Learning and Instruction 46 (2016) 12–20

Contents lists available at ScienceDirect

Learning and Instruction

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

How accurately can students evaluate the quality of self-generated examples of declarative concepts? Not well, and feedback does not help☆

Amanda Zamary*, Katherine A. Rawson, John Dunlosky

Department of Psychological Sciences, Kent State University, USA

ARTICLE INFO

Article history:
Received 22 October 2015
Received in revised form
19 May 2016
Accepted 1 August 2016

Keywords:
Example generation
Metacognition
Monitoring
Overconfidence
Feedback

ABSTRACT

Students are commonly asked to learn declarative concepts in many courses. One strategy students report using involves generating concrete examples of abstract concepts. If students have difficulties evaluating the quality of their generated examples, then instructors will need to provide students with appropriate scaffolds or feedback to improve judgmentaccuracy. No prior research has investigated if students can accurately evaluate the quality of the examples they generate, which was the first aim of the current research. The second aim of this research was to investigate the extent to which providing feedback while students evaluate their generated examples can improve the accuracy of their example-quality judgments. In two experiments, students generated examples for declarative concepts from social psychology and then judged the quality of their examples. When making judgments, students received no feedback (in which they were only given the key term), full definition feedback (in which they were shown the definition of the declarative concept) or idea unit feedback (in which they first evaluated if they represented each idea unit of the definition within their example). Outcomes showed that students were overconfident when judging the quality of their examples, specifically for commission errors (i.e., examples that were entirely incorrect). Surprisingly, full definition and idea unit feedback did not help students improve the accuracy of their example-quality judgments. Thus, until scaffolds are discovered to reduce student overconfidence, instructors will need to assist in evaluating generated examples as students use this strategy to learn declarative concepts.

Published by Elsevier Ltd.

1. Introduction

Pick up a textbook from just about any content domain, turn to the end of any chapter, and more often than not, you will find a list of concept terms. Indeed, one of the most common kinds of information students are expected to learn in many courses are *declarative concepts.* Declarative concepts are abstract concepts denoted by key terms with short (usually 1–2 sentence) definitions of the meaning of the concept (Rawson, Thomas, & Jacoby, 2015). For example, a student in a social psychology course might be asked to learn the declarative concept *social norms* and its corresponding definition, *explicit or implicit conventions that dictate appropriate behavior in social situations.* Learning declarative concepts is particularly important because they often serve as the foundation for understanding more complex ideas presented in a course.

To help students learn declarative concepts, instructors can encourage their students to use many study strategies. One strategy that can be used to support declarative concept learning is *example generation*, a strategy in which students generate concrete examples of a declarative concept. This strategy has intuitive appeal for learning declarative concepts because these concepts are abstract in nature and can be applied to many different concrete situations. For example, with the declarative concept used above (social norms), a student may generate the example “Everyone at the funeral was wearing black. It was the social norm” (this example was generated by a student in the current study). Students report generating examples while studying (Gurung, Weidert, & Jeske, 2010; Weinstein, Lawrence, Tran, & Frye, 2013). Moreover, instructors may instruct their students to generate examples as part

☆ The research reported here was supported by a James S. McDonnell Foundation 21st Century Science Initiative (JSMF #220020166) in Bridging Brain, Mind and Behavior Collaborative Award.

* Corresponding author. Kent State University, P.O. Box 5190, Kent, OH 44242-0001, USA.
E-mail address: azamary@kent.edu (A. Zamary).

http://dx.doi.org/10.1016/j.learninstruc.2016.08.002
0959-4752/Published by Elsevier Ltd.

of in-class activities or as homework assignments, because example generation may benefit students' comprehension of the concepts and also because one instructional goal of teaching declarative concepts is to help students use them outside of the classroom, which requires applying concepts in new contexts. Effective student self-regulation is important inside the classroom as well, such as when instructors assist students during class (de Bruin & Van Gog, 2012).

While using almost any strategy, effective self-regulated learning depends in part on accurate monitoring and effective control (Dunlosky, Hertzog, Kennedy, & Thiede, 2005; Thiede, 1999; Winne & Hadwin, 1998). Monitoring refers to evaluating one's ongoing learning or performance. Control refers to many different decisions that may influence how one proceeds, such as deciding what to restudy (e.g., selecting content that has been judged as less well known), how long to restudy, and what strategies to use (Nelson & Narens, 1990; Winne, 2001). Self-regulated learning theories also include a broad array of other processes that can be triggered when students learn, including goal setting, emotional responses to learning progress, help seeking, among others (for various perspectives on self-regulated learning, see Zimmerman & Schunk, 2001). Even so, fundamental to many of these theories are feedback loops in which students' monitoring is used to guide subsequent control processes (see especially, Winne, 2001). Therefore, if students do not monitor accurately, it will likely undermine the effective control of study (e.g., by not selecting the appropriate information to study, by not spending an appropriate amount of study time, or by not using effective study strategies).

In the present case, if students accurately monitor when they are generating poor examples, they could effectively control their subsequent learning by trying to generate a better example or by seeking help to better understand the target concept (e.g., from a textbook or instructor). By contrast, if students do not monitor accurately and are overconfident in the quality of the examples they generate, they may poorly control study, prematurely discontinue practice, fail to seek help to better understand the concepts, and so on. As a result of these suboptimal control decisions, students in turn may perform poorly on subsequent tests because they did not spend enough time on concepts they had not yet fully learned (Dunlosky & Rawson, 2012). For instance, a student in the current study who was learning the concept social norms generated the example "The woman acted very strange in a crowd of people at the mall". This example is entirely incorrect (i.e., a *commission error*) because it does not illustrate any of the essential ideas within the concept definition; "acting strange" is not a conventional or appropriate behavior in the social situation of a crowd in a mall. If the student judges this example to be entirely correct or partially correct, they would be displaying overconfidence. *Overconfidence* in this context is defined as the degree to which a student believes their example is of greater quality than it actually is. Such overconfidence in example quality would likely curtail further study because students may stop studying concepts that they believe they fully understand.

Thus, for generating examples to benefit students' learning (either through in- class activities or during self-regulated study outside of class), students will need to be able to accurately evaluate the quality of their examples so that they can identify when they have not yet learned a concept well. Unfortunately, no prior research has investigated whether students can accurately evaluate the quality of their generated examples. Addressing this issue was the first aim of the current research. The second aim was to investigate the extent to which providing feedback to students during their evaluation of their generated examples can improve the accuracy of their example-quality judgments. Critically, if students are overconfident even with feedback, instructors will need to help their students better evaluate the quality of generated examples.

1.1. Question 1: how well can students evaluate the quality of their examples?

Although no research has directly investigated how well students can evaluate the quality of the examples they generate, research on metacomprehension indicates that students often have trouble evaluating their own learning and comprehension of text material (Dunlosky & Lipko, 2007; Thiede, Griffin, Wiley, & Redford, 2009). Within this literature, the studies most relevant for present purposes have investigated how well students can evaluate the accuracy of their own cued recall responses when prompted to retrieve declarative concept definitions. In a study by Rawson and Dunlosky (2007), students read text passages on several different topics (e.g., psychological measurement) and were then asked to recall definitions of concepts contained in the passages. After recalling each definition, students were prompted to make the following judgment: "If the correctness of the definition you just wrote was being graded, do you think you would receive no credit, partial credit, or full credit?" For recall responses that were commission errors (i.e., entirely incorrect, based on experimenter scoring), participants judged that they were partially or fully correct 83% of the time (see also Dunlosky & Rawson, 2012; Dunlosky, Hartwig, Rawson, & Lipko, 2011). Although this prior research involved evaluation of cued recall responses for definitions rather than generated examples of declarative concepts, the consistent pattern of outcomes across studies provides empirical support for the expectation that students will also be overconfident in the quality of their generated examples.

In addition to empirical support for the prediction that students will have difficulty evaluating the quality of their generated examples, the literature on metacognitive monitoring provides a theoretical account for *why* students may not accurately evaluate the quality of their responses. According to the accessibility hypothesis (Koriat, 1993), learners base their self-evaluations on the sheer amount of information that is retrieved prior to making a judgment. According to this hypothesis, the more information accessed at the time of judgment leads to greater confidence in the accuracy of what is retrieved, regardless of the quality of what is retrieved. Thus, if a student recalls a great deal about a particular concept, they will judge it as relatively accurate even if all the information retrieved is incorrect – that is, they will show overconfidence for commission errors. Accessibility does partly explain students' overconfidence in judging the accuracy of their recall of key term definitions (Dunlosky, Rawson, & Middleton, 2005; Rawson & Dunlosky, 2007), and it can also be applied to example generation. Namely, students' evaluations of their examples may be based on the amount of information that they generate when developing an example; if so, when they generate long examples that are incorrect, their judgments would be expected to demonstrate substantial overconfidence.

1.2. Question 2: is judgment accuracy influenced by the kind of feedback provided?

No prior research has directly investigated the impact of feedback on evaluating the quality of self-generated examples. However, indirect evidence from the metacomprehension literature provides empirical support for the prediction that giving students feedback will improve their judgment accuracy. Lipko et al. (2009) had middle-school students study declarative concepts from the topics of genetics and literary non-fiction. After initial self-paced study, students were asked to recall the definitions of concepts.