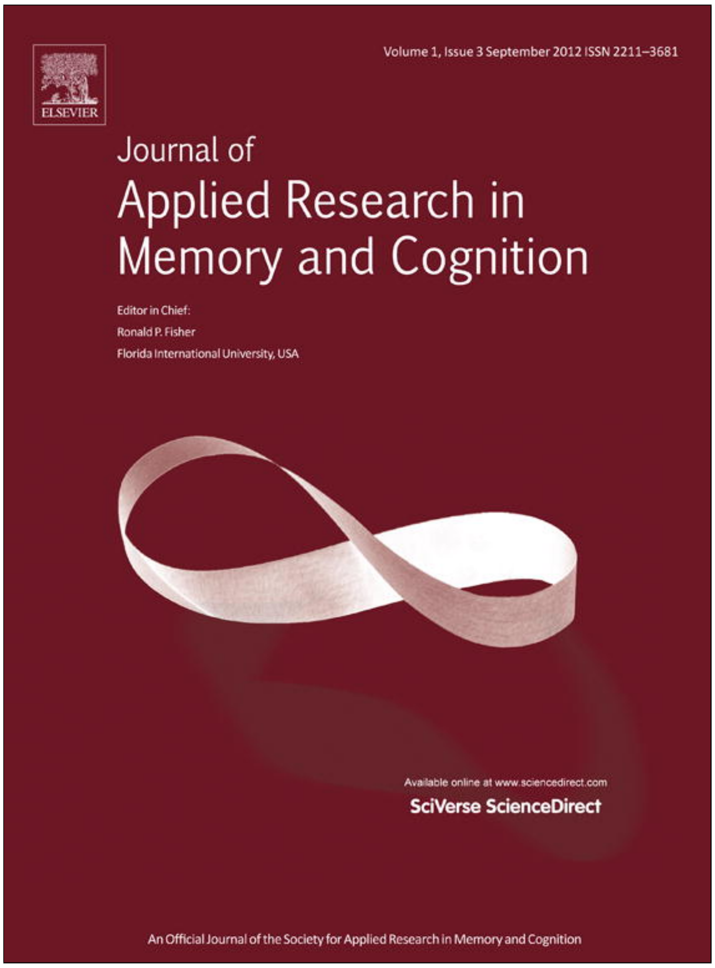
Volume 1, Issue 3 September 2012 ISSN 2211-3681
ELSEVIER
Journal of
Applied Research in
Memory and Cognition
Editor in Chief:
Ronald P. Fisher
Florida International University, USA
Available online at www.sciencedirect.com
SciVerse ScienceDirect
An Official Journal of the Society for Applied Research in Memory and Cognition

Journal of Applied Research in Memory and Cognition 1 (2012) 145–151

Contents lists available at SciVerse ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Literature Review

Metamemory and memory efficiency: Implications for student learning

Bennett L. Schwartz[a,*], Anastasia Efklides[b]

[a] *Florida International University, USA*
[b] *School of Psychology, Aristotle University of Thessaloniki, Greece*

ARTICLE INFO

Article history:
Received 19 December 2011
Received in revised form 1 June 2012
Accepted 5 June 2012
Available online 13 June 2012

Keywords:
Metacognitive experiences
Metamemory
Judgments of learning
Resolution
Calibration
Memory strategies

ABSTRACT

Judgments of learning (JOLs) are metamemory judgments about the likelihood of remembering later an item that we are currently studying. Much research has documented that JOLs are accurate at discriminating easy from difficult items (resolution), but often fail to account for other factors such as the advantages of retrieval practice, overlearning, spacing, and desirable difficulty, thus resulting in poor calibration. This implies that JOLs should not be used to determine *how long* or *in what manner* to study, although they can be used to determine *what* to study. To counter poor calibration, explicit metamemory knowledge can be learned and applied in order for students to become more efficient learners. Our goal is to focus on how metamemory influences decisions about studying, and how we might use metamemory to improve our learning efficiency.

1. Introduction

Decisions on learning and studying are complex. They depend on task demands and individual characteristics such as prior knowledge, motivation, and affect (Ariel, Dunlosky, & Bailey, 2009; Efklides, 2011). Metamemory concerns both the micro-level, the metamemory processes that prompt people to focus on some items rather than others, and the macro-level, the strategic metacognitive knowledge that guides the regulation of our studying, learning, and remembering (Dunlosky & Rawson, in press; Efklides, 2006, 2011; Nelson & Narens, 1990). Our goal is not an exhaustive review of the literature on metamemory (see Dunlosky & Metcalfe, 2009; Efklides, 2011; Rhodes & Tauber, 2011) but rather to examine the problems in how micro-level metamemory guides learning and remembering and how metamemory guides our study, given the complexity of learning.

Metamemory experiences are the moment-to-moment feelings that we have as we study, learn, and retrieve. Functionally, metamemory experiences are the output of monitoring processes that check the progress of learning and the likelihood of remembering. Metamemory experiences allow us to “feel” what we will learn, remember, and forget. For example, a person “feels” as though he has mastered the endings for “re” verbs in the present tense in French. Metamemory experiences can form the basis for control decisions, which then translate metamemory into action. For instance, one feels as if the French verbs are well remembered, so no further study is required. The focus of this paper is on how the metamemory experience of judgments of learning (henceforth, JOLs) allows us to monitor and subsequently control our learning, and how JOLs might be employed to improve learning efficiency.

1.1. How closely do metamemory studies mirror what real learners actually do?

A concern among psychologists is the extent to which lab findings generalize to real-world situations. In this regard, metamemory research sets a fine example, moving back and forth from the lab to the classroom (see Efklides, 2011; Rawson & Dunlosky, 2011). In particular, research focuses on how students use JOLs or JOL-like processes in ordinary learning and how confidence and study time interact in test-takers. To apply what we learn from the lab to real-world learning, it is important to know what students do when confronted with information that they must master.

Various studies suggest that people employ studying strategies some of which mirror “best practices” and some of which do not (see Rawson & Dunlosky, 2011; Roediger, Putnam, & Smith, 2011). For example, Kornell and Bjork (2007) found that students generally (a) do not employ spacing (or distributed practice), even though it is potentially beneficial, and (b) do not appreciate the improvements in learning that can occur from self-testing. McCabe (2011) examined undergraduate’s metacognitive knowledge about factors that cognitive research has shown to affect learning (e.g., self-testing,

* Corresponding author at: Department of Psychology, Florida International University, University Park, Miami, FL 33199, USA. Tel.: +1 305 348 4025.
E-mail address: bennett.schwartz@fiu.edu (B.L. Schwartz).

2211-3681/$ – see front matter © 2012 Society of Applied Research in Memory and Cognition. Published by Elsevier Inc.All rights reserved.
http://dx.doi.org/10.1016/j.jarmac.2012.06.002

cognitive load, generation effects). She found that, in most cases, students endorsed the wrong situation as more productive and efficient for learning. Providing students with specific metacognitive knowledge about variables that affect memory, such as the spacing effect and the self-testing effect, may allow students to improve learning efficiency.

1.2. Judgments of learning (JOLs)

JOLs are made during learning and are estimates concerning the likelihood of later remembering a target item (Rhodes & Tauber, 2011). Although metacognitive knowledge plays a role in JOLs, they are considered a measure of metacognitive experience. In experimental research, a JOL consists of a judgment concerning a word, a word-pair, a photograph of a face, or brief passage of text. JOLs are made either in the presence of just the cue or both the cue and the target. Moreover, they are made immediately after a study episode or later, or after other items have been studied, as in delayed JOLs (see Dunlosky & Nelson, 1992; Rhodes & Tauber, 2011 for a discussion of the "delayed JOL effect").

There is a consensus that metacognitive experiences, including JOLs, are based on heuristics (Dunlosky & Metcalfe, 2009; Koriat, 1997; Tauber & Rhodes, 2010). People use cues (e.g., fluency, size of material, amount of material) to non-consciously infer the strength of JOLs. Many cues lead to remarkably high accuracy in predicting performance, especially when the JOL is made at a delay after study and without the cue word present (Dunlosky & Nelson, 1992). This means that the heuristic mechanism successfully monitors the ability of our memory systems to retain information. Heuristic systems work because, under normal circumstances, the heuristics correlate with what they are intended to monitor, namely retrieval. The cues are correlated with and therefore predictive of the learning and remembering processes that metamemory monitors. We assert that the heuristic/inferential nature of metamemory judgments leaves open the possibility that people can focus on more diagnostic cues, leading to better monitoring and hence better control of memory.

1.3. The distinction between relative and absolute accuracy

It is critical to differentiate between relative and absolute accuracy (Benjamin & Diaz, 2008; Schwartz & Metcalfe, 1994). Relative accuracy (resolution) concerns the ability of JOLs to discriminate among items, whereas absolute accuracy (calibration) concerns the ability of JOLs to predict overall performance level. These are two different dimensions of accuracy that, at least in theory, may not be correlated with each other. For example, if a participant gives only 90% and 100% ratings to items but then scores 20% on the test, that participant will have low calibration. However, if the participant remembers more of the items given 100% than items given 90%, resolution will be relatively high. All that is needed for resolution is that higher judgments correspond to better performance than lower judgments. It is also possible for someone to show good absolute accuracy but show poor relative accuracy. A participant may think she will remember the weaker items and not think she will remember the stronger items, but her overall mean JOL reflects her actual performance. We contend that in order for JOLs to be useful for the individual learner, JOLs need to be accurate in both the relative (resolution) and absolute (calibration) sense. JOLs must show relative accuracy – so students know which items are relatively more difficult than others. JOLs must also show absolute accuracy – so students will know if their overall level of preparedness matches their goals for performance. One of our themes is that JOLs tend to show good resolution but poor calibration.

1.4. A theoretical framework

Koriat (1997) argued that JOLs are heuristic judgments based on three classes of cues: intrinsic, extrinsic, and mnemonic. Intrinsic cues are properties of the to-be-learned material, such as size, font, sound of words, vividness of a picture, concreteness of words, and imagability. Extrinsic cues are based on properties of the learning environment, such as instructions to use imagery, rote encoding or other study techniques. Extrinsic cues include the amount of time allotted to study items and the spacing of items across study. Finally, mnemonic cues are based on the person's subjective experience, such as the fluency with which items are perceived during study and the fluency with which items are retrieved at test.

Koriat (1997) argued that people attend more to intrinsic cues than to extrinsic cues during study. Indeed, many illusions that affect JOLs arise from an unjustified attention to intrinsic characteristics and a lack of attention to extrinsic characteristics (Kornell, Rhodes, Castel, & Tauber, 2011). Koriat's framework suggests several critical applications to everyday learning. People pay more attention to the intrinsic characteristics than they should, and should pay more attention to extrinsic characteristics than they do.

Numerous studies have supported Koriat's (1997) prediction that intrinsic cues will influence JOLs more than extrinsic cues do. For example, Carroll, Nelson, and Kirwan (1997) compared the relative contribution of overlearning (extrinsic) and cue-target relatedness (intrinsic) on memory and JOLs. Carroll et al.'s participants overlearned the unrelated pairs, but only studied the related pairs to criterion, thus putting the two variables into conflict with each other for subsequent JOLs. Under these circumstances, overlearning of unrelated pairs produced better recall than criterion learning of related pairs (53–42%). However, the JOLs predicted the opposite, that the related pairs would be remembered better despite the extra study for the unrelated pairs. In keeping with Koriat's prediction, the intrinsic cues influenced JOLs more than the extrinsic cues, which led to better retrieval (see Carroll, 2008).

Rhodes and Castel (2008a) demonstrated that JOLs are subject to an illusory effect based on perceptual manipulations. They asked participants to make JOLs on word pairs, some of which were in smaller fonts and others in larger fonts. The larger fonts created an illusion of better learning, as manifested by larger JOLs. Rhodes and Castel (2009) showed the related finding that JOLs are influenced by how loudly stimuli are presented. Word pairs that were read more loudly to participants were given higher JOLs than those that were presented more softly. It is possible to interpret these data in terms of intrinsic cues, as the participants were relying on the size or loudness of the words to determine their JOLs. This makes sense if the participants anticipate that the larger or louder stimuli were more memorable (see Rhodes & Castel, 2008a). However, these findings suggest that the relative ease of processing of the larger or louder stimuli resulted in the experience that these items were easier to learn. This would classify these data as supporting the role of mnemonic cues. Either way, the higher JOLs were illusory in that they were not mirrored by better performance (see Table 1).

Other researchers have proposed other heuristics to account for mechanisms underlying JOLs in addition to Koriat's triad of intrinsic, extrinsic, and mnemonic. Finn and Metcalfe (2007) suggested that one heuristic for making JOLs is the memory or belief of whether the target item has been successfully recalled in the past, which they labeled the "memory-for-past-test" heuristic (also see Ariel & Dunlosky, 2011). Having retrieved an item in the past is the basis of judging an item to be retrievable in the future. To demonstrate this, Finn and Metcalfe gave participants repeated study attempts and repeated JOLs. In fact, Trial 2 JOLs were better predicted by Trial 1 recall than they were by Trial 2 recall. In Koriat's (1997) framework, it is likely that memory for past test would qualify as a mnemonic cue, as participants base their JOL on their experience of remembering that they recalled the item earlier.

1.5. Common metamemory flaws

One goal of educators and cognitive psychologists is to bring the flaws and inaccuracies of metacognitive judgments to the attention of learners so that they can compensate for errant metacognitive experiences. The next section will examine several critical flaws in our ability to make accurate JOLs. Table 1 shows variables that affect memory performance but not JOLs, variables that do not affect memory performance but influence JOLs, and those few variables that affect memory performance and JOLs in the same manner.

2. Stability bias: we underestimate both learning and forgetting

One of the most dangerous errors in human metacognition is the stability bias, the tendency to expect memory to remain constant over time (Kornell, 2011). As such, we underestimate future forgetting and overestimate future remembering based on what we can retrieve now. Stability bias has two components: we discount the advantages of future studying, and we discount the likelihood that we will forget what we remember now. This may appear counterintuitive because at the level of belief (or metamemory knowledge), we know we will forget much and that studying promotes learning, but at the micro-level of metacognitive experience of individual items, the stability bias is strong (Kornell, 2011).

Kornell and Bjork (2009) conducted 12 experiments related to stability bias. Participants were instructed to study some paired associates one time and other paired associates four times, and then they took a cued-recall test. Not surprisingly, recall increased across the learning trials with more trials leading to better recall performance. But JOLs did not track this; JOLs given on individual paired associates did not increase across the learning sessions. Moreover, in some of the experiments, in which the advantages of study were made salient, JOLs did increase, but not enough to match the rate of learning, a phenomenon known as the underconfidence-with-practice effect (Koriat, Ma'ayan, Sheffer, & Bjork, 2006). In another study, Kornell et al. (2011) showed no interaction between the font-size effect and the anticipated study opportunities. Font size influenced JOLs, but not memory, whereas study opportunities affected memory but not JOLs (also see McDonough & Gallo, 2012). These data are consistent with Koriat (1997), as the intrinsic variable (font size) influenced JOLs even though it was non-diagnostic, but the extrinsic variable (study repetitions) did not influence JOLs even when it could have been diagnostic.

Table 1
Variables and their effect on memory and JOLs.

Name of effect	Effect on memory	Effect on metamemory
	Affects memory but not metamemory	
Studying multiple trials[a]	Improves memory	No effect on JOLs
Retention interval[b]	Decreases memory performance	No effect on JOLs
Retrieval Practice/Self-testing[c,d]	Improves memory performance	No effect on JOLs
Serial position effects[e]	Creates primacy and recency effects	No effect on JOLs
Inverting words[f]	Inverted words are recalled better than upright words	No effect on JOLs
Overlearning[g]	Overlearning strengthens memory	Small effect on JOLs
Part-set cueing[h]	Non-restudied items show reduced memory relative to control s who do not restudy	No effect on JOLs
Spacing[i,j]	Spaced items recalled better than massed items	Massed items are given higher JOLs than spaced items
	Affects memory more strongly than metamemory	
List length[k]	Items in shorter lists are recalled better than items in longer lists	JOLs track effect, but still strong overconfidence for longer lists
	Does not affect memory but does affect metamemory	
Size of stimuli[l,m,n]	Does not affect memory	Larger stimuli get higher JOLs
Loudness of presentation[o]	Does not affect memory	Louder stimuli get higher JOLs
	Memory and metamemory match	
Past test performance[p]	Retrieved items likely to be retrieved again.	JOLs are higher for initially retrieved items than items not retrieved.
Directed forgetting[q]	Directed forgetting produces lower recall	JOLs are lower in directed forgetting condition
Relatedness[g]	Related items recalled better than unrelated items	JOLs are higher for related items than for unrelated items

[a] Kornell and Bjork (2009).
[b] Koriat et al. (2004).
[c] Karpicke (2009).
[d] Kornell and Son (2009).
[e] Castel (2008).
[f] Sungkhasetee et al. (2011).
[g] Carroll et al. (1997).
[h] Rhodes and Castel (2008b).
[i] Zechmeister and Shaughnessy (1980).
[j] Son (2004).
[k] Tauber and Rhodes (2010).
[l] Rhodes and Castel (2008a).
[m] Kornell et al. (2011).
[n] McDonough and Gallo (2012).
[o] Rhodes and Castel (2009).
[p] Finn and Metcalfe (2007).
[q] Friedman and Castel (2011).

If participants do not see the advantage of further study, they may choose not to study even though continued study will improve performance. It leads to the proverbial complaint among students that they "studied really hard," but "did not do well on the test." This experience may be, in part, fueled by the failure to anticipate the benefits of continued study. "Studying hard" does not mean studying enough.

Stability bias also means that we underestimate forgetting, that is, participants do not anticipate how much information they will forget over time. For example, Koriat, Bjork, Sheffer, and Bar (2004) examined people's abilities to anticipate forgetting. Two groups of participants studied paired associates and then made either JOLs or judgments of forgetting (see Finn, 2008 for a discussion of judgments of forgetting). One group's JOLs were directed at a test to be taken in 10 min whereas the other groups JOLs were directed at a test to be taken 1 week later. Despite the difference in the anticipated retention interval, the JOLs were identical, even though those in the 1-week retention interval recalled fewer items than those in the 10-min retention interval. The JOLs did not predict the certainty of forgetting, especially over longer retention intervals. When participants made judgments of forgetting rather than JOLs, their stability bias was reduced but not eliminated. The judgments of forgetting increased for the 1-week condition relative to JOLs, but not enough to accurately reflect the forgetting that would occur. Focusing people on forgetting helps, but does not eliminate the stability bias. In sum, stability bias demonstrates both a failure to appreciate leaning and a failure to anticipate forgetting.

3. We underestimate the advantages of self-testing

Retrieval practice, also known as self-testing, is a powerful memory enhancer. To practice retrieval means to learn by recalling information from memory (e.g., learn by taking tests). Retrieval practice can lead to more learning than re-studying (for reviews, see Roediger, 2009; Roediger & Butler, 2011). In some cases, given equal total time, retrieval practice can result in over 80% better retention than restudying the items (see Roediger et al., 2011).

Roediger and Karpicke (2006a, 2006b) asked participants to read short prose passages concerning scientific information. One group of participants was given four opportunities to read a passage. A second group was given one reading period, but then had three opportunities to practice recalling the passage. Participants tested immediately did better if they had been in the re-reading condition. But one week later, the outcome reversed: the retrieval practice group outperformed the restudy group by more than 50%. Thus, employing retrieval practice as part of their study regimen could benefit people tremendously (Roediger, 2009; Roediger & Butler, 2011).

Nevertheless, recent studies have shown that people do not recognize the benefits of tests (e.g., Karpicke, 2009; Kornell & Son, 2009). This failure to recognize the benefits of self-testing appears in both JOLs and metamemory control decisions about how to study (see Table 1). For example, Karpicke (2009) required students to learn Swahili–English word pairs. After an initial study period and an initial test period, participants made JOLs for the word pairs. They were then given subsequent test periods. After each item had been tested, the participants could choose to drop items from future study, to study items, or to continue to self-test. Participants then had a final study/review session. They returned a week later for a final test. Consistent with the self-testing effect, final recall was best for the self-tested items. However, the JOLs predicted a different pattern. Participants anticipated remembering more of the studied items than either the tested items or the dropped items. Thus, there was a dissociation between memory and metamemory. Self-testing improves memory, but it is not reflected in the JOLs.

Even though students may not know that retrieval practice helps learning, they use it to evaluate the state of their memory. Kornell and Son (2009) replicated the basic effect: JOLs were higher following a study trial than following a test trial. However, when given a choice to re-study or test, more participants chose to test than study, leading to a more adaptive strategy. Post-experimental questionnaires suggest that participants made the choice to test because they wanted to evaluate their learning rather than improve the efficiency of their learning. Knowing which items one has learned or not learned is critical in using metacognition to improve learning efficiency. Testing oneself allows one to identify those items that require more study, which can then direct the choice of items in subsequent study (or ideally, subsequent self-testing). Nevertheless, it reflects that the participants were not aware that self-testing is good for learning as well as evaluating that learning (see Roediger et al., 2011).

To summarize, self-testing is an effective manner of learning, producing large advantages over re-reading (Roediger & Karpicke, 2006a, 2006b). However, the heuristics that influence JOLs are not sensitive to this factor, and JOLs tend to be lower for self-testing trials than for restudy trials. When participants are given control over learning in experiments, they may adopt self-testing but not to improve the efficiency of their learning. The intuition that restudy leads to better retrieval than self-testing must be countered with metacognitive knowledge that the opposite is true.

4. Judgments of learning mispredict the spacing effect

Students know that massing their study is inefficient, but they mass anyway because they do not organize their time well enough. Students generally shun distributed study because they find it difficult to use, and most students study on a "crisis" basis (Kornell & Bjork, 2007). In fact, the bias against spacing runs deeper. Zechmeister and Shaughnessy (1980) showed that participants give higher JOLs for free recall items that they massed than those items that they spaced. Thus, the distributed study led to better recall, but the JOLs were higher for the massed items (see Table 1).

The relation between JOLs and distributed study is more complicated because students do choose to distribute study under some conditions (Benjamin & Bird, 2006; Son, 2004, 2010; Son & Kornell, 2009; Toppino, Cohen, Davis, & Moors, 2009). When participants are given control over decisions to mass or space study, they do so based on the difficulty of the items (Son, 2010; Toppino & Cohen, 2010). For example, Son (2004) found that items given low JOLs were more likely to be followed by massing decisions. However, higher JOLs were correlated with greater spacing decisions, allowing for better retention of those items. Although a more adaptive strategy is to space all items, Son found a relation between JOLs and control decisions about spacing. Son (2010) found a similar pattern except that maximum spacing occurred at intermediate levels of JOLs, and the highest level of JOLs showed a return to massing. Son argued that these strategies are adaptive. Indeed, when people were forced to space, their final recall was poorer than when they used their own distribution of spacing and massing. Interestingly, when incentives are high, participants tend to space rather than mass, suggesting that participants know that spacing has learning advantages (Toppino & Cohen, 2010). Thus, participants do anticipate the benefits of spacing, and when given control over learning, they will employ it in some situations.

5. Other memory variables for which metamemory is insensitive

Primacy and recency. Castel (2008) asked participants to make JOLs about single words presented for later free recall. As expected, free recall of single-item lists yielded robust primacy and recency effects. However, JOLs given by participants anticipated neither the primacy nor the recency effect. Castel found that people could become sensitive to the primacy and recency effects. In the experiment, additional information, such as explicitly providing information (that is, giving people metacognitive knowledge) about serial position, caused participants to give JOLs that reflect the serial position curve. When participants made JOLs before they saw the target and could use only serial position as a cue, their JOLs demonstrated primacy and recency effects. Castel's data support the idea that intrinsic properties of items drown out extrinsic properties of test situations, and they support our view that explicit metacognitive knowledge can offset biases in metacognitive experience.

Part-set cueing. Part-set cueing refers to the phenomenon in which cueing part of a list of items interferes with retrieving the rest of the list items (Brown, 1968; Slamecka, 1968). Rhodes and Castel (2008b) found that part-set cueing resulted in a decrease in recall relative to a free recall condition, but that predictions of recall were identical across both conditions. This was true for both a semantic (Experiment 1) and episodic (Experiment 2) version of part-set cueing. In one version of part-set cueing employed by Rhodes and Castel (Experiment 1), participants in the part-set condition were allowed to view a list of 25 of the United States and then were asked to retrieve as many of the remaining states as possible. In the control condition, participants were simply asked to recall as many states as possible. Participants in the control condition (free recall) recalled more of the remaining 25 states than did the participants who were given the partial list (see Basden, Basden, & Morales, 2003; Brown, 1968). However, their predictions of how many items they would remember were the same across conditions. That is, metamemory judgments were insensitive to the effects of part-set cueing. With practice, however, participants became aware of the part-set cueing effect and lowered their JOLs for non-reviewed items. Here we see another variable that affects memory, but not JOLs, until explicit attention was brought to it.

Desirable difficulty. Participants may fail to understand the principle of desirable difficulty, that is, tasks requiring effort create better retention than those that do not. Sungkhasetee, Friedman, and Castel (2011) showed that presenting inverted words led to better recall of those words than words presented normally. However, participants' JOLs did not reflect this memory effect. It is likely that the inverted words were more difficult to process, thus leading to better memory because of the extra effort but received lower JOLs because they were processed more slowly. Thus, JOLs do not reflect desirable difficulties, a component of improving studying efficiency (see Metcalfe, 2011).

List length. JOLs mirror the pattern of recall produced by differences in list length, but the participants do not anticipate the extent to which list length affects recall. The list length effect refers to the observation that the probability of recall for individual items within a longer list will be less for individual items within a shorter list. Tauber and Rhodes (2010) examined if JOLs would be sensitive to the effect of list length. Koriat's (1997) model predicts that participants attend to intrinsic properties of items before they attend to the extrinsic properties of the situation. This suggests that each item will be judged by its own difficulty, but the list length will be mostly ignored. In fact, that is what Tauber and Rhodes found. Percent recall was superior for individual items in the shorter list. The basic pattern was also found for JOLs, but participants were still more overconfident for longer lists than for shorter lists. Participants were not able to compensate for the overconfidence when given instructions about the difficulty of long lists. Only when participants received specific instructions followed by practice did their overconfidence diminish. However, it is worth pointing out the effects of instructions and training given here are similar to Rhodes and Castel's (2008b) finding that practice made people aware of part-set cueing effect. Practice enhances awareness of factors that are involved in memory retrieval, and this awareness can potentially inform future studying decisions.

Directed forgetting. In contrast with other variables presented here, JOLs are sensitive to at least one extrinsic variable that affects memory. In a typical experiment using directed forgetting, the experimenter instructs participants to ignore or forget some items, usually with a cover story that they were accidentally given the wrong list. Research shows that participant's ability to recall to-be-forgotten items is impaired (Bjork, 1970). Friedman and Castel replicated this effect, but demonstrated that JOLs were sensitive to the manipulation. JOLs were lower for to-be-forgotten items than they were for to-be-remembered items.

What can we conclude about how JOLs are influenced by memory effects and the promise of improving learning by attending more to our own JOLs? JOLs tend to show strong correlations with performance within conditions (strong resolution or relative accuracy). Thus, JOLs may be potentially useful in guiding self-directed study if we use them judiciously. However, JOLs are insensitive to several variables that affect the overall level of memory, thus leading to problems with calibration or absolute accuracy. JOLs tend to reflect overconfidence in that the mean JOL tends to be higher than the actual level of recall. We reviewed how JOLs are insensitive to effects such as inverting words, part-set cueing, retention interval, retrieval practice, serial-position effects, and studying multiple trials. Instructors can help their students' learning by alerting them to these factors. Trained instructors can advise their students in how to overcome these shortcomings in metamemory: do not trust your confidence, engage in retrieval practice, and study past the point that you think you have mastered the materials.

If the heuristics that underlie JOLs miss so many memory variables, why then do JOLs show relative accuracy? That is, why are people are more likely to recall those items given higher JOLs than those given lower JOLs (Dunlosky & Metcalfe, 2009; Rhodes & Tauber, 2011)? It is likely that relative accuracy is the result of the efficacy of intrinsic cues. Overconfidence, on the other hand, occurs because people are insensitive to some variables, usually extrinsic characteristics that affect memory performance. Thus, to summarize, relative accuracy is a function of good sensitivity to intrinsic characteristics, but overconfidence is produced by poor sensitivity to extrinsic characteristics.

Overconfidence. One of the greatest impediments to learning is the illusion that material has already been learned when it has not. Students are overconfident with respect to real test performance, just as they are overconfident in the lab. For example, Hacker, Bol, Horgan, and Rakow (2000) asked students in an undergraduate educational psychology class to predict their performance before they began an actual test. Students also gave confidence estimates for the exam after they were finished. Hacker et al. then divided the sample into five groups based on test performance, each group corresponding to one quintile from highest performers to lowest performers. Based on their pre-test predictions, the top quintile showed no overconfidence, and in fact, showed slight underconfidence. However, the lower 80% of students were overconfident, with the lowest performers being the most overconfident (nearly 30%). This pattern was also evident in the post-dictions: the top 20% were well-calibrated, but the remaining 80% were overconfident. Hacker et al.'s data suggest that most students do not adequately anticipate their performance nor can they assess how much material they have mastered. This study was replicated and extended by Miller and Geraci (2011), who found that even with feedback, students' predictions remained overconfident. Even when students received "extra credit" incentives for accuracy, they maintained overconfidence in their exam predictions (for a review, see Hacker, Bol, & Keener, 2008).

5.1. How can judgments be used to improve the nature of control?

What decisions do people make when they choose to study and can we improve upon them? Students much decide what to study, how long to study, and how to study (Kornell & Bjork, 2007). We claim that people can improve their studying efficiency by understanding the strengths and weaknesses of their own JOLs. In keeping with the themes of this paper, students can be instructed to attend to their JOLs when determining which items to study, but to mistrust their JOLs when it comes to how long to study.

One model that has been offered to understand metacognitive control as well as guide student learning is the Region of Proximal Learning theory. This theory states that students should study the unrecalled items with the highest JOLs, that is, the subjectively easiest items that are not yet mastered (Metcalfe, 2002, 2009, 2011; Kornell & Metcalfe, 2006). Metcalfe (2002) claims that students should not commit time to reviewing easy items (which they may know at test anyway) or fussing over difficult items. Reviewing easy items wastes time because these are already known, and reviewing difficult items implies that resources may be wasted on items that may not be learned despite the study. Because JOLs show high resolution, JOLs can be used to differentiate difficult and easy items.

However, JOLs are not well calibrated, which can potentially lead to problems with the region-of-proximal-learning framework. The region-of-proximal-learning perspective capitalizes on people's ability to accurately use intrinsic cues to achieve good JOL resolution. However, the region-of-proximal-learning perspective does not, by itself, offer a response to the failure of most people to anticipate the effects of extrinsic characteristics, nor does it provide a buttress against the problems inherent with overconfidence. To counter the effects of stability bias (overestimating remembering and underestimating forgetting) and overconfidence, learners should continue to study items that they gave high JOLs to, learned and not yet learned, in contrast to the original predictions of the region of proximal learning.

Can we instruct students to trust their JOLs to discriminate among items (resolution), but not trust JOLs to make overall assessments of how much has been learned (calibration)? Research suggests that students do not easily change their metacognitive knowledge. McCabe (2011) compared five groups of undergraduate students who were given different levels of instruction about metacognition varying from no instruction to a "seminar" on cognition and education in which they read the relevant articles and had in-depth discussion on the topics. Only the group that had enrolled in the seminar benefitted from instruction. McCabe's finding suggests that changing declarative knowledge about learning strategies is difficult, but not impossible. Dunlosky and Rawson (in press) also found that extensive feedback to participants can reduce overconfidence in metacognitive judgments, also suggesting that training and feedback can positively affect people's learning strategies by improving calibration.

6. Conclusions

In real-world learning, metacognitive experiences dictate which items are felt to be difficult or easy (Efklides, 2011). In the current paper, we reviewed factors that influence JOLs, and we discussed when JOLs may be accurate and what contributes to inaccuracy. Much still needs to be examined to determine how we can best use metamemory toward improving learning efficiency. Real-world learning may involve learning more complicated material than paired associates. In many circumstances what matters may be an understanding of the material and not necessarily memorizing it. Nonetheless, we think there is much of applied value to be gleaned from research on JOLs.

Koriat's (1997) theory of the etiology of JOLs accounts for the relative accuracy of JOLs; they are based on intrinsic features, which reflect item difficulty. Koriat's theory accounts for the poor absolute accuracy, the tendency towards overconfidence. This occurs because people do not adequately account for extrinsic or mnemonic factors. In fact, Table 1 shows several extrinsic variables that produce JOLs that run counter to their effects on memory (see Table 1).

What can we tell our students so that they can better use their metamemory to direct study? In some situations, metamemory experiences are valid. We are good at assessing the items that are difficult and the items that are not, and our intuition when making JOLs can be trusted here. However, our JOLs are overconfident and in many cases fail to reflect variables that scientific research finds can improve memory efficiency. So, in choosing how long to study, when to study, and what methods to study, we are often best not trusting our intuition and instead following the guidelines set forth here and in other empirically based programs of improving learning efficiency.

References

Ariel, R., & Dunlosky, J. (2011). The sensitivity of judgment-of-learning resolution to past test performance, new learning, and forgetting. *Memory & Cognition*, *39*, 171–184.

Ariel, R., Dunlosky, J., & Bailey, H. (2009). Agenda-based regulation of study-time allocation: When agendas override item-based monitoring. *Journal of Experimental Psychology: General*, *138*, 432–447.

Basden, B. H., Basden, D. R., & Morales, E. (2003). The role of retrieval practice in directed forgetting. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 389–397.

Benjamin, A. S., & Bird, R. (2006). Metacognitive control of the spacing of study repetitions. *Journal of Memory and Language*, *55*, 126–137.

Benjamin, A. S., & Diaz, M. (2008). Measurement of relative metamnemonic accuracy. In J. Dunlosky, & R. A. Bjork (Eds.), *Handbook of Memory and Metamemory* (pp. 73–94). New York, New York: Psychology Press.

Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior*, *9*, 255–268.

Brown, J. (1968). Reciprocal facilitation and impairment of free recall. *Psychonomic Science*, *10*, 41–42.

Carroll, M. (2008). Metacognition in the classroom. In J. Dunlosky, & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 411–427). New York: Psychology Press.

Carroll, M., Nelson, T. O., & Kirwan, A. (1997). Trade-off of semantic relatedness and degree of overlearning: Differential effect on metamemory and long-term retention. *Acta Psychologica*, *95*, 239–253.

Castel, A. D. (2008). Metacognition and learning about primacy and recency effects in free recall: The utilization of intrinsic and extrinsic cues when making judgments of learning. *Memory & Cognition*, *36*, 429–437.

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Thousand Oaks, CA: Sage Publications, Inc.

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*, 373–380.

Dunlosky, J., & Rawson, K. A. Overconfidence produces underachievement: Inaccurate self-evaluations undermine students' learning and retention. *Learning and Instruction*, in press.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process. *Educational Research Review*, *1*, 3–14.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, *46*, 6–25.

Finn, B. (2008). Framing effects on metacognitive monitoring and control. *Memory & Cognition*, *36*, 813–821.

Finn, B., & Metcalfe, J. (2007). The role of memory for past test in the underconfidence with practice effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *33*, 238–244.

Friedman, M. C., & Castel, A. D. (2011). Are we aware of our ability to forget? Metacognitive predictions of directed forgetting. *Memory & Cognition*, *39*, 1448–1456.

Hacker, D. J., Bol, L., & Keener, M. C. (2008). Metacognition in education: A focus on calibration. In J. Dunlosky, & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 429–455). New York: Psychology Press.

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test predictions and performance in a classroom context. *Journal of Educational Psychology*, *92*, 160–170.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General*, *138*, 469–486.

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., Bjork, R. A., Sheffer, L., & Bar, S. (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., 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.

Kornell, N. (2011). Failing to predict future chances in memory: A stability bias yields long-term overconfidence. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 365–386). New York, NY: Psychology Press.

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, *14*, 219–224.

Kornell, N., & Bjork, R. A. (2009). A stability bias in human memory: Overestimating remembering and underestimating learning. *Journal of Experimental Psychology: General*, *138*, 449–468.

Kornell, N., & Metcalfe, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*, 609–622.

Kornell, N., & Son, L. K. (2009). Learners' choices and beliefs about self-testing. *Memory*, *17*, 493–501.

Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S. K. (2011). The ease of processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. *Psychological Science*, *22*, 787–794.

McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition*, *39*, 462–476.

McDonough, I. M., & Gallo, D. A. (2012). Illusory expectations can affect retrieval-monitoring accuracy. *Journal of Experimental Psychology.*, *38*, 391–404.

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning. *Journal of Experimental Psychology: General*, *131*, 349–363.

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science*, *18*, 159–163.

Metcalfe, J. (2011). Desirable difficulties and studying in the region of proximal learning. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 259–276). New York: Psychology Press.

Miller, T. M., & Geraci, L. (2011). Training metacognition in the classroom: The influence of incentives and feedback on exam predictions. *Metacognition and Learning*, *6*, 303–314.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 125–173). New York: Academic Press.

Rawson, K. A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General*, *140*, 283–302.

Rhodes, M. G., & Castel, A. D. (2008a). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*, 615–625.

Rhodes, M. G., & Castel, A. D. (2008b). Metacognition and part-set cuing: Can interference be predicted at retrieval? *Memory & Cognition*, *36*, 1429–1438.

Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review*, *16*, 550–554.

Rhodes, M. G., & Tauber, S. K. (2011). The influence of delaying judgments of learning (JOLs) on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin*, *137*, 131–148.

Roediger, H. L. III. (2009). The critical role of retrieval in enhancing long-term memory: From the laboratory to the classroom. *Keynote address at the 50th annual psychonomics meeting*. Boston, MA. Available at http://bethereglobal.com/content/442_psychonomics/play%20flash.htm.

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*, 20–27.

Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255.

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*, 181–210.

Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. In J. Mestre, & B. Ross (Eds.), *Psychology of learning and motivation: Cognition in education* (pp. 1–36). Oxford: Elsevier.

Schwartz, B. L., & Metcalfe, J. (1994). Methodological problems and pitfalls in the study of human metacognition. In J. Metcalfe, & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 93–114).

Slamecka, N. J. (1968). A methodological analysis of shift paradigms in human discrimination learning. *Psychological Bulletin*, *69*, 423–438.

Son, L. K. (2004). Spacing one's study: evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 601–604.

Son, L. K. (2010). Metacognitive control and the spacing effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*, 255–262.

Son, L. K., & Kornell, N. (2009). Simultaneous decisions at study: Time allocation, ordering, and spacing. *Metacognition and Learning*, *4*, 237–248.

Sungkhasetee, V. W., Friedman, M. C., & Castel, A. D. (2011). Memory and metamemory for inverted words: Illusions of competency and desirable difficulties. *Psychonomic Bulletin & Review*, *18*, 973–978.

Tauber, S. K., & Rhodes, M. G. (2010). Are judgments of learning (JOLs) sensitive to the amount of material to-be-remembered? *Memory*, *18*, 351–362.

Toppino, T. C., & Cohen, M. S. (2010). Metacognitive control and spaced practice: Clarifying what people do and why. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*, 1480–1491.

Toppino, T. C., Cohen, M. S., Davis, M. L., & Moors, A. C. (2009). Metacognitive control over the distribution of practice: When is spacing preferred. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *35*, 1352–1358.

Zechmeister, E. B., & Shaughnessy, J. J. (1980). When you know that you know and when you think that you know but you don't. *Bulletin of the Psychonomic Society*, *15*, 41–44.