Computers in Human Behavior 54 (2016) 73–82

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Full length article

Boosting metacomprehension accuracy in computer-supported learning: The role of judgment task and judgment scope

Julia Vössing [a,*], Christian Stamov-Roßnagel [b]

[a] Freie Universität Berlin, Germany
[b] Jacobs University, Germany

ARTICLE INFO

Article history:
Received 5 June 2015
Received in revised form
27 July 2015
Accepted 30 July 2015
Available online 11 August 2015

Keywords:
Metacomprehension accuracy
Judgments of learning
Computer-supported learning

ABSTRACT

Although learners' judgments of their own learning are crucial for self-regulated study, judgment accuracy tends to be low. To increase accuracy, we had participants make combined judgments. In Experiment 1, 247 participants studied a ten-chapter expository text. In the simple judgments group, participants after each chapter rated the likelihood of answering correctly a knowledge question on that chapter (judgment of learning; JOL). In the combined judgments group, participants rated text difficulty before making a JOL. No accuracy differences emerged between groups, but a comparison of early-chapter and late-chapter judgment magnitudes showed that the judgment manipulation had induced cognitive processing differences. In Experiment 2, we therefore manipulated judgment scope. Rather than predicting answers correct for an entire chapter, another 256 participants rated after each chapter the likelihood of answering correctly a question on a specific concept from that chapter. Both judgment accuracy and knowledge test scores were higher in the combined judgments group. Moreover, while judgment accuracy dropped to an insignificant level between early and late chapters in the simple judgments group, accuracy remained constant with combined judgments. We discuss implications for research into metacomprehension processes in computer-supported learning and for adaptive learner support based on judgment prompts.

1. Introduction

Learners' judgments how well they have mastered the learning material at hand are a crucial ingredient of self-regulated study and underlies their study decisions, e.g., in terms of time allocation (Metcalfe, 2009; Thiede, Anderson, & Therriault, 2003; Winne & Hadwin, 1998). To make those judgments, people continuously monitor their learning (Nelson & Leonesio, 1988) and use a variety of cues from that monitoring to control learning (Koriat, 1997; Metcalfe, 2009). Consistent with this assumption, numerous studies have shown that people's judgments of their learning predict actual learning performance (see Dunlosky & Lipko, 2007; Maki, 1998), reflecting *metacomprehension*, which refers to one's ability to judge one's learning and/or comprehension of text materials (Dunlosky & Lipko, 2007). Such *judgments of learning* are typically elicited by having learners gauge the likelihood they will correctly answer a test question on a passage of text they previously studied.

Early research mostly used paired-associates word lists and short passages of expository texts to study metacomprehension processes (see Dunlosky & Metcalfe, 2009). More recently, researchers have become interested in metacomprehension in computer-supported learning (e.g., Mengelkamp & Bannert, 2009) that may show systematic differences to metacomprehension with other materials. As one example, Ackerman and Goldsmith (2011) reported a screen inferiority effect such that learners showed worse performance in a knowledge test when learning on screen than on paper. The authors concluded that on-screen learning was characterized by less accurate predictions of performance and more erratic study-time regulation (see also Lauterman & Ackerman, 2014). Also, Serra and Dunlosky (2010) showed that epistemic beliefs specific to computer-supported learning may influence metacomprehension. Participants judged their learning to be better when studying text with diagrams, relative to studying text alone, and their test performance was indeed greater. However, in another group receiving pictures that did not support learning, JOL were as high as those in the effective pictures group. The data on epistemic beliefs

* Corresponding author.
E-mail address: julia.voessing@fu-berlin.de (J. Vössing).

http://dx.doi.org/10.1016/j.chb.2015.07.066
0747-5632/

showed that JOL were influenced by a "multimedia heuristic" supporting learners' belief that multimedia learning will invariably lead to good performance.

In sum, those findings indicate the need for a better understanding of the metacognitive regulation of computer-supported learning. Judgments of learning (JOL) might be a suitable tool to study metacomprehension processes, but judgment accuracy – defined as the correlation between JOL and learning performance – tends to be low (Dunlosky & Lipko, 2007; Lin & Zabrucky, 1998; Maki, 1998). Consequently, there have been calls for research to improve metacomprehension accuracy (e.g., Redford, Thiede, Wiley, & Griffin, 2012; Serra & Dunlosky, 2010). Such research may start from the finding that characteristics of the judgment task itself (e.g., Finn, 2008; Rhodes & Tauber, 2011) can constrain accuracy.

Against this backdrop, we conducted two experiments that sought to increase metacomprehension accuracy through modifying the judgment task. Participants studied a ten-chapter expository text and afterward took a knowledge test on that text. They made judgments of learning after each chapter. Rather than directly rating the likelihood of answers correct, participants in the experimental group made a judgment of text difficulty beforehand. Our research question was whether this manipulation would increase experience-based processing and thus JOL accuracy. We describe major influences on JOL accuracy in the next section before describing two experiments that tested the usefulness of two-step judgments. In the final section, we discuss implications and limitations of our approach together with some avenues for further research.

1.1. Judgment task, experience-based processing, and metacomprehension accuracy

To increase the accuracy of judgments of learning, several researchers manipulated the learning processes that precede judgments. Such pre-judgment measures include, for instance, rereading (e.g., Dunlosky & Rawson, 2005; Rawson, Dunlosky, & Thiede, 2000), summarizing (e.g., Anderson & Thiede, 2008; Thiede & Anderson, 2003), or generating keywords before making JOL (e.g., Dunlosky, Rawson, & Middleton, 2005a). More recent research has shown concept mapping to be an effective pre-judgment measure (e.g., Redford et al., 2012). Finally, Carpenter and Olson (2012) reduced overconfidence in word-pair learning through retrieval practice, i.e. by presenting word pairs three times and having participants make a JOL on each trial, or by simply warning participants not to be overconfident.

In this paper, we explore a manipulation of the judgment task itself to increase accuracy, guided by evidence that different cognitive processes seem to underlie different types of judgment tasks. For instance, comparing learners' standard judgments of learning with their judgments of text comprehension, Dunlosky, Rawson, and McDonald (2002) obtained higher ratings of comprehension than of performance, i.e. the two judgments differed in magnitude. Finn (2008) had her participants either indicate the likelihood they would remember or forget word pairs. Participants were more overconfident when indicating the likelihood they would remember, rather than forget word pairs. Ozuru, Kurby, and McNamara (2012) had their participants read an expository text and after each sentence make either a judgment of difficulty or a prediction of performance in an upcoming test. Text difficulty predicted judgments of difficulty, but not performance predictions. The latter were predicted by reading ability, which had no significant influence on difficulty judgments. Ozuru et al. (2012) concluded that engaging in different metacomprehension judgments may alter the information sources to which learners attend, which in turn influences metacognitive accuracy.

This raises the question if having participants make both kinds of judgments would influence accuracy. Son and Metcalfe's (2005) findings are consistent with that view. At study, participants were to memorize word pairs. At test, participants were given one word from a pair as a cue and had to recall the other word that served as target. One group was instructed to attempt to recall the target before making a JOL ("instructed JOL"), whereas another group made JOL alone. Son and Metcalfe (2005) found that for low-magnitude items, i.e. when learners indicated low probabilities of correct recall, JOL alone were made faster the lower their magnitudes were. Instructed low JOL, however, were made more slowly the lower their magnitudes were. This difference did not emerge with high-magnitude items. Son and Metcalfe (2005) concluded that JOL are based on a two-stage process (see Alter & Oppenheimer, 2009; Dunlosky, Serra, Matvey, & Rawson, 2005b) comprising a quick pre-retrieval stage at which stimulus familiarity is analyzed, and a second stage in which judgments of learning are based on retrieval fluency or on stimulus characteristics that become apparent after attempted retrieval. Important for our purpose, recall performance was significantly higher in the instructed JOL condition, this effect was most pronounced with low JOL. Apparently, the retrieval instruction had induced additional processing that increased recall.

On the background of Son and Metcalfe (2005) and Ozuru et al.'s (2012) findings, we had one group make two-step judgments. Participants estimated the difficulty of a given text passage prior to making a JOL (*combined JOL*). Control group participants made standard judgments of learning (*simple JOL*). The combined-JOL manipulation sought to have participants attend to characteristics of the text and thus promote *experience-based processing*. In this processing mode, learners base their judgments of learning on cues (e.g., familiar terms) from the ongoing processing of the learning material. Alternatively, judgments may rely on *theory-based processing* if participants respond on the basis of general beliefs (e.g., their ability self-perceptions), rather than assessing the knowledge they gain from a text (see Glenberg & Epstein, 1987). If under combined judgments learners rely more on their specific experiences with the text than on their global self-perceptions, we would expect them to more accurately gauge their learning, resulting in higher correlations between JOL and performance in the knowledge test. Consequently, we assume that.

Hypothesis 1. JOL accuracy will be higher with combined judgments, relative to simple judgments.

Furthermore, we expect differences between experience-based judgments and theory-based judgments when comparing judgments on earlier chapters to judgments on later chapters. This assumption is based on Zhao and Linderholm (2008) anchoring-and-adjustment account of metacomprehension accuracy. It assumes that because learners are uncertain about an upcoming test, the initial judgments in a series of JOL form an anchor, from which learners adjust away based on their experiences with current tasks. The more theory-based judgments are, the less they will be adjusted; conversely, adjustment should be higher for more experience-based judgments. Serra and Dunlosky (2010) found such adjustment effects when they compared participants' judgments before, during, and after studying the six chapters of an expository text. Participants' pre-study judgments were significantly higher in magnitude than both the averaged judgments after reading a chapter and the post-study judgments, which may reflect "participants' realization that the task will be more difficult than they had expected" (p. 704). By analogy, if simple and combined JOL differ in their degrees of experienced-based processing, we would