bias: enhancing learners' sensitivity to mnemonic cues pertaining to ease of retrieval and inducing learners to resort to theory-based judgments as a basis for JOLs. Both procedures proved effective in mending metacognitive illusions—as reflected in JOLs and self-regulation of study time—but only theory-based debiasing yielded transfer to new items. The results support the notion that improved metacognition is 1 key to optimizing transfer but also that educating subjective experience does not guarantee generalization to new situations.

Keywords: metacognition, judgments of learning (JOLs), transfer of learning, illusions of competence, foresight bias

In this study, we examined the extent to which interventions that are designed to mend spoiled metacognitions exhibit transfer beyond the original context of training. Specifically, we evaluate the effectiveness of two types of procedures for alleviating the illusions of competence that learners sometimes experience when studying new material. Whereas both types of procedures were expected to improve metacognitive monitoring and metacognitive control in a second presentation of the same materials, they were hypothesized to exhibit differential effects with regard to transfer to new material.

In our previous work (Koriat & Bjork, 2005, in press), we documented a metacognitive bias that is largely unavoidable whenever learners, during the study of new materials, need to assess the likelihood that they will be able to recall that material on a subsequent memory test. This bias, which we have termed a *foresight bias*, derives from a characteristic that is inherent to the relationship between study and test sessions—namely, that learners typically assess their degree of mastery of the studied material in the presence of the information that they might be asked to recall during testing. Foresight biases, and accompanying overconfidence, are most likely to occur when an answer that is solicited during testing is judged to be natural or obvious when presented along with the question but is actually less likely to come forward when the question is presented alone.

Asher Koriat, Department of Psychology, University of Haifa, Haifa, Israel; Robert A. Bjork, Department of Psychology, University of California, Los Angeles.

This research was supported by United States–Israel Binational Science Foundation Grant No. 97-34. We thank Hana Struminger for programming the experiments, Rinat Gil for conducting the experiments, and Limor Sheffer for the statistical analyses. We are indebted to John Dunlosky for his comments on a previous draft.

Correspondence concerning this article should be addressed to Asher Koriat, Department of Psychology, University of Haifa, Haifa 31905, Israel. E-mail: akoriat@research.haifa.ac.il

Such a discrepancy between the study and test situations can be conceptualized in terms of the distinction between a priori and a posteriori associations. Consider, for example, a paired-associates task in which participants are required to study the pair and to assess, at the end of each trial, the probability that they will be able to recall the target (response) word when presented with the cue word. Foresight bias will be strongest when the a posteriori association, which refers to the perceived relationship between the cue and target when both are present (as during study), is inordinately strong relative to the a priori association, which refers to the probability that the cue word, when presented alone, will elicit the target word. Whereas a posteriori association is best measured by subjective judgments of the degree of relatedness between the cue and the target, a priori association is best measured by word-association norms. In general, when the cue word tends to elicit many associates other than the target, the a posteriori cue–target association will tend to be inflated relative to the a priori association.

To illustrate, consider the case of word pairs that have differing strengths of backward and forward associations, such as the pair *umbrella–rain*. The likelihood of *umbrella* eliciting *rain*, according to word association norms, is .70, whereas that of *rain* eliciting *umbrella* is only .04 (D. L. Nelson, McEvoy, & Schreiber, 1999). We found that backward-associated pairs (such as *rain–umbrella*) produced very inflated judgments of learning (JOLs), presumably because the to-be-recalled target activates aspects of the cue that are less likely to come forward when the cue is presented alone, as in cued-recall test.

In other experiments, we used pairs such as *citizen–tax* for which the association is purely a posteriori, that is, pairs in which neither word is an a priori associate of the other but which appear related when presented together. These pairs also produced inflated JOLs. Furthermore, across several experiments, words that were randomly paired and also had no association between them ("unrelated") consistently yielded inflated JOLs, presumably be-

cause participants could form connections between the words, and such connections induced illusions of competence (Koriat & Bjork, 2005, in press). Altogether, the results suggest that illusions of competence do not occur across the board. Rather, JOLs are particularly inflated when the target word presented during study brings to the fore aspects of the cue word that are less likely to emerge during testing, when the cue word is presented alone.

In this study, we focused on debiasing procedures that can help reduce the illusions of competence that derive from misleading a posteriori associations. In particular, we examined the question of whether the effects of such procedures generalize beyond the original context in which they were first applied. This question was motivated by discussions in the literature stressing the role that metacognitive skills play in transfer of training. Given the preponderance of findings testifying for failures to achieve transfer of learning (see Perkins & Salomon, 1994), it has been argued that awareness of one's cognitive processes and the monitoring and regulation of these processes can nevertheless contribute to transfer (Bransford & Schwartz, 1999; De Corte, 2003; Mayer & Wittrock, 1996; see also Bransford, Brown, & Cocking, 2000). Indeed, metacognitive reflection on one's mental processes has been found to promote transfer of skills (Lin & Lehman, 1999; Lucangeli, Galderisi, & Cornoldi, 1995). Thus, training programs that incorporate metacognitive activities, such as reciprocal teaching (Palincsar & Brown, 1984) and procedural facilitation (Scardamalia, Bereiter, & Steinbach, 1984), have proven successful in increasing the degree to which students transfer their learning to new settings and events. These studies suggest that interventions that improve the effectiveness of metacognitions should enhance transfer of training to new materials and, furthermore, that the improved metacognitive skills themselves also transfer readily to new situations (e.g., Salomon, Globerson, & Guterman, 1989). As noted by De Corte (2003), however, the mechanisms underlying the beneficial effects of metacognitive skills on transfer of learning are not clear.

Theory-Based and Mnemonic-Based Metacognitive Judgments

The proposed link between metacognitive skills and transfer of training requires a reassessment in light of the current distinction in metacognition research between theory-based (or information-based) and mnemonic-based (or experience-based) metacognitive judgments (Koriat & Levy-Sadot, 1999; see also Dunlosky & Nelson, 1997; Kelley & Jacoby, 1996; Koriat, 1993, 1997; Koriat, Bjork, Sheffer, & Bar, 2004; Strack, 1992). Theory-based judgments rely on a deliberate, explicit deduction from rules and theories retrieved from memory in making inferences about one's state of knowledge. Mnemonic-based judgments are also inferential in nature, but they rely on a variety of internal cues that are used automatically and unconsciously to give rise to sheer subjective feelings. Such feelings are then used as the immediate basis for metacognitive judgments.

Consider, for example, the monitoring of one's own knowledge during study. It has been proposed that JOLs can be based on an analytic inference that draws on beliefs or theories, such as "I have poor memory" (Dunning, Johnson, Ehrlinger, & Kruger, 2003) or "memory for studied information is better if tested soon after study than if tested after a long delay" (Koriat et al., 2004). Analytic inferences are presumably not very different from those underlying many everyday predictions (e.g., which team is likely to win a football game). Clearly, the accuracy of theory-based JOLs should depend on the validity of the underlying theory or belief.

JOLs, however, can also be based on heuristics that rely on internal, mnemonic cues. For example, it has been proposed that JOLs monitor the fluency of encoding the information or the fluency of retrieving it during learning (Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Benjamin & Bjork, 1996; Benjamin, Bjork, & Schwartz, 1998; Dunlosky & Nelson, 1992; Hertzog, Dunlosky, Robinson, & Kidder, 2003; Koriat, 1997; Koriat & Ma'ayan, 2005; T. O. Nelson, Narens, & Dunlosky, 2004). The assumption is that processing fluency gives rise to a sheer subjective feeling of knowing, which can then serve as the basis for a recall prediction (see Schwarz, 2002). The validity of such intuitive feeling depends, of course, on the validity of the mnemonic cues in predicting future memory performance.

The aim of this study was to compare the effectiveness of theory-based and mnemonic-based procedures in mending metacognitive illusions. As will be detailed below, we propose that only debiasing procedures that draw on one's theories and beliefs should be expected to also ensure some degree of transfer to new materials. The effects of mnemonic-based debiasing, in contrast, should be confined to the original training material with little generalization to new material. In fact, however, experimental researchers in the area of metacognition have put a much greater emphasis on mnemonic-based than on theory-based JOLs (e.g., Benjamin & Bjork, 1996), whereas discussions of metacognitive skills that stress their contribution to transfer of training seem to relate more to theory-based rather than mnemonic-based metacognitive processes.

Debiasing Procedures That Can Mend Metacognitive Illusions

Let us return to the foresight bias. If, as we suggest, this bias is a largely unavoidable consequence of intrinsic differences between typical study and test situations, how can faulty monitoring and the accompanying overconfidence be mended? As sketched in the foregoing discussion, the following two approaches seem promising: increasing tuning to the mnemonic cues that are pertinent to the retrieval of the response term during testing (mnemonic-based debiasing), or helping learners formulate an effective theory that can serve as an alternative basis for JOLs, replacing misleading mnemonic cues (theory-based debiasing).

In a previous study (Koriat & Bjork, in press), we focused on the former type of debiasing and explored two procedures that were hypothesized to reduce the foresight bias by providing learners with mnemonic cues that are diagnostic of recall performance: delayed JOLs and study–test practice. The delayed-JOL procedure was based on the finding that JOLs are more accurate when they are delayed until shortly after study than when they are made immediately after study (T. O. Nelson & Dunlosky, 1991; see also Dunlosky & Nelson, 1992, 1994). Presumably, whereas immediate JOLs rely heavily on encoding fluency and are therefore strongly affected by the presence of the target, delayed JOLs, when prompted by the cue alone, tend to rely primarily on retrieval fluency, which is more diagnostic of the criterion performance (cued recall; Koriat & Ma'ayan, 2005). Therefore, we expected

delayed JOLs to be better protected against the contaminating effects of inflated a posteriori associations that occur in the presence of the target. Indeed, delaying JOLs was found to reduce the foresight bias that is characteristic of immediate JOLs (Koriat & Bjork, in press, Experiment 4; see also Koriat, Ma'ayan, Sheffer, & Bjork, 2006).

The second debiasing procedure was study–test practice. Several observations suggest that the experience of studying and recalling a list of items provides learners with mnemonic cues about the relative fluency of encoding and retrieving these items, so that when they study these items again, the accuracy of their JOLs in predicting recall improves. Matvey, Dunlosky, Shaw, Parks, & Hertzog (2002) have referred to the benefit that ensues from previous study–test blocks as knowledge updating based on task experience. The results of Koriat (1997) and Koriat, Ma'ayan, and Nussinson (2006) suggest that two changes occur with repeated study of a list of paired associates. First, learners increasingly rely on internal mnemonic cues (such as processing fluency) in making JOLs rather than on such cues as preexperimental judgments of item difficulty. Second, the validity of the internal cues in predicting recall increases. Both of these contribute to the improvement in JOL-recall correlation with study–test practice. We hypothesized that such practice should also help alleviate the foresight bias. Indeed, with repeated study–test blocks, learners appeared to gain information about the relative difficulty of learning and remembering the forward and backward pairs; the difference in JOLs between the two types of pairs became gradually closer to their difference in recall. Study–test practice was also found to mend the flawed allocation of study time that presumably ensues from the foresight bias. Under self-paced learning conditions, participants spent about the same time studying backward-associated and forward-associated pairs on the first presentation of the list, but on subsequent presentations, a larger proportion of the time was allocated to the study of the backward pairs (Experiment 5). Additional experiments (Experiments 2 and 3) established that it is specifically test experience rather than study experience that yields the most benefit in terms of bringing JOLs into closer alignment with recall, consistent with the assumption that such illusions derive from the inherent discrepancy between the study and test situations. Possibly, then, repeated testing is one way in which the subjective experience underlying JOLs can be educated.

One complication that we encountered in assessing the metacognitive benefits that ensue from study–test experience should be mentioned, because it is relevant to the experimental work to be reported. This complication derives from the underconfidence-with-practice (UWP) effect that has been documented in several studies (Finn & Metcalfe, 2004; Koriat, Ma'ayan, et al., 2006; Koriat, Sheffer, & Ma'ayan, 2002; Meeter & Nelson, 2003; Scheck & Nelson, 2005; Serra & Dunlosky, 2005). When participants are presented with the same list of paired associates for several study–test cycles, their JOLs exhibit a shift toward marked underconfidence from the second study–test cycle on. Thus, for example, unrelated word pairs typically exhibit an overconfidence bias in the first study–test block but an underconfidence bias in the second block (Koriat & Bjork, in press; Koriat et al., 2002). In fact, in Koriat and Bjork's study (Experiment 1), even the backward pairs, which yielded a marked overconfidence bias in the first study–test block, exhibited a marked underconfidence bias in the second study–test block (Koriat & Bjork, in press). The sources of the UWP effect are not yet entirely clear, but its occurrence made it difficult to describe the effects of practice in terms of improved metacognitive accuracy, although the reduction in overconfidence with practice was clearly stronger for the pairs with inflated a posteriori association. Therefore, the strategy adopted by Koriat and Bjork (in press) was to assess the benefits of practice not in terms of the JOL-recall discrepancy but rather in terms of the ability to discriminate between items that are prone to yield a foresight bias and those that do not. Thus, as already noted, it was found that indeed study–test practice improved the discrimination between forward and backward pairs such that the differences between them in JOLs tended to mirror more closely their difference in recall.

In sum, our previous study, which focused on mnemonic-based debiasing, indicated that metacognitive illusions can be remedied by providing learners with mnemonic cues pertaining to ease of retrieval. Soliciting JOLs after some delay and providing participants with study–test experience helped to reduce the inordinately high JOLs associated with items that induce a foresight bias.

The Aims of the Experiments

The aims of Experiment 1 of this study were twofold. First, in addition to a mnemonic-debiasing procedure, which involved study–test practice with the same list of items, we included a theory-debiasing procedure in which learners were induced to relinquish the misleading mnemonic cues underlying JOLs and to switch to theory-based JOLs. Second, in order to examine transfer of training, we manipulated whether the list to be studied in the second phase of the experiment was the same as the list in the first phase or a new list. The study lists included forward-associated pairs, which have been previously shown to yield good JOL-recall correspondence, and backward-associated pairs, which are known to yield inflated JOLs in the first study–test cycle.

We hypothesized that both debiasing procedures would be effective in increasing discrimination between the forward and backward pairs when the same list is used in the two blocks, but only the effects of theory-based debiasing would transfer to a new list of items. The rationale for this prediction is that because mnemonic-based metacognitive judgments are mediated by nonanalytic heuristics (see Jacoby & Brooks, 1984) that operate below full consciousness, they achieve their effects implicitly and automatically, without the learner's ability to spell out the underlying principle. Therefore, the improved monitoring that ensues from study–test experience is item-specific and should not generalize to new items that were not included in the training session. In contrast, theory debiasing is expected to help participants articulate a general rule, which they can then apply to a new set of items.

Whereas Experiment 1 concerned metacognitive monitoring, Experiment 2 focused on metacognitive control. Specifically, the goal of that experiment was to examine whether theory-based and mnemonic-based debiasing procedures have effects on the allocation of study time that are consonant with their effects on JOLs. Will such procedures increase the investment of study time in items that typically produce a foresight bias, and, if so, will such effects generalize to new materials given theory-based debiasing but not mnemonic-based debiasing? Such a pattern would demonstrate that any monitoring effects observed in Experiment 1 are not mere epiphenomena but that they do actually exert a causal influence on behavior (see Koriat, 2000; T. O. Nelson, 1996).

The expected interaction between type of metacognitive debiasing and stimulus material (same vs. different list) would demonstrate, on the one hand, that metacognitive training that helps educate subjective experience need not automatically ensure generalization to new situations and, on the other hand, that a good theory that helps replace faulty metacognitions can have better practical benefits (see Koriat et al., 2004).

Experiment 1

In Experiment 1, participants were presented with two blocks, each including the study of a list of paired associates followed by a cued-recall test. The lists consisted of Hebrew word pairs with asymmetric associative strength, pairs that were similar to the English pairs used in Experiment 2 of Koriat and Bjork (2005). That is, the association from Word A to Word B was much stronger than the association from Word B to Word A. The mnemonic-debiasing procedure capitalized on the finding that practice studying a list of paired associates helps alleviate the foresight bias for backward associated pairs (Koriat & Bjork, in press). Thus, the first block consisted of a study–test cycle on one list of items, whereas the second block (which took place after a filler task) involved a study–test cycle using either the same list as in the first block or a second, new list.

The theory-debiasing procedure also included a training session that was intended to make participants aware of the conditions that lead to a foresight bias. After the first study–test block but before the second, participants were presented with a short list of asymmetric word-association pairs and were asked to guess the percentage of participants who would be likely to give the target word in response to the cue word. They were then given the actual percentages (based on norms) and were induced to consider the foresight bias as a source of the discrepancy between the estimated and actual percentages of responding. Finally, they were given a second study–test cycle using the same or a different list of paired associates. Thus, the design of Experiment 1 conformed to a Debiasing Procedure (mnemonic-based vs. theory-based) × List Repetition (same list vs. different list) factorial.

Method

Participants. The participants were 96 Hebrew-speaking University of Haifa undergraduates (61 women and 35 men); 18 served for course credit and 78 were paid for their participation.

Materials. Two lists of Hebrew paired-associates were used. These were compiled from a norming study that was conducted by Koriat and Bjork (in press;, Experiment 3) in an attempt to identify Hebrew word pairs with asymmetric associations. Each list included 36 word pairs composed of two equal sets of 18 pairs that were matched in terms of the strength of the forward and backward associations. For the first list, associative strength for the forward and backward directions averaged .493 and .012, respectively, for one set, and .488 and .018, respectively, for the other set. For the second list, the respective means were .510 and .016 for one set and were .482 and .013 for the other set. For each list, one set was assigned to the forward direction (with the strongest association being from the cue word to the target word), and the other set was assigned to the backward direction (with the assignment then counterbalanced across participants). In addition to the 36 asymmetrical pairs, 18 unrelated pairs were included in each list, each with zero associative strength.

Apparatus and procedure. The experiment was conducted on a Silicon Graphics personal computer. The stimuli were displayed on the computer screen, and JOLs and recall, spoken by the participants, were entered by the experimenter on a keyboard.

The procedure included two study–test blocks. Participants were instructed that they would have to study 54 paired associates and assess the chances that they would be able to recall the target word in response to the cue word in a subsequent test that would take place after the whole list has been presented. During the study phase, the two words appeared side by side for 3 s and were replaced by the statement "Probability to recall: _____." Participants provided JOLs on a 0%–100% scale, expressing their prediction of recalling the target word in response to the cue, and the next pair was presented as soon as the experimenter recorded the data on the keyboard. During the test phase, the 54 cue words were presented one after the other in a random order. Participants had to say the response word within 6 s. After the first study–test cycle, half of the participants were randomly assigned to the theory-debiasing condition, and the rest were assigned to the mnemonic-debiasing condition. In addition, half of the participants in each group were assigned to the same-list condition, and the remaining participants were assigned to the different-list condition.

In the theory-debiasing group, participants were told that the focus of the study was on the feeling of knowing that learners have when studying new material and, in particular, on the reasons for the inflated feelings of competence that they sometimes experience during study in comparison with their performance during testing. It was specified that one likely source of the inflated sense of competence derives from the type of associations that come to mind while studying new material. They were then shown a list of 10 new asymmetric pairs (taken from the same preliminary norming study as the experimental pairs) and were asked to estimate for each pair the percentage of people who would produce the target word when asked to say the first word that comes to mind in response to the cue word in a word-association test.

After participants in the theory-debiasing groups completed the estimation task, they were shown the actual percentages. The experimenter then focused on several pairs for which there was a large gap between the estimated and actual percentages and asked the participant to try to identify the source of such discrepancies, after which the experimenter introduced the foresight-bias hypothesis. It was pointed out that the presence of the target in conjunction with the cue sometimes inflates its perceived likelihood in comparison with when the cue appears alone and that word associations are not necessarily symmetrical. The participant was then asked to try to apply the reasoning underlying the foresight bias in explaining the discrepancy between estimated and actual percentages for the other word pairs.

Finally, participants in the theory-debiasing group were told that the same type of foresight bias as described with regard to the word-association task may have occurred when they had made JOLs during the study of paired associates, because in the study phase, both the cue word and target word were present on each trial, whereas only the cue word was presented at test. Because the cue word could elicit many other possible responses, the recall estimates may have been inflated exactly as in the word-association test.

In the mnemonic-debiasing group, participants engaged in two filler tasks before moving to the second block. Each participant was yoked to one theory-debiasing participant in the same list-repetition condition (same list vs. different list) and was administered the filler tasks for the same amount of time that the yoked participants took to go through the theory-debiasing procedure.

At the beginning of the second study–test blocks, participants in the theory-debiasing condition were reminded to apply the theory in making their JOLs.

Results

Table 1 presents mean JOLs and recall as a function of block (first vs. second) for all combinations of list repetition (same list

Table 1
Experiment 1: Mean Judgments of Learning (JOLs) and Recall as a Function of Block (First vs. Second) for the Unrelated, Backward, and Forward Pairs for All Combinations of Condition (Mnemonic vs. Theory Debiasing) and List Repetition (Same List vs. Different List).

	Pair type					
	Unrelated		Backward		Forward	
Debiasing condition measure	First block	Second block	First block	Second block	First block	Second block
Same list across study–test blocks						
Mnemonic debiasing						
JOL	26.1	35.9	69.2	76.1	73.4	84.2
Recall	17.7	50.3	64.2	89.1	86.3	98.6
Theory debiasing						
JOL	34.6	24.4	77.5	68.9	82.0	82.1
Recall	17.1	39.5	62.2	88.6	85.2	96.8
Different list across study–test blocks						
Mnemonic debiasing						
JOL	34.5	28.0	72.0	69.4	79.0	76.5
Recall	17.0	21.3	66.7	72.4	89.1	93.3
Theory debiasing						
JOL	22.0	9.4	72.5	52.5	79.0	71.5
Recall	17.3	22.0	70.1	73.2	90.7	91.2

vs. different list) and condition (mnemonic debiasing vs. theory debiasing). In the first set of analyses, we focused on the first study–test block for which the procedure was the same for all participants.

The JOL-recall correspondence for the first study–test block. The results for the first study–test block were unexpected in that the forward pairs yielded an underconfidence bias unlike what has been observed in previous studies. Thus, in three previous experiments that contrasted forward and backward pairs (in both Hebrew and English), JOLs for the forward pairs were either well calibrated (Koriat & Bjork, 2005, Experiment 2) or exhibited a small overconfidence bias of about 4% (Koriat & Bjork, in press, Experiment 1; Koriat, Ma'ayan, et al., 2006, Experiment 2, immediate-JOL condition). JOLs for the backward pairs, in contrast, were consistently inflated, demonstrating a marked overconfidence bias of about 20% on average. Here, in contrast, the backward pairs yielded an overconfidence bias as in previous studies, $t(95) = 3.12$, standard error of difference of means ($SEdm$) = 2.25, $p < .005$, but the forward pairs exhibited a significant underconfidence bias, $t(95) = 5.51$, $SEdm = 1.72$, $p < .0001$. The reason for the underconfidence bias in this experiment is unclear.

It should be noted that in the three previous experiments mentioned above, the unrelated pairs also yielded an overconfidence bias of about 13%. As noted earlier, the inflated JOLs for such pairs seem to derive from the tendency of participants to perceive a relationship between words that are not directly related according to word-association norms. In Experiment 1, JOLs were significantly inflated for the unrelated pairs, $t(95) = 6.34$, $SEdm = 1.90$, $p < .0001$.

In the following analyses, we focus on the difference between the forward and backward pairs, examining the extent to which the mnemonic and theory-debiasing procedures helped in increasing the discrimination between them. Consider first the results for the first study–test block. It can be seen that the forward–backward differences in JOLs were much smaller than the corresponding differences in recall: Across all participants, JOLs averaged 78.3% and 72.8% (a 5.5% difference) for the forward and backward pairs, whereas percent recall averaged 87.8% and 65.8% (a 22.0% difference), respectively. An Associative Direction (forward vs. backward) × Measure (JOL vs. recall) analysis of variance (ANOVA) yielded $F(1, 95) = 172.94$, $MSE = 105.46$, $p < .0001$, for associative direction; $F < 1$ for measure; and $F(1, 95) = 90.28$, $MSE = 72.49$, $p < .0001$, for the interaction.

We also repeated the two-way ANOVA separately for the mnemonic-debiasing and theory-debiasing groups. The interaction was significant for each group, $F(1, 47) = 48.73$, $MSE = 68.48$, $p < .0001$, and $F(1, 47) = 41.12$, $MSE = 78.02$, $p < .0001$, respectively. A three-way ANOVA, Associative Direction (forward vs. backward) × Measure (JOL vs. recall) × Condition (mnemonic vs. theory), yielded $F < 1$ for the triple interaction, suggesting that there were no systematic differences between the mnemonic and theory groups in the susceptibility to the foresight bias.

In sum, the results for the first study–test block indicated that participants failed to fully appreciate the effects of associative direction on recall, presumably because of the inflated a posteriori associations of the backward pairs.

The UWP effect for the same-list conditions. Before examining the effects of the two debiasing procedures on the discrimination between the forward and backward pairs, we should note that the results, in general, disclose a trend in the direction of underconfidence, consistent with the UWP effect (Koriat et al., 2002). This trend is most clearly seen for the unrelated pairs. For the

mnemonic-debiasing condition, these pairs yielded a near-significant overconfidence bias in the first block, $t(23) = 2.04$, $SEdm = 4.13$, $p < .06$, which changed to a significant underconfidence bias in the second block, $t(23) = 2.94$, $SEdm = 4.89$, $p < .01$. Similarly, for the theory-debiasing condition, an overconfidence bias was found in the first block, $t(23) = 4.36$, $SEdm = 4.03$, $p < .001$, but an underconfidence bias was evident in the second block, $t(23) = 4.45$, $SEdm = 3.38$, $p < .001$.

The backward pairs also exhibited a crossover interaction similar to that observed for the unrelated pairs. The results for the forward pairs, in contrast, are not as clear, possibly because recall performance for these pairs approached ceiling. These results reinforce our strategy of focusing on the sensitivity to associative direction as the main dependent variable rather than on the absolute values of JOLs and recall.

Comparing the effects of the two debiasing procedures: Same lists. We compare now the effects of the two debiasing manipulations, focusing first on the results for the groups that received the same list of words in both blocks. The results indicate that although practice studying the same list induced or intensified an underconfidence bias, it helped increase sensitivity to the effects of associative direction (see also Koriat & Bjork, in press, Experiment 1). Whereas in the first block, the JOLs associated with the forward and backward pairs differed by only 4.2% and 4.5% in the mnemonic-debiasing and theory-debiasing conditions, respectively, the corresponding differences on the second block were 8.1% and 13.2%, respectively. A three-way ANOVA, Block (first vs. second) × Condition (mnemonic vs. theory) × Associative Direction (forward vs. backward), yielded $F(1, 46) = 1.43$, $MSE = 180.66$, *ns*, for block; $F < 1$ for condition; and $F(1, 46) = 59.44$, $MSE = 45.23$, $p < .0001$, for associative direction. The interaction between block and associative direction, however, was also significant, $F(1, 46) = 17.02$, $MSE = 27.93$, $p < .001$, indicating increased sensitivity to associative direction with practice. Note that, in these analyses, we did not take into account recall performance, because recall on the second block approached ceiling for the forward pairs in some of the cells.

The Condition × Block interaction was also significant, $F(1, 46) = 11.38$, $MSE = 180.86$, $p < .005$, reflecting the observation that JOLs increased from the first to the second block in the mnemonic-debiasing condition (as is generally the case, see Koriat et al., 2002) but decreased in the theory-debiasing condition. The decrease observed for the theory-debiasing condition may stem from the fact that JOLs in the first block were relatively high for that condition, but it is also possible that the theory-debiasing manipulation, which emphasized the bias induced by seeing the cue and target pairs together, reduced predictions for the unrelated and forward pairs as well.

In sum, practice studying the same list seems to increase sensitivity to associative direction, and there was no significant difference in this respect between the mnemonic and theory conditions. The results for the mnemonic group replicate the previous finding reported by Koriat and Bjork (in press), and the novel results for the theory group did not differ in this respect.

Comparing the effects of the two debiasing procedures: Different lists. We turn next to the two groups receiving a different list on the second block. As can be seen in Table 1, JOLs associated with the forward and backward pairs in the first block differed by 7.0% and 6.5% in the mnemonic-debiasing and theory-debiasing conditions, respectively. In the second block, in contrast, the respective differences were 7.1% and 19.0%. Thus, whereas in the theory-debiasing condition the effects of associative direction increased for the different-list group as it did for the same-list groups, mnemonic-debiasing failed to yield such increase.

These conclusions were supported by a three-way ANOVA on JOLs, Block (first vs. second) × Condition (mnemonic vs. theory) × Associative Direction (forward vs. backward). This analysis yielded $F(1, 46) = 19.40$, $MSE = 163.46$, $p < .0001$, for block; $F(1, 46) = 1.58$, $MSE = 880.23$, *ns*, for condition; and $F(1, 46) = 70.88$, $MSE = 66.03$, $p < .0001$, for associative direction. The following interactions were significant: Condition × Block, $F(1, 46) = 9.26$, $MSE = 163.46$, $p < .005$; Condition × Associative Direction, $F(1, 46) = 5.92$, $MSE = 66.03$, $p < .05$; Block × Associative Direction, $F(1, 46) = 19.40$, $MSE = 24.68$, $p < .0001$; and Condition × Block × Associative Direction, $F(1, 46) = 18.24$, $MSE = 24.68$, $p < .0001$.

The Condition × Block interaction derives from an observation that is somewhat similar to that found for the same-list groups. Whereas in the mnemonic-debiasing group, JOLs on the second list (second block) did not differ significantly from those made for the first list (first block), $t(23) = 1.34$, $SEdm = 1.89$, $p < .20$, in the theory-debiasing condition, JOLs were significantly lower (by 13.8%) on the second list, $t(23) = 4.34$, $SEdm = 3.18$, $p < .001$, and even lower for the backward pairs.

Of greater importance, the triple interaction (as well as the Condition × Associative Direction interaction) reflects the observation that only the theory manipulation increased sensitivity to the forward–backward distinction, whereas the mnemonic condition yielded no such effect. Thus, in the theory-debiasing condition, JOLs decreased from the first to the second list by 7.5% for the forward pairs and by 20.0% for the backward pairs. For that condition, a Block × Associative Direction ANOVA yielded $F(1, 23) = 30.98$, $MSE = 29.98$, $p < .0001$, for the interaction. In the mnemonic-debiasing condition, in contrast, JOLs decreased by 2.5% and 2.6%, respectively, for the forward and backward pairs, and neither the effects of block nor the interaction were significant, $F < 1$ and $F(1, 23) = 1.76$, $MSE = 85.87$, $p < .21$, respectively.

Comparing the same-list and different-list conditions. Examination of the results presented in Table 1 indicates that the theory manipulation, in fact, exerted a similar beneficial effect whether the same list or a different list was used in the second block. Thus, for the theory-debiasing condition, a Block (first vs. second) × List Repetition (same vs. different) × Associative Direction (forward vs. backward) ANOVA on JOLs yielded significant effects for block, $F(1, 46) = 17.28$, $MSE = 224.17$, $p < .0001$; for associative direction, $F(1, 46) = 83.87$, $MSE = 66.36$, $p < .0001$; and for the Block × Associative Direction interaction, $F(1, 46) = 46.00$, $MSE = 29.08$, $p < .0001$. These effects reflect the observations that JOLs decreased from the first to the second block and decreased more so for the backward than for the forward pairs. The effects of list repetition and the interaction between list repetition and block were also significant, $F(1, 46) = 4.78$, $MSE = 770.33$, $p < .05$, and $F(1, 46) = 4.85$, $MSE = 224.17$, $p < .05$, respectively, but the triple interaction was not, $F(1, 46) = 1.46$, $MSE = 29.08$.

Mnemonic debiasing, in contrast, seems to have exerted different effects when the same list was repeated than when a different list was used. Thus, as noted earlier, when a different list was used

in the second block, JOLs decreased by about the same amount for the forward and backward pairs (2.5% and 2.6%, respectively). In contrast, when the same list was repeated, JOLs increased by 10.8% for the forward pairs compared with only 6.9% for the backward pairs. However, the triple interaction between block, list repetition, and associative direction was not significant, $F(1, 46) = 1.76$, $MSE = 23.52$, $p < .20$.

Changes in monitoring resolution. We also examined the changes in monitoring resolution that occurred with practice. Resolution refers to the extent to which participants' JOLs discriminate between items that are eventually recalled and those that are not and is generally indexed by the within-subject JOL-recall gamma correlation (see T. O. Nelson, 1984). Focusing first on the participants who received the same list in both blocks, gamma correlations across all items averaged .62 and .81 for the first and second blocks of the mnemonic group, respectively, $t(23) = 5.44$, $SEdm = 0.036$, $p < .0001$. The respective means for the theory group were .62 and .82, $t(23) = 6.78$, $SEdm = 0.029$, $p < .0001$. These results are consistent with previous findings that practice studying the same list improves resolution (e.g., Cull & Zechmeister, 1994; Koriat, 1997; Koriat et al., 2002). It is somewhat surprising that the theory-based group did not exhibit a stronger improvement in resolution from the first to the second block than the mnemonic-based group. We have previously observed that the improvement in resolution that occurs as a result of practice does not derive solely from increased sensitivity to differences between pair types (e.g., forward vs. backward pairs) but also from increased sensitivity to interitem differences within each class of pairs (Koriat, Ma'ayan, et al., 2006). Nevertheless, the results raise the question whether reliance on rules and theories conflicts with the ability to use internal mnemonic cues about interitem differences in processing fluency (see Wilson & Schooler, 1991).

In contrast, for participants who received a different list in the second block, there was no improvement in resolution from the first to the second list. For the mnemonic condition, gamma correlations for the first and second lists averaged .66 and .68, respectively, $t(23) = 0.49$, $SEdm = 0.034$, *ns*. The respective means for the theory condition were .72 and .70, $t(23) = 0.33$, $SEdm = 0.044$, *ns*. Thus, the improvement in resolution that occurs with practice (see Cull & Zechmeister, 1994; Koriat, 1997; Shaughnessy & Zechmeister, 1992) appears to be item specific—it is obtained only when the same list is repeated and does not transfer to a new list.

Discussion

Experiment 1 yielded results that were generally consistent with our predictions. First, the results for the first block replicated the foresight bias that we have reported previously, with backward-associated pairs and unrelated pairs yielding inflated JOLs. Unexpectedly, however, the forward pairs exhibited an underconfidence bias. Second, both debiasing procedures proved effective in increasing sensitivity to associative direction when the same list of items was restudied. Finally, and more important for present purposes, there were differential effects of the two debiasing procedures on transfer to a new list. Whereas the theory-debiasing procedure was effective in reducing the discrepancy between the forward and backward pairs in JOL-recall correspondence, the mnemonic-debiasing procedure yielded little such improvement. This pattern of results brings to the fore a fundamental difference between the two types of debiasing and lends further support to the usefulness of the distinction between theory-based and mnemonic-based metacognitive processes: Whereas the effects of mnemonic-debiasing appear to be entirely item-specific (confined to the materials used in the first block), the effects of theory-debiasing are generalizable to new materials.

This pattern of results has important implications for transfer of training, which is considered to be one of the most fundamental goals of educators (De Corte, 2003; Perkins & Salomon, 1994). It would seem that when learners discover and formulate a rule on the basis of their study experience, they can apply that rule beyond the initial learning context (Gick & Holyoak, 1983). Such is not the case for mnemonic-based debiasing, in which, apparently, participants' study–test experience equips them with useful mnemonic cues about the recallability of different items to the extent of improving their monitoring on a repeated study of these items but provides them with little insight into what they have learned from study–test practice. The benefits from mnemonic-based debiasing appear, therefore, to be contextualized—that is, tied to the specific items used in training.

Experiment 2

Whereas Experiment 1 concerned metacognitive monitoring, Experiment 2 focused on metacognitive control. Thus, Experiment 2 was similar to Experiment 1, except that study time was self-paced rather than fixed, so that participants could regulate the amount of study time invested in each item. Also, no JOLs were solicited. Our focus in this experiment was on the consequences that mnemonic and theory debiasing might have on mending a biased allocation of study time during self-paced learning

A common assumption among students of metacognition is that monitoring is not an epiphenomenon but actually plays a causal role in guiding cognitive processes and behavior (Koriat, in press; Koriat & Goldsmith, 1996; T. O. Nelson & Narens, 1990). Thus, judgments of learning have been assumed to affect the allocation of study time during self-paced learning, so that more study time is invested in items that are associated with lower JOLs (see T. O. Nelson & Leonesio, 1988; Son & Schwartz, 2002). Thiede, Anderson, and Therriault (2003), for example, observed that a manipulation that improved learner's monitoring accuracy in studying text resulted in a more effective regulation of study and, in turn, in overall better test performance. Thus, learners seem to rely on their metacognitive feelings in regulating their behavior, and, to the extent that these feelings are accurate, such reliance can sometimes help improve memory performance.

If so, we should expect the illusion of competence resulting from the foresight bias to have detrimental effects on the allocation of study time, causing learners to allocate relatively less study time to items with inflated a posteriori associations. The two debiasing procedures used in Experiment 1 are expected to mend the allocation of study time toward increasing the proportion of study time invested in the backward pairs when given the opportunity to restudy the same list of items (see Koriat & Bjork, in press). When participants are required to study a new list of items, however, the shift in the policy of study time allocation should be found only following theory debiasing and not following mnemonic debiasing. Such results would indicate that the effects of debiasing

procedures on metacognitive monitoring have their parallel effects on metacognitive control.

Method

Participants. Participants were 48 Hebrew-speaking University of Haifa undergraduates (36 women and 12 men); 2 served for course credit and 46 were paid for their participation.

Materials. Two new lists of Hebrew paired-associates were used. They were compiled from a norming study that was conducted by Rubinsten, Anaki, Henik, Drori, and Faran (2005).[1] Each list included 36 asymmetrically associated word pairs, which were divided into two equal sets that were matched in terms of the strength of the forward and backward associations. For the first list, associative strength for the forward and backward directions averaged .390 and .041, respectively, for one set and .383 and .045, respectively, for the other set. For the second list, the respective means were .391 and .041 for one set and were .388 and .045 for the other set. For each list, one set was assigned to the forward condition and the other to the backward condition, and the assignment was counterbalanced across participants. In addition, a set of 36 unrelated pairs was included in each list, each pair with zero associative strength.

Apparatus and procedure. The experiment was conducted on an IBM compatible personal computer. The procedure was the same as that of Experiment 1 with the following exceptions. First, study time was self-paced rather than fixed; participants were instructed to study each item for as long as they needed and to press the left key of the mouse when they were through, at which time the next pair would appear on the screen. They were instructed to invest the exact time they needed for studying each word pair, no more and no less. They were also told that their success in performing the task would depend on their ability to recall as many words as possible from the list at test while keeping the total time invested in studying the entire list as short as possible. They were informed that the list included 72 paired associates.

Second, no JOLs were solicited. Thus, during the study phase, each word pair remained on the screen until the left mouse key was pressed, and 1 s thereafter, the next pair was shown. The procedure for the test phase was the same as in Experiment 1. Third, in both blocks, a 6-min filler task was introduced between the study and test phases to reduce the likelihood of a ceiling effect on recall as a result of the self-paced feature of the experiment. Finally, the procedure for the theory-debiasing condition was similar, but not identical, to the procedure used in Experiment 1 (because of the elimination of JOLs). Participants were given the same theoretical explanation, as in Experiment 1, including practicing the word-association test. However, instead of detailing the implications for JOLs, participants were told that the amount of time spent studying each item is likely to be influenced by its judged difficulty. They were urged to consider the foresight-bias hypothesis in allocating study time to the various items on the second block.

Results

Examination of the study-time distribution revealed a very large between-participant variation. Mean study time for the first block, calculated across all participants, averaged 6.61 s/item. However, this mean ranged from 2.02 to 27.21 across participants.

Effects of associative direction on study time and recall in the first block. Assuming that study time is affected by participants' covert monitoring of the relative difficulty of the items, then it should be relatively insensitive to the direction of association in the first block, whereas recall testing should evidence better memory for the forward than for the backward pairs. Indeed, this is precisely the pattern that was consistently found across the four groups. As can be seen in Figure 1 (left panel), mean study time per item was practically identical for the forward and backward pairs for each of the four groups (4.26 s and 4.60 s, respectively, across all groups) and both means were substantially lower than the mean time spent studying the unrelated pairs (8.78 s). In contrast, memory performance yielded a different pattern (see Figure 1, right panel), with the backward pairs exhibiting consistently inferior performance than the forward pairs (58.9% and 83.1%, respectively, across all groups).

Statistical analyses confirmed these impressions. A two-way ANOVA on the results for the first block, Pair Type (backward, forward, unrelated) × Measure (study time vs. recall), yielded a significant interaction, $F(2, 94) = 175.71$, $MSE = 90.09$, $p < .0001$. With regard to study time, a t test confirmed that more time was spent studying the unrelated than the backward pairs, $t(47) = 7.93$, $SEdm = 0.53$, $p < .0001$, suggesting that participants controlled study time allocation according to the judged difficulty of the items (see T. O. Nelson & Leonesio, 1988), but there was no difference between the amount of time spent studying the backward and the forward pairs, $t(47) = 1.78$, $SEdm = 0.19$, *ns*. This pattern was equally observed for all four experimental groups. A three-way ANOVA on study time, Condition (theory vs. mnemonic) × List Repetition (same vs. different) × Associative Direction using only the forward and backward pairs, yielded no significant main effects or interactions. The difference between the forward and backward pairs yielded $F(1, 44) = 2.99$, $MSE = 0.92$, $p < .10$.

As far as recall is concerned, the backward pairs yielded markedly lower recall than the forward pairs, $t(47) = 11.13$, $SEdm = 2.17$, $p < .0001$, although they exhibited better recall than the unrelated pairs, $t(47) = 6.88$, $SEdm = 3.29$, $p < .0001$. The inferior recall for the backward versus the forward pairs was found for each of the experimental groups. A Condition × List Repetition × Associative Direction ANOVA using only the forward and backward pairs yielded a significant effect for associative direction, $F(1, 44) = 116.85$, $MSE = 119.92$, $p < .0001$, and no other main or interaction effects.

Mending metacognitive illusions. Because of the large between-participant variation in response time, we focused on the relative amount of time invested in the study of the forward and backward pairs in each of the two blocks. For each participant, we calculated the difference between the average times spent studying the backward and forward pairs as a proportion of the average amount of time spent on all forward and backward pairs. The means of these averages are plotted in Figure 2 (top panel) for each group as a function of block. A Block (2) × Group (4) ANOVA yielded a significant effect for block, $F(1, 44) = 18.32$, $MSE = 0.05$, $p < .0001$, but not for group, $F(3, 44) = 2.00$, $MSE = 0.07$, $p < .14$. The interaction, however, was significant, $F(3, 44) = 3.48$, $MSE = 0.05$, $p < .05$. It can be seen that the mnemonic-different group exhibited little increase in the time allocated to the backward pairs relative to the forward pairs, $t(11) = 0.37$, $SEdm = 0.05$, *ns*. In contrast, such an increase was significant for the mnemonic-same group, $t(11) = 3.61$, $SEdm = 0.09$, $p < .005$, and for the theory-same group, $t(11) = 2.87$, $SEdm = 0.10$, $p < .05$,

[1] These norms, based on a relatively large number of participants, were not available when Experiment 1 was conducted.

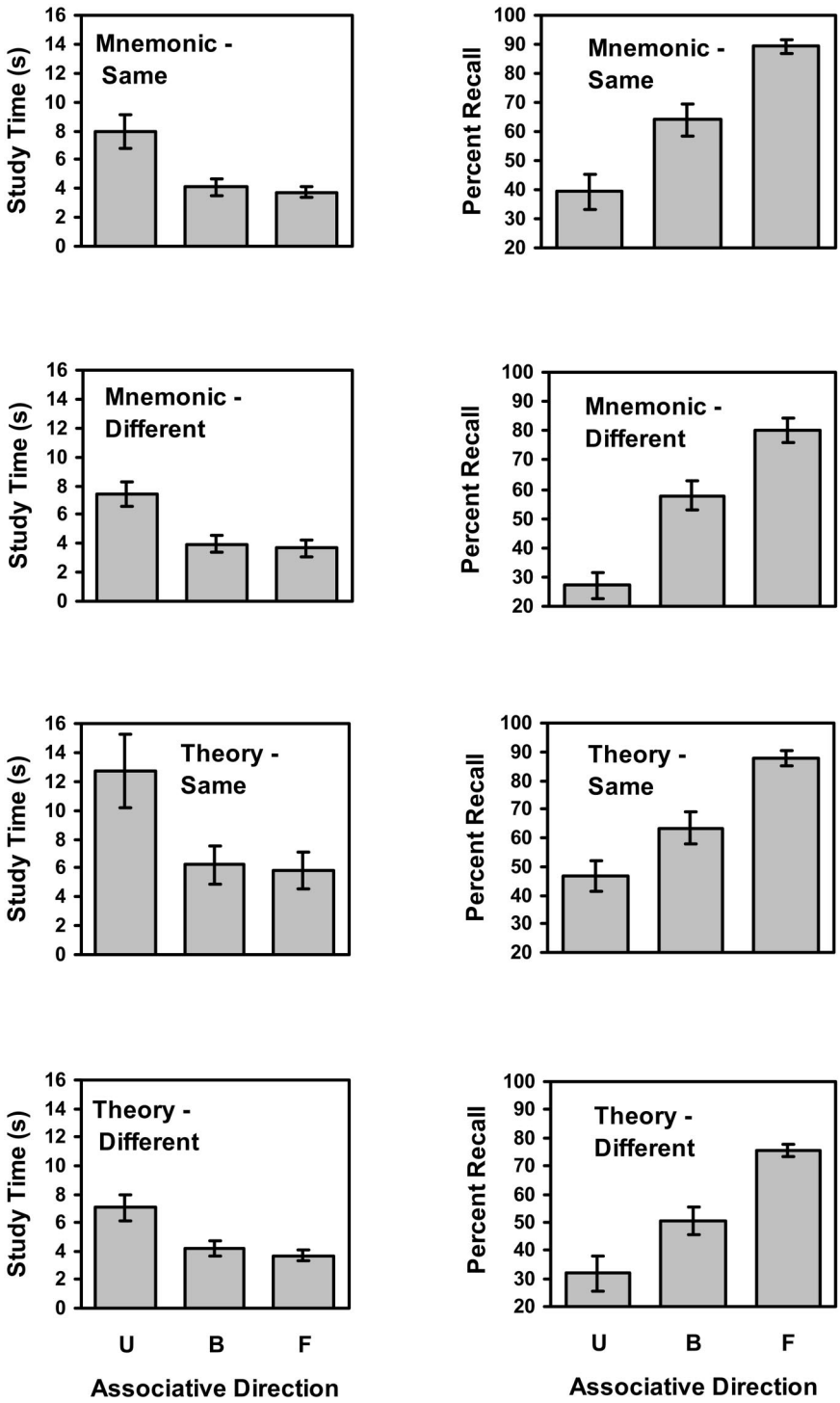

Figure 1. Mean study time per item (left panel) and percent recall (right panel) for unrelated (U), forward (F), and backward (B) pairs for the first block for each of the four experimental groups. The error bars represent ± 1 *SEM* (Experiment 2).

and was near significant for the theory-different group, $t(11) = 1.49$, $SEdm = 0.16$, $p < .17$.

For comparison purposes, in the bottom panel of Figure 2, we present the corresponding results for JOLs in Experiment 1. These results reflect the JOL difference between forward and backward pairs as a function of block. The results mimic closely those for study time in Experiment 2. The similarity between the two panels of Figure 2 suggests that the two debiasing procedures exert

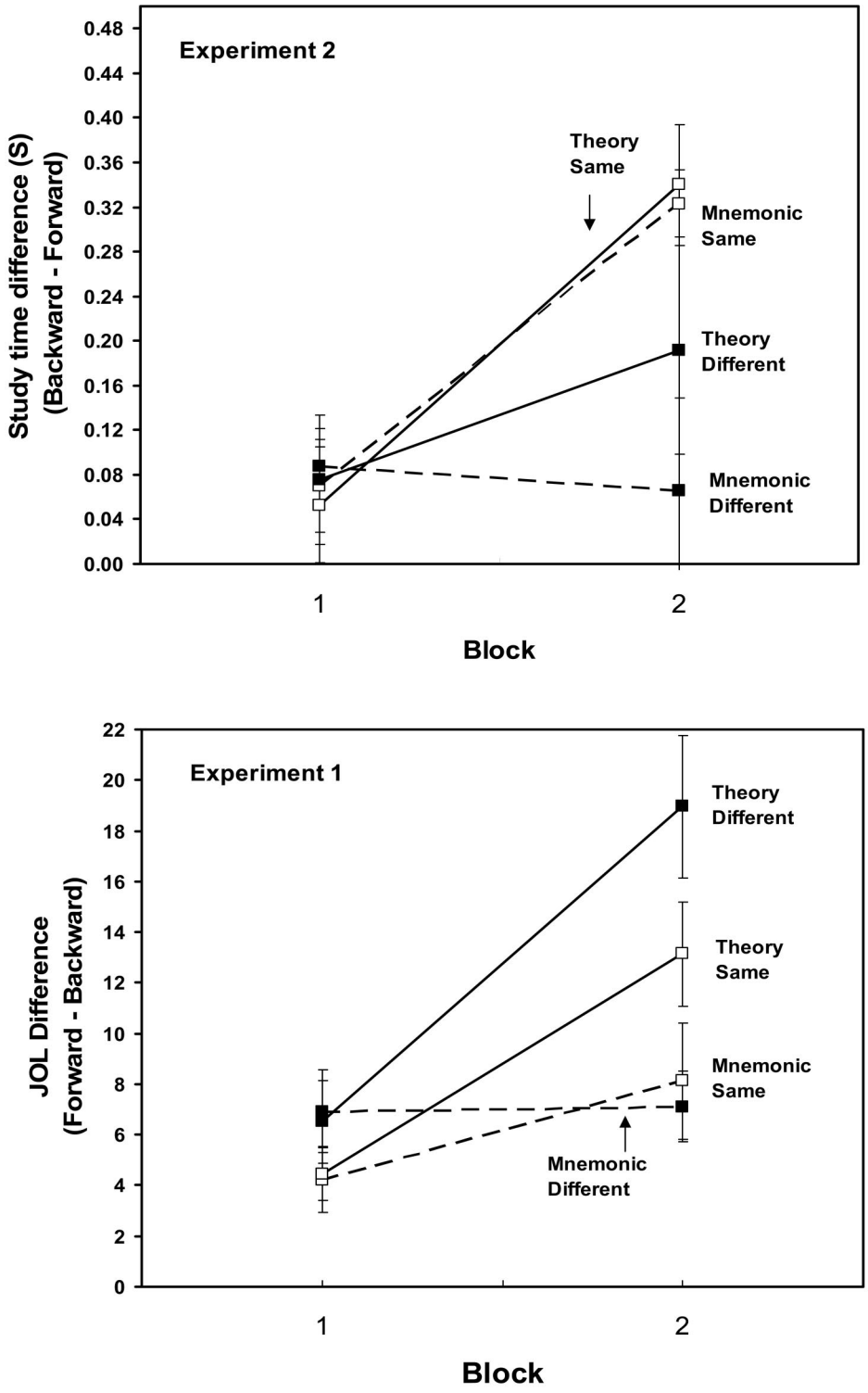

Figure 2. Mean difference between the backward and forward pairs in study time allocated to each pair, calculated as a proportion of the mean time allocated to backward and forward pairs (top panel: Experiment 2), and mean difference between the forward and backward pairs in judgments of learning (JOLs; bottom panel: Experiment 1) as a function of block for each of the four experimental groups. The error bars represent ± 1 *SEM*.

similar effects on study time allocation (control) as they do on JOLs (monitoring).

A question that arises is whether the changes that occurred in the allocation of study time in the second block had any effects on recall. We examined that question focusing only on the different-list groups for which there was a differential effect of condition on study time allocation, A Condition × Block ANOVA yielded $F(1, 22) = 12.02$, $MSE = 93.54$, $p < .005$, for block, but the effects of condition and the interaction were not significant, $F < 1$ and $F(1, 22) = 1.63$, $MSE = 93.54$, $p < .22$, respectively. For the mnemonic group, recall for the first and second blocks averaged 68.9% and 75.0%. The respective means for the theory group were 62.9% and 76.2%. Thus, there is no indication that the increased investment of study time in the backward pairs reduced the forward–backward difference in recall. This result is consistent with similar observations documenting a "labor-in-vain effect" (see T. O. Nelson & Leonesio, 1988).

Discussion

Experiment 2 examined the expected consequences of the foresight bias and its alleviation on the allocation of study time during self-paced learning. As predicted, the effects of associative direction mimicked those observed for JOLs in Experiment 1 in three critical respects. First, on the first block, learners did not appropriate more study time to the backward than to the forward pairs, presumably because of the foresight bias induced by the former pairs. Second, when the same list was presented for restudy, both debiasing procedures were effective in inducing a change in the policy of study time allocation—a relatively greater amount of time was allocated to the backward than to the forward pairs. Third, however, when a different list was used in the second block, such a change in study time allocation was observed only for theory debiasing but not for mnemonic debiasing. These results suggest that the conclusions from Experiment 1 regarding the conditions that produce transfer of training on metacognitive monitoring hold true with regard to metacognitive control as well.

General Discussion

In this study, we focused on the contrast between two different bases for metacognitive judgments and between two corresponding ways in which illusions of knowing can be debiased. Although our results lend further support to the importance of the distinction between theory-based and mnemonic-based metacognitive judgments, they also bring to the fore its implications for the education of subjective experience and the alleviation of metacognitive illusions.

Clearly, one way in which metacognition can be trained is by providing people with valid knowledge about the operation of the cognitive system and inducing them to apply that knowledge in forming metacognitive judgments. This approach has been prevalent in developmental studies of metacognition (see Koriat & Shitzer-Reichert, 2002). The assumption underlying these studies is that part of the learning and memory deficits observed in young children derive from ignorance about the factors that affect cognitive performance or from the failure to apply one's knowledge to concrete learning situations (see Schneider & Pressley, 1997). Our results not only demonstrate the benefits of explicit metacognitive training but also indicate its potential usefulness in fostering transfer to new situations.

Subjective experience, however, may also be trained by procedures that improve the quality of the mnemonic cues on which participants base their judgments. In fact, many of the dissociations reported in the literature between metacognitive judgments and memory performance appear to derive from reliance on mnemonic cues that are poorly diagnostic of performance (e.g., Benjamin et al., 1998; Koriat, 1995; Reder & Ritter, 1992). As far as JOLs are concerned, two procedures have proved effective in improving metacognitive accuracy—delaying JOLs and providing participants with study–test experience. These manipulations apparently equip learners with mnemonic cues that are better tuned to the retrieval of the to-be-remembered items.

The finding that mnemonic-based debiasing is item specific has important theoretical and practical implications. This finding supports the assumption regarding the nonanalytic, implicit nature of the process underlying mnemonic-based metacognitive judgments (Kelley & Jacoby, 1996; Koriat, 1997; Whittlesea, Jacoby, & Girard, 1990; but see Matvey, Dunlosky, & Guttentag, 2001). It suggests that the foresight bias emanating from inflated a posteriori associations is not based on a faulty belief or on the application of an erroneous rule. Rather, it is based on the overall fluency that derives from the cue–target pair as a whole, which gives rise to a sheer subjective feeling, not to an explicit inference.

Presumably, metacognitive training procedures that capitalize on mnemonic cues for educating subjective experience produce their effects by increasing sensitivity to diagnostic cues about each of the studied items without learners being able to spell out the principle that cuts across different items (e.g., a rule that states that inflated a posteriori associations induce an illusion of competence). Therefore, the effects of mnemonic debiasing do not transfer to new materials.

In contrast, the theory-debiasing procedure induced participants to discover and articulate a general rule and to apply that rule in a subsequent study–test cycle. This procedure was found to foster transfer beyond the original learning context. Not only did theory debiasing help mend metacognitive judgments for the new materials, but it also resulted in a more effective allocation of study time in studying these materials. These results demonstrate the beneficial role of insight and understanding for transfer of training. In fact, the theory-debiasing procedure we used incorporates some of the basic ingredients of the procedure that promote transfer in its broader, more active sense (see De Corte, 2003). Whereas transfer has been conceived traditionally as the immediate application of knowledge acquired in one situation to another, recent analyses emphasize the importance of metacognitive skills and incorporate the notion that transfer must entail active preparation for future learning (Bransford & Schwartz, 1999; Mayer & Wittrock, 1996; see also Bransford et al., 2000).

Consistent with that view, it has been proposed that one of the conditions for transfer is the assessment of one's competencies, so that learners can make an accurate appraisal of the effort needed to accomplish a task successfully. Indeed, De Corte (2003) reported results indicating that a well-designed training program that fosters self-judging and stimulates reflection on learning can yield successful transfer to a new situation. Such conditions would seem to be absent in the mnemonic-debiasing condition.

Concluding Comment

The present study demonstrated differential effects of two debiasing procedures on measures of both metacognitive monitoring and metacognitive control. These results have two broad and related implications. First, they demonstrate that metacognitive training that helps educate subjective experience need not—and perhaps typically does not—ensure generalization to new situations. And second, that a good theory that helps replace faulty metacognitions does tend to generalize, supporting Lewin's (1945, p. 129) dictum that "nothing is so practical as a good theory."

References

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language, 28,* 610–632.

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.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (Vol. 24, pp. 61–100). Washington, DC: American Educational Research Association.

Cull, W. L., & Zechmeister, E. B. (1994). The learning ability paradox in adult metamemory research: Where are the metamemory differences between good and poor learners? *Memory & Cognition, 22,* 249–257.

De Corte, E. (2003). Transfer as the productive use of acquired knowledge, skills, and motivations. *Current Directions in Psychological Science, 12,* 142–146.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20,* 374–380.

Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language, 33,* 545–565.

Dunlosky, J., & Nelson, T. O. (1997). Similarity between the cue for judgments of learning (JOL) and the cue for test is not the primary determinant of JOL accuracy. *Journal of Memory and Language, 36,* 34–49.

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12,* 83–87.

Finn, B., & Metcalfe, J. (2004, November). *Multitrial judgments of learning.* Poster presented at the 45th annual meeting of the Psychonomic Society, Minneapolis, MN.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1–38.

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, and Cognition, 29,* 22–34.

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 18, pp. 1–47). San Diego, CA: Academic Press.

Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language, 35,* 157–175.

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. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition, 9,* 149–171.

Koriat, A. (in press). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *Cambridge handbook of consciousness.* New York: 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, and Cognition, 31,* 187–194.

Koriat, A., & Bjork, R. A. (in press). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory & Cognition.*

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 133,* 643–656.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103,* 490–517.

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 483–502). New York: Guilford Publications.

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and 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, and Cognition, 3,* 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.

Koriat, A., & Shitzer-Reichert, R. (2002). Metacognitive judgments and their accuracy: Insights from the processes underlying judgments of learning in children. In P. Chambres, M. Izaute, & P.-J. Marescaux (Eds.), *Metacognition: Process, function, and use* (pp. 1–17). New York, NY: Kluwer.

Lewin, K. (1945). The research center for group dynamics at Massachusetts Institute of Technology. *Sociometry, 8,* 126–135.

Lin, X., & Lehman, J. D. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching, 36,* 837–858.

Lucangeli, D., Galderisi, D., & Cornoldi, C. (1995). Specific and general transfer effects following metamemory training. *Learning Disabilities Research & Practice, 10,* 1–21.

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.

Matvey, G., Dunlosky, J., Shaw, R. J., Parks, C., & Hertzog, C. (2002). Age-related equivalence and deficit in knowledge updating of cue effectiveness. *Psychology and Aging, 17,* 589–597.

Mayer, R. C., & Wittrock, M. C. (1996). Problem solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 47–62). New York: Macmillan Reference Books.

Meeter, M., & Nelson, T. O. (2003). Multiple study trials and judgments of learning. *Acta Psychologica, 113,* 123–132.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). *The University of South Florida word association, rhyme and word fragment norms.* Retrieved December 2, 1999, from http://luna.cas.usf.edu/~nelson/

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95,* 109–133.

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51,* 102–116.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science, 2,* 267–270.

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, and Cognition, 14,* 676–686.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation*: *Advances in research and theory* (Vol. 26, pp. 125–140). San Diego, CA: Academic Press.

Nelson, T. O., Narens, L., & Dunlosky, J. (2004). A revised methodology for research on metamemory: Pre-judgment recall and monitoring (PRAM). *Psychological Methods, 9,* 53–69.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1,* 117–175.

Perkins, D. N., & Salomon, G. (1994). Transfer of learning. In T. Husén & T. N. Postlethwait (Eds.), *International encyclopedia of education, 2nd ed.* (Vol. 11, pp. 6452–6457). Oxford, England: Pergamon Press

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 435–451.

Rubinsten, O., Anaki, D., Henik, A., Drori. S., & Faran, Y., (2005). Norms for free associations in the Hebrew language. In A. Henik, O. Rubinsten, & D. Anaki (Eds.), *Word norms for the Hebrew language* (pp. 17–34). Be'er Sheva, Israel: Ben Gurion University of the Negev.

Salomon, G., Globerson, T., & Guterman, E. (1989). The computer as a zone of proximal development: Internalizing reading-related metacognitions from a reading partner. *Journal of Educational Psychology, 81,* 620–627.

Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflective processes in written composition. *Cognitive Science, 8,* 173–190.

Scheck, P., & Nelson, T. O. (2005). Lack of pervasiveness of the underconfidence-with-practice effect: Boundary conditions and an explanation via anchoring. *Journal of Experimental Psychology: General, 134,* 124–128.

Schneider, W., & Pressley, M. (1997). *Memory development between 2 and 20.* New York: Springer-Verlag.

Schwarz, N. (2002). Feelings as information: Moods influence judgments and processing strategies. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 534–547). New York: Cambridge University Press.

Serra, M. J., & Dunlosky, J. (2005). Does retrieval fluency contribute to the underconfidence-with-practice effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1258–1266.

Shaughnessy, J. J., & Zechmeister, E. B. (1992). Memory-monitoring accuracy as influenced by the distribution of retrieval practice. *Bulletin of the Psychonomic Society, 30,* 125–128.

Son, L. K., & Schwartz, B. L. (2002). The relation between metacognitive monitoring and control. In T. J. Perfect & B. S. Schwartz (Eds.), *Applied metacognition* (pp. 15–38). Cambridge, United Kingdom: Cambridge University Press.

Strack, F. (1992). The different routes to social judgments: Experiential versus informational strategies. In I. I. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 249–275). Hillsdale, NJ: Erlbaum.

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95,* 66–73.

Whittlesea, B. W., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language, 29,* 716–732.

Wilson, T. D., & Schooler, J. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60,* 181–192.

Received July 6, 2005
Revision received December 29, 2005
Accepted March 6, 2006 ■