Learning and Instruction xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Learning and Instruction

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

When analogies harm: The effects of analogies on metacomprehension☆

Jennifer Wiley*, Allison J. Jaeger[1], Andrew R. Taylor[2], Thomas D. Griffin

University of Illinois at Chicago, 1007 W Harrison St MC 285, Chicago, IL, 60607, USA

ARTICLE INFO

Keywords:
Metacomprehension
Analogical examples
Expository science texts

ABSTRACT

The main goal of the present research was to test whether the presence of analogies would affect the relative accuracy of metacognitive judgments about learning from expository science texts, and whether any effect would depend on the type of cues that readers used as the basis for their judgments of comprehension. In a series of experiments, students read texts that either contained or did not contain analogies; were asked to judge how well they understood each text; took comprehension tests for each topic; and were asked to self-report the basis for their judgments. Relative metacomprehension accuracy was computed as the intra-individual correlation between judgments and test performance. Results showed that the presence of analogies can lead to poor relative metacomprehension accuracy for students who fail to use situation-model-based cues to judge their understanding of text.

1. Introduction

Many individuals suffer from poor metacomprehension accuracy for science texts, meaning they are generally quite poor at gauging how well they have understood what they have read (Dunlosky & Lipko, 2007; Maki, 1998; Wiley, Griffin, & Thiede, 2005). A major strand of research on metacomprehension accuracy follows the lead of initial studies by Maki and Berry (1984) and Glenberg and Epstein (1985). To assess how well people can evaluate their own understanding of texts, they developed a paradigm to assess the accuracy of comprehension monitoring (metacomprehension) in which individuals study a set of expository texts, judge their understanding for each of the texts, and then complete comprehension tests. From this data, they computed *relative metacomprehension accuracy* (Maki, 1998) as a measure of whether a person knows which texts they understood the best, relative to those they understood the least. *Relative metacomprehension accuracy* is operationalized as the within-person or intra-individual correlation between an individual learner's judgments of understanding for a number of different topics, and that same learner's actual understanding for each of those topics as assessed by objective tests of comprehension.

Relative accuracy mimics the regulation of learning that students need to engage in on a daily basis. Students are routinely tasked with learning about a variety of topics during any given night of homework, as well as gauging their understanding across all topics when they study for cumulative tests. Being able to have an accurate sense of one's relative levels of understanding for various concepts is critical to effectively regulating how much effort to devote to studying and restudying on each topic. Most studies examining *relative metacomprehension accuracy* using this standard paradigm have used sets of expository texts on scientific topics consisting of 4–6 texts over 500 words each. Typically, baseline conditions (without additional instructions or manipulations) in these studies have found low levels of relative accuracy hovering around $r = 0.27$ (see Dunlosky & Lipko, 2007; Maki, 1998; Thiede, Griffin, Wiley, & Redford, 2009; and Griffin, Mielicki, & Wiley, in press; for reviews). Perfect accuracy would be a correlation between judgments and performance of 1.00.

1.1. Why do individuals tend to have poor relative metacomprehension accuracy?

To explain poor relative metacomprehension accuracy, theories of metacomprehension such as the *situation-model approach to metacomprehension* (Wiley, Thiede, & Griffin, 2016) have integrated the cue-utilization approach (Koriat, 1997) which makes a distinction between the use of valid and invalid cues as a basis for monitoring judgments, with Kintsch's (1994, 1998) comprehension paradigm which makes a distinction between memory for text and learning from text. In combination, these frameworks suggest that accurate comprehension monitoring depends on the use of valid cues, and specifies that situation-model-based cues will be most valid when learning will be

☆ This research was completed at the University of Illinois at Chicago.

* Corresponding author.
E-mail address: jwiley@uic.edu (J. Wiley).
[1] Allison J. Jaeger is now at the Spatial Intelligence Learning Center (SILC), Temple University.
[2] Andrew R. Taylor is now at the Division of Public Health Sciences, Henry Ford Health Systems, Detroit, Michigan.

http://dx.doi.org/10.1016/j.learninstruc.2017.10.001
Received 16 May 2017; Received in revised form 7 October 2017; Accepted 14 October 2017

evaluated through comprehension questions that require representing key relations among ideas (i.e. the situation model). Following Kintsch's comprehension framework, readers should derive their predictive judgments of performance on comprehension questions based on their situation models and not based on their memory for discrete ideas or details mentioned in the text. Thus, these theories predict that situation-model cues will be the most valid basis for making accurate predictions of performance on upcoming tests to the extent that comprehension is tested by measures that depend on the quality of the situation model.

In light of these theoretical assumptions, one prevailing explanation for the low levels of relative metacomprehension accuracy that are typically observed is that students do not seem to know what kind of information they should use as a basis for their judgments of comprehension. Koriat's (1997) cue-utilization framework asserts that learners engage in an inferential process when they are asked to make predictive judgments of how they will do on upcoming tests. That is, learners attempt to gauge their likely test performance from a variety of different types of cues. These cues can include features of the learning context that are intrinsic or extrinsic to the learning materials (e.g., amount of information, nature of the items, or number of study episodes), but they can also include cues that reflect subjective experiences of attempting to process the information. Considering both intrinsic and extrinsic features of a learning context results in theory-based inferences, where learners use their beliefs or assumptions about the effect that such features may have on learning to determine their judgments. These kinds of cues can be thought of as rule-based or heuristic because they rely on *a priori* knowledge and assumptions about how contextual features should impact the learning process in general. On the other hand, internal subjective indicators of performance derived from the actual learning experience provide a basis on which to determine the extent to which a particular item has been learned (Fischer & Mandl, 1984; Flavell, 1979; Griffin, Jee, & Wiley, 2009; Griffin, Wiley, & Salas, 2013; Nelson & Narens, 1990). In Koriat's (1997) research on judgments of learning in a metamemory context, he demonstrated that relative accuracy improves as learners shift from greater reliance on theory-based cues to greater reliance on experience-based cues when making predictive judgments about future performance on a memory test. Conceptually, this makes sense because *a priori* theories will tend to be general and thus will fail to predict variation in performance within the same person from one instance to the next.

Extending this logic to a context in which a reader is tasked with learning from texts, the *situation-model approach to metacomprehension* (Wiley et al., 2016) further posits that subjective *experience-based* cues that are anchored in the quality of the mental representation constructed during reading will provide the most valid basis from which to infer the extent to which one has understood a text on a particular topic (Rawson, Dunlosky, & Thiede, 2000; Weaver, Bryant, & Burns, 1995; Wiley et al., 2005). That is, greater reliance on experience-based cues should also increase the relative accuracy of predictive judgments about future performance on a comprehension test. Further, Kintsch's comprehension framework (1994, 1998) suggests that not all subjective experiences will be equally valid indicators of understanding. Some perceptions, such as a sense of fluency, an apparent ease of processing while reading, or the feeling that one will remember details from a text, can sometimes be useful; but in other cases (i.e. when the comprehension will be tested with questions assessing understanding of key relations among ideas, rather than memory-for-details questions) they will be less useful because they are not indicative of the quality of the mental model of the phenomena that is being constructed. Therefore these perceptions would be less valid as cues for making accurate judgments about an individual's understanding of the topic. Similarly, heuristic cues such as the length or difficulty of a text would also be less valid, because they are not invariably linked to the quality of an individual's understanding about the topic. Cues that are specifically based in the quality of the situation model, or the mental model of the scientific system or process that one constructs while reading, will be the most valid basis for making accurate predictions about performance on upcoming comprehension tests that include questions assessing understanding of key relations among ideas.

Some examples of ways that readers can generate valid cues are to engage in self-explanation about how or why a process works, and to think about the extent to which a causal explanation is complete and coherent (c.f. Griffin, Wiley, & Thiede, 2008; Jaeger & Wiley, 2014; Wiley, Griffin, Jaeger, et al., 2016). Similarly, readers can think about whether or not they could explain how and why the process works or phenomenon happens to another student. Or they could consider whether they might be able to make new inferences based on the information, or apply the principles in the text to a new situation. Importantly, readers do not need to possess a high quality situation model or need to be able to construct a complete and coherent causal explanation in order to have good relative metacomprehension accuracy. Readers just need to *attempt* to engage in these activities because the results of these attempts represent *experiences* that will provide them with cues about the quality of their mental models, which in turn will provide them with a valid basis for making their judgments of understanding.

On the other hand, relying on superficial cues such as text length, fluency, interest in the topic, or perceived difficulty should be less likely to lead to accurate judgments of understanding, and thus these superficial cues can be seen as less valid. Thiede, Griffin, Wiley, and Anderson (2010) found evidence consistent with these theoretical assumptions by asking readers to self-report the types of cues that they used as a basis for their judgments of understanding. They found that students who reported using heuristic cues were generally less accurate than those who reported using experience-based cues. Similarly, Jaeger and Wiley (2014) found that students who reported using comprehension-based cues (using their ability to explain, summarize or make connections while reading the text, or whether they thought they could answer questions like those they saw on the practice test as the basis for their judgments) had higher levels of relative metacomprehension accuracy than those who used non-comprehension-based cues.

1.2. How do instructional adjuncts affect metacomprehension accuracy?

To date, most studies on metacomprehension have investigated accuracy in the context of reading non-illustrated expository text passages. Work has just begun to assess how different adjuncts to text, or features of texts, might influence metacomprehension accuracy. The main question for the current line of investigation is whether the presence of analogies, a common instructional device, might impact metacomprehension accuracy.

Because not much work has been done yet with instructional analogies and metacomprehension, it is useful to consider a related line of research exploring the effects of including visualizations alongside text. In some domains like geology, biology and chemistry, a popular way of supporting understanding is through providing visualizations such as diagrams, schematics, animations, or simulations that are thought to be helpful in conveying invisible phenomena or relations that may be implicit in written text (Ainsworth & Loizou, 2003; Butcher, 2006; Larkin & Simon, 1987). Despite their intuitive promise, many studies have shown that visualizations can fail to improve or even impede learning outcomes compared to plain text (Hegarty, 2014; Höffler & Leutner, 2007). Much depends on the topic, the quality of the visualizations, and the learning task, as well as on the learner. The theory of multimedia learning suggests that when designing multimedia-learning materials, such as texts accompanied by visualizations, the presented material should have a coherent structure and be presented in a manner that provides guidance as to how to build an accurate mental model. In particular, Mayer (2005) lays out several principles of multimedia instructional design that are aimed at minimizing the effects of extraneous processing by emphasizing main ideas