Learning and Individual Differences 32 (2014) 40–53

Contents lists available at ScienceDirect

Learning and Individual Differences

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

Development of a new reading comprehension assessment: Identifying comprehension differences among readers ☆

Sarah E. Carlson [a,*], Ben Seipel [b], Kristen McMaster [c]

[a] University of Oregon, Center on Teaching and Learning, United States
[b] California State University, Chico, United States
[c] University of Minnesota, Twin Cities, United States

ARTICLE INFO

Article history:
Received 28 September 2012
Received in revised form 9 January 2014
Accepted 17 March 2014

Keywords:
Reading comprehension assessment
Individual differences
Reading comprehension processes

ABSTRACT

The purpose of this study was to evaluate the Multiple-choice Online Cloze Comprehension Assessment (MOCCA), designed to identify individual differences in reading comprehension. Data were collected with two sets of 3rd through 5th grade students during two years: 92 students participated in Year 1 and 98 students participated in Year 2 to address primary research questions, and an additional 94 ($N = 192$) students participated in Year 2 to address the limitation of test administration time. Participants were group administered the MOCCA and a standardized reading proficiency assessment, and individually administered other reading measures. Preliminary analyses indicated that the MOCCA produced reliable and valid scores as a new reading comprehension assessment for identifying types of comprehension processes used during reading, as well as for identifying individual differences in the types of comprehension processes used during reading. Findings are discussed in terms of developing a new measure to identify cognitive reading comprehension processes used during reading. Future research is needed to provide additional support for the technical adequacy of the assessment.

1. Introduction

Many students struggle with reading, and in particular, reading comprehension. As students advance in school, they transition from learning to read (e.g., learning to decode and developing fluency and comprehension skills) to reading to learn (e.g., using comprehension skills to learn from text; Chall, 1996). This transition is often most evident in the upper elementary grades, when many readers begin to encounter difficulties with new comprehension requirements (Shanahan & Shanahan, 2008, 2012).

Assessments are needed to determine why readers experience comprehension difficulties in order to develop appropriate instruction to meet their individual needs, yet few such assessments are available. Thus, the purpose of this study was to report preliminary findings from a new reading comprehension assessment, the Multiple-choice Online Cloze Comprehension Assessment (MOCCA), developed to identify individual differences in reading comprehension. In this paper, we first discuss theories of reading comprehension that guided the development of MOCCA. Second, we describe existing reading comprehension assessments used to measure specific aspects of comprehension, and how they have informed the development of MOCCA. Finally, we report initial evidence of the reliability and validity of MOCCA, and discuss how the present study extends the reading comprehension assessment literature.

1.1. Reading comprehension theories and assessments

Reading comprehension is a complex and multidimensional construct; thus, the development of reading comprehension assessments should be guided by theory (August, Francis, Hsu, & Snow, 2006; Fletcher, 2006). Reading comprehension theories help identify constructs that work during the process of comprehension and specify the relationships among them so that researchers can better operationalize the dimensions to be assessed.

Reading comprehension theories suggest that successful reading comprehension involves the extent to which a reader can develop a coherent mental representation of a text through developing a coherent situation model (e.g., Graesser, Singer, & Trabasso, 1994; Kintsch, 1998; McNamara, Kintsch, Songer, & Kintsch, 1996; van den Broek, Rapp, & Kendeou, 2005). A situation model is comprised of the situations that take place in a text (e.g., time, space, characters, and causality) (van Dijk & Kintsch, 1983; Zwaan, Magliano, & Graesser, 1995). For instance,

☆ This research was supported by Grant #R305C050059 from the Institute of Education Sciences (IES), U.S. Department of Education, to the University of Minnesota and through the Interdisciplinary Education Sciences Predoctoral Training Program, "Minnesota Interdisciplinary Training in Education Sciences (MITER)" for data collection and resources, as well as by Grant #R305b110012 from the IES, U.S. Department of Education, to the Center on Teaching and Learning at the University of Oregon, through a Postdoctoral Fellowship for writing resources. The opinions expressed are those of the authors and do not necessarily represent views of the IES or the U.S. Department of Education.

* Corresponding author at: University of Oregon, 1600 Millrace Drive, Suite 207, Eugene, OR 97403, United States. Tel.: +1 541 346 8363; fax: +1 541 346 8353.
E-mail address: carlsons@uoregon.edu (S.E. Carlson).

http://dx.doi.org/10.1016/j.lindif.2014.03.003
1041-6080/

a reader may track *causality* by keeping track of the goal of the text (Trabasso & van den Broek, 1985; van den Broek, Lynch, Naslund, Ievers-Landis, & Verduin, 2003). The following example describes a causal connection: "Jimmy wanted to buy a bike. He got a job and earned enough money. He went to the store to buy the bike. Jimmy was happy." In this example, a reader could make a causal connection by generating an inference that Jimmy was happy because he reached his goal and bought a bike.

Researchers have found that many poor comprehenders (i.e., readers with adequate word reading skills but with poor comprehension skills compared to peers with similar word reading skills) fail to make causal inferences while reading, which may stem from failure to track causal relations and goals in a text (e.g., Cain & Oakhill, 1999, 2006; McMaster et al., 2012; Rapp, Broek, McMaster, Kendeou, & Espin, 2007; van den Broek, 1997). To provide appropriate instruction to improve such inference generation, it is important that reading comprehension assessments identify the specific processes with which poor comprehenders struggle.

Researchers have assessed reading comprehension *processes* to understand how readers build connections (i.e., inferences) and track relations during reading to develop a coherent representation of a text, and have assessed reading comprehension *products* to evaluate the result of the representation of the text. The *products* are the 'end result' of reading, or what the reader learned or stored in memory from the text after reading (i.e., *offline*). Reading products are typically assessed using recall, questioning activities, and traditional multiple-choice assessments.

In contrast, reading *processes* occur during the act of reading (i.e., *online*) and can be somewhat more difficult to assess because the examiner must infer what is taking place during reading. Methods to assess online reading comprehension processes include eye-tracking methods, reading time measures, and think-aloud tasks (e.g., Ericsson & Simon, 1993; Kaakinen, Hyönä, & Keenan, 2003; Linderholm, Cong, & Zhao, 2008). Think-aloud tasks, for example, are used to identify specific reading comprehension processes (e.g., causal, bridging, elaborative inferences; paraphrases) that readers use during reading (Ericsson & Simon, 1993). Findings from think-aloud studies indicate that readers use different types of comprehension processes during reading to develop coherent situation models (e.g., Laing & Kamhi, 2002; Trabasso & Magliano, 1996a,b; van den Broek, Lorch, Linderholm, & Gustafson, 2001). Although think-aloud data provide fruitful information about the processes that readers use during comprehension, they are laborious, time consuming, and impractical for practitioners to use to identify reading comprehension differences among their students for instructional purposes.

1.2. Identifying comprehension differences

Researchers who have assessed reading comprehension processes using think-aloud methods have identified individual processing differences among readers at different levels of comprehension skill (McMaster et al., 2012; Rapp et al., 2007). Specifically, McMaster et al. (2012) administered a think-aloud task to fourth grade readers at different levels of comprehension skill (i.e., good, average, and poor). They identified two types of poor comprehenders: (1) paraphrasers: poor comprehenders who mostly paraphrased during reading; and (2) elaborators: poor comprehenders who elaborated about the text, including information that was connected to background knowledge that was not always relevant to the text. These findings were consistent with previous research that found similar types of poor comprehenders, and support other researchers' conclusions that poor comprehenders may struggle with reading in different ways (Cain & Oakhill, 2006; Nation, Clarke, & Snowling, 2002; Perfetti, 2007; Rapp et al., 2007).

McMaster et al. (2012) also found that the two types of poor comprehenders responded to intervention in different ways. Specifically, they compared two questioning interventions: one that prompted readers to answer causal questions (*Why* questions that prompted readers to make causal connections during reading), and one that prompted readers to answer general questions (questions that prompted readers to make any kind of connections during reading). The researchers found that paraphrasers benefitted more from the general questioning intervention than elaborators did, whereas elaborators benefited more from the causal questioning intervention than paraphrasers did. These findings suggest that different types of poor comprehenders may respond differently to intervention.

Though researchers have employed methods to assess reading comprehension processing differences among readers (e.g., think-aloud tasks), most traditional school-based reading comprehension assessments (e.g., reading proficiency assessments, standardized measures) have not been designed to detect such processes or to identify individual comprehension differences. In addition, many of these methods assess the *product* of reading comprehension rather than the *process*, limiting the types of conclusions that can be drawn about *how* readers comprehend differently. For example, Keenan, Betjemann, and Olson (2008) found that commonly used standardized reading comprehension assessments measure aspects of reading such as decoding and word recognition, but not necessarily reading comprehension, and what is measured varies depending on the age of the reader. Thus, such traditional assessments may be insufficient for identifying specific reading comprehension differences; yet, educators often make instructional decisions based on their outcomes (Keenan et al., 2008).

Researchers have begun to develop other methods to help address the shortcomings of traditional reading assessments and measure how readers comprehend text rather than only assessing the product of comprehension. For instance, Magliano and colleagues developed the Reading Strategy Assessment Tool (RSAT; Magliano, Millis, Development Team, Levinstein, & Boonthum, 2011), which measures a subset of the comprehension processes found to lead to a coherent representation of a text. RSAT is an automated computer-based assessment in which readers read texts one sentence at a time, and are asked either indirect questions (i.e., "What are your thoughts regarding your understanding of the sentence in the context of the passage?") or direct questions (i.e., *Why* questions related to a target sentence). Readers type their responses, which are later analyzed for types of comprehension processes (e.g., paraphrases, bridging inferences, elaborations) and content words (e.g., nouns, verbs, adjectives, adverbs) used during reading.

Magliano et al. (2011) identified unique types of comprehension processes that readers used during reading using RSAT, and also found that RSAT predicted scores on measures of reading comprehension. However, the measure is limited in several ways. First, RSAT uses an open-ended response format where participants type their responses to questions, limiting its use to older participants who have developed appropriate typing skills. Second, linguistic algorithms used to identify the types of comprehension processes produced in responses may be limited in capturing the quality of responses and identifying individual profiles of readers. Finally, like think alouds, the open-ended response task used in RSAT can produce a large amount of variability in how readers interpret the task instructions, especially the instructions for answering the *indirect* question which could be interpreted differently from reader to reader. Thus, it seems useful to develop an assessment that capitalizes on the strengths of RSAT (e.g., identify comprehension processes during reading), but is also familiar to readers in terms of testing format, efficient for educators to administer and score, and can be used for making instructional decisions with children.

Other recently developed assessments, such as the Diagnostic Assessment of Reading Comprehension (DARC; August et al., 2006) and The Bridging Inferences Test, Picture Version (Bridge-IT, Picture Version; Pike, Barnes, & Barron, 2010), measure individual differences in reading comprehension processes for readers in Grades 2–6. The DARC requires readers to remember newly read text, connect to and integrate relevant background knowledge, and generate bridging