Learning and Instruction 22 (2012) 271–280

Contents lists available at SciVerse ScienceDirect

Learning and Instruction

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

Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention

John Dunlosky*, Katherine A. Rawson

Kent State University, Psychology Department, Kent, OH 44242, USA

ARTICLE INFO

Article history:
Received 2 February 2011
Received in revised form
19 August 2011
Accepted 21 August 2011

Keywords:
Judgment accuracy
Retention
Metacognition
Self regulation
Metacomprehension

ABSTRACT

The function of accurately monitoring one's own learning is to support effective control of study that enhances learning. Although this link between monitoring accuracy and learning is intuitively plausible and is assumed by general theories of self-regulated learning, it has not received a great deal of empirical scrutiny and no study to date has examined the link between monitoring accuracy and longer-term retention. Across two studies, college students paced their study of key-term definitions (e.g., "Proactive interference: Information already stored in memory interferes with the learning of new information"). After all definitions were studied, participants completed practice cued recall tests (e.g., "What is proactive interference?") in which they attempted to type the correct definition for each term. After each test trial, participants judged how much of their response was correct. These study-test-judgment trials continued until a definition was judged as correct three times. A final cued recall test occurred two days later. In Study 1, judgment accuracy was manipulated experimentally, and in Study 2, individual differences in accuracy were examined. In both studies, greater accuracy was associated with higher levels of retention, and this link could not be explained by differential feedback, effort during study, or trials to criterion. Results indicate that many students could benefit from interventions aimed at improving their skill at judging their learning.

1. Introduction

Given the amount of material that students must learn across a wide variety of courses, they should strive to achieve durable learning and also to use their time efficiently. To help foster efficient and durable learning, researchers have recommended the use of *accurate* monitoring to guide learning (e.g., Dunlosky, Hertzog, Kennedy, & Thiede, 2005; Pashler et al., 2007). One reason for this recommendation is that students who are overconfident in their evaluations of learning may fall short of their learning goals, whereas accurate evaluations of one's own learning can be used to more effectively guide further study (Rawson & Dunlosky, in press). Although intuitive, the evidence for this link between overconfidence and achievement is limited and may exaggerate the importance of accurate monitoring for effectively achieving durable retention. Before we describe the available evidence, we briefly discuss general theories of self-regulated learning, which predict that monitoring accuracy will be related to learning and retention. We then consider previous research in relationship to the present studies, which were primarily designed to answer the following question: Does overconfidence produce underachievement when students are learning key-term definitions?

1.1. Is better monitoring accuracy related to better acquisition and retention?

1.1.1. Theory of self-regulated learning

Metacognitive theory of self-regulated learning inextricably links monitoring to control (e.g., Boekaerts, 1997; Flavell, 1979; Koriat & Goldsmith, 1996; Metcalfe, 2009; Nelson & Narens, 1990; Winne & Hadwin, 1998). The relationship between monitoring and control was eloquently illustrated by Nelson and Narens (1990; for historical roots, see Miller, Galanter, & Pribram, 1960), who separated cognitive processes into two interrelated levels, called the meta-level and the object-level. Nelson and Narens (1990) proposed.

> two dominance relations, called 'control' and 'monitoring', which are defined in terms of the direction of the flow of information between the meta-level and the object-level…The basic notion underlying control—analogous to speaking into a telephone handset—is that the meta-level *modifies* the object-level, [and] the basic notion underlying monitoring—analogous

* Corresponding author. Tel.: +1 330 672 2207; fax: +1 330 672 3786.
E-mail address: jdunlosk@kent.edu (J. Dunlosky).

0959-4752/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.learninstruc.2011.08.003

to listening to the telephone handset—is that the meta-level is *informed* by the object-level (pp. 126-127; italics in original).

In this model, the output from control processes (from the object-level) informs monitoring (cf. Koriat, Ma'ayan, & Nussinson, 2006), and most important here, monitoring in turn influences control decisions.

Without this latter link, in which a learner's assessments of ongoing learning are used to make decisions about what and how to study, metacognitive monitoring would be inert—an epiphenomenon perhaps worthy of discussion by philosophers of the mind but not relevant to understanding how human thought and action is controlled (Lieberman, 1979; Nelson, 1996). Fortunately, correlational evidence (e.g., Hines, Touron, & Hertzog, 2009) and experimental evidence (e.g., Metcalfe & Finn, 2008; Thiede & Dunlosky, 1999) converge on the conclusion that learners' monitoring does influence control decisions about what to study. Whereas the influence of monitoring on the control of study has been firmly established, the present studies were designed to evaluate the under-explored question of whether the *accuracy* of monitoring is related to better retention. It is critical to not conflate monitoring with monitoring accuracy. That is, many studies have demonstrated that learners of all ages and abilities use *monitoring* to guide their study (for reviews, see Dunlosky& Ariel, 2011; Son & Metcalfe, 2000), but the key issue here is whether the *accuracy* of learner's monitoring is related to better acquisition and retention.

Because learners use monitoring to efficiently obtain their learning goals (for reviews, see Benjamin, 2007; Castel, 2007; Dunlosky & Ariel, 2011), better monitoring accuracy is expected to be related to more effective learning and higher levels of retention. Although this *accuracy-influences-memory* (AIM) hypothesis appears relatively straightforward, it relies on two auxiliary assumptions. First, it assumes that monitoring is used to control learning in a relatively effective manner (Thiede, 1999), because even if learners' monitoring were perfectly accurate, if they did not (or could not) appropriately use their monitoring to control learning across items (Son & Sethi, 2006), then better accuracy will not benefit retention (Nelson, Dunlosky, Graf, & Narens, 1994). The second assumption is a corollary of the first; namely, an association between monitoring accuracy and learning may not be observed if learners differ in how they use monitoring to control study time—for example, a learner who monitors accurately but uses this monitoring inappropriately to control study may be outperformed by a learner with somewhat lower monitoring accuracy who makes more effective control decisions (Thiede, 1999). The relevance of these assumptions will be highlighted in the following review.

1.1.2. Evidence relevant to the AIM hypothesis

Only a few studies have investigated the link between monitoring accuracy and memory (Begg, Martin, & Needham, 1992; Bisanz, Vesonder, & Voss, 1978; Nietfeld, Cao, & Osborne, 2006; Thiede, 1999; Thiede, Anderson, & Therriault, 2003). Begg et al. (1992) had college students study word pairs (e.g., railroad – mother), and this study trial was followed by a delayed judgment of learning (JOL) in which the students predicted the chances of remembering the pair. These JOLs were prompted with either the cue only (railroad –?) or the cue-target pair (railroad – mother); cue-only prompts lead to substantially greater levels of monitoring accuracy. After making a JOL, the entire pair was presented for restudy for 4 s, and after all pairs had been restudied, a criterion test of paired-associate recall occurred. Despite the greater accuracy of cue-only JOLs, test performance was no different when participants made cue-only JOLs versus cue-target JOLs. This evidence led Begg et al. (1992) to conclude that "memory monitoring does not make a valuable contribution to memory" (p. 212). Note, however, that the first auxiliary assumption introduced above was not met: Because the restudy trials were brief and experimenter paced, participants could not use monitoring to effectively control learning. Similarly, Rhodes and Tauber (2011) conducted a meta-analysis on papers examining judgment accuracy, which demonstrated that conditions supporting higher levels of judgment accuracy were not related to greater memory performance. Most relevant here, the studies in their meta-analysis did not employ methods that allowed participants to use their judgments to control study (in fact, judgments were made *after* studying was completed), so these studies also fail to meet the first auxiliary assumption and hence are not relevant to evaluating the AIM hypothesis.

To sidestep such problems, Thiede (1999) used a procedure similar to the one used by Begg et al. (1992), but with two key changes. First, accuracy was not manipulated between groups; instead, the focus was on whether individual differences in accuracy were related to better learning. Second, all participants made delayed cue-only JOLs on the first study-test trial, and during each of the next four study-judgment-test trials, participants selected those items that they wanted to restudy. Thus, participants could use their monitoring to control their learning in a (potentially) effective manner. Thiede (1999) computed the *relative* accuracy of the delayed JOLs on Trial 1, which involved correlating each participant's JOLs with his or her own recall performance across items. As predicted by the AIM hypothesis, individual differences in relative accuracy on Trial 1 were positively correlated with subsequent boosts in learning across trials (i.e., participants who were better at discriminating items they knew from items they did not yet know profited more from the subsequent opportunities to select items for restudy). However, the relationship between accuracy and learning arose only on later trials and was relatively weak. One reason for the weak effect is that participants selected items for restudy, so that some benefits of better accuracy may have been offset by individual differences in how well monitoring was used to allocate study time (as per the second auxiliary assumption described above). Although investigating individual differences in how learners use monitoring to control study is important, holding control processes constant across participants can provide a more valid estimate of the influence of monitoring accuracy on retention. Accordingly, in the present studies, the same algorithm was used to control study decisions for all participants.

The present studies also critically differ from research by Thiede and his colleagues (Thiede, 1999; Thiede et al., 2003) in that they focused on relative accuracy (see also, Bisanz et al., 1978), whereas we are focusing on absolute accuracy. Relative accuracy refers to the degree to which judgments predict which items (relative to others) are more likely to be recalled on the criterion test and is measured by the aforementioned intra-individual correlations. By contrast, *absolute* accuracy refers to the extent to which people's judgments demonstrate overconfidence (judgments are more optimistic than actual performance) versus underconfidence (judgments are less optimistic than actual performance; for details about other conceptualizations of absolute accuracy, see Boekaerts & Rozendaal, 2010; Keren, 1991; Schraw, 2009). Arguably, absolute accuracy is critical for obtaining efficient and durable learning. The theoretical claim here is straightforward, but it has major implications for student success. Overconfidence should lead to premature termination of study and hence yield lower levels of learning during practice, which in turn will translate to lower levels of retention on the final test. Put differently, students who are more overconfident are not expected to learn as much during practice as compared to students who are better calibrated, and these initial differences in learning are expected to persist on the final retention test.

The relationship between absolute accuracy and performance has been investigated in two studies (Nietfeld et al., 2006; Shin,