

18

THE READING STRATEGY ASSESSMENT TOOL

A Computer-Based Approach for Evaluation Comprehension Processes during Reading

*Joseph P. Magliano, Melissa Ray,
and Keith K. Millis*

Introduction and Overview of the Reading Strategy Assessment Tool (RSAT)

When teachers, administrators, and researchers think of comprehension tests, they typically picture a reading passage or set of reading passages followed by a set of multiple-choice questions. Students are expected to read each passage and then select the best answer to each question. When answering these questions, they often are able to refer back to the passages, which may have implications on what aspects of comprehension these test formats actually assess (Farr et al., 1990). They may be better suited for assessing how well students can take tests on a text, as opposed to providing a direct assessment of comprehension ability.

RSAT (Magliano, Millis, The RSAT Development Team, Levinstein, & Boonthum, 2011) adopts a very different approach. RSAT is a computer-administered test that is designed to assess comprehension *while* students read and the *strategies* that students use to support their comprehension during reading (Gilliam et al., 2007; Magliano et al., 2011). It was originally developed in a post-secondary environment, but work is underway to develop a version that is appropriate for students in middle school. RSAT is well grounded in theory and empirical research of how students comprehend texts (Magliano et al., 2011), which indicate that readers must make connections between sentences (bridge) and use their world knowledge (elaborations) to construct a coherent text representation. These processes are assumed by most theories of comprehension in psychology to be important for building a durable memory for a text.

In RSAT, students read texts one sentence at a time. Only the sentence that they are currently reading is available onscreen, and students click a button to continue to the next sentence. While it may seem odd to have students only see one sentence at a time, this decision was made to force students to draw upon

their memory for what they have read so far to answer the questions. We have shown that this approach provides a better indicator of a student's comprehension skill than allowing the student to have full access to what they have read so far (Gilliam et al., 2007), albeit it is not a naturalistic way of presenting text content (Magliano & Graesser, 1991). After pre-selected sentences are read, an open-ended question appears. When a question appears onscreen, the text title is displayed at the top of the screen, but no sentences from the text are visible. Just below the question, a textbox also appears, and students must type their responses to the questions before moving onto the next sentence.

There are two types of questions: direct and indirect. *Direct questions* are why-, what-, and how-questions that assess the extent to which students comprehend the text (e.g., "Why do we hear thunder after seeing the lightning strike?" on text about lightning). Ideal answers to direct questions can be derived from information presented in the prior text. *Indirect questions* always ask students "What are you thinking now?" Students are instructed to answer this question by typing their thoughts regarding their understanding of what they are reading and are given practice on how to do this before taking RSAT. This question answering process is akin to thinking aloud (Trabasso & Magliano, 1996). Figure 18.1 contains sample screen shots for two sentences of a text and the direct or indirect question that would appear after the second sentence (students see either a direct or an indirect

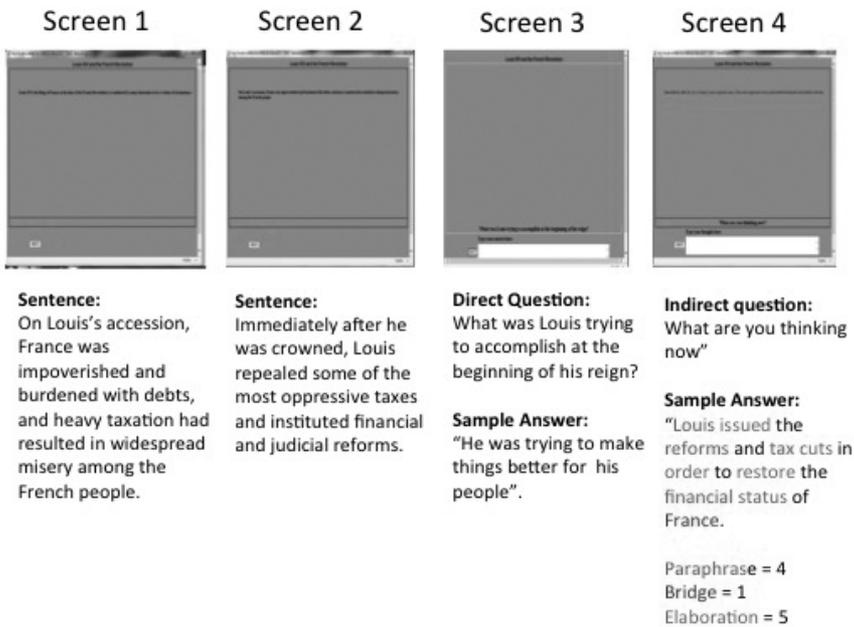


FIGURE 18.1 Sample Screen Shots for Two Sentences of a Text and the Direct or Indirect Question that Would Appear after the Second Sentence.

question). Screen 3a contains a direct question and sample answer. Similarly, screen 3b contains an indirect question and sample response.

Answers to the direct questions feed into the global comprehension score, and answers to the indirect questions feed into the strategy scores. RSAT provides a measure of three strategies involved in comprehension: bridging inferences (connecting sentences with one another), elaborative inferences (connecting the text with what they already know), and paraphrases (producing key words from the sentence being read). The scores are derived from computer-based algorithms that automatically score answers to direct and indirect questions. Screen 3b in Figure 18.1 shows how the RSAT algorithms use the words in an answer to an indirect question to generate strategy scores: paraphrasing of the current sentences (red), bridging to prior text concepts (green), or elaborating upon the texts (blue).

The current version of RSAT contains both fiction and nonfiction, informational texts (history and science). The texts have between 20 and 30 sentences, with about 6 questions per text (about half of each type of question). Some forms of RSAT only ask one type of question (e.g., only indirect questions) and other forms have both types of questions (with approximately an equal number of each). These forms were created, because researchers and practitioners may be interested in assessing comprehension or comprehension processing. Fiction texts are folktales, and informational texts are similar to brief encyclopedia articles. The texts have Flesch-Kincaid readability scores between 9th and 10th grade. This readability level was chosen so that texts would be *within* the reading skill of college students (the population for which RSAT was developed). As a result, RSAT would be able to assess comprehension processes when students read texts within their reading abilities. With more difficult texts comprehension might be negatively affected by things such as complex sentence structure or low frequency words, which may mask or interfere with the intended strategies of bridging, elaborations, and paraphrases.

What evidence is there for the validity of RSAT? First, the RSAT approach has been shown to be as correlated with curriculum-like tests of comprehension (e.g., end of chapter questions). Additionally, there is evidence that the RSAT comprehension score is correlated with performance on well-established standardized tests of comprehension, such as ACT (Magliano et al., 2011). Moreover, the strategy measures (i.e., paraphrase, bridging, and elaboration scores) are correlated with these measures of comprehension roughly to the same extent that they correlate with one another (Gilliam et al., 2007; Magliano et al., 2011). Finally, we have demonstrated that RSAT scores are highly correlated with human judgments of the use of comprehension processes (Magliano et al., 2011).

Why is RSAT Needed?

National and international assessment data indicate that today's adolescents fail to master important skills in core areas, such as reading (Carnegie Council on

Advancing Adolescent Literacy, 2010). The development of effective screening tools that can be used by both teachers and school psychologists is needed to address this problem (see Al Otaiba & Fuchs, 2006). Practitioners are typically stuck with inauthentic multiple-choice type reading measures that provide evidence only of the outcome of reading. While such assessments are easy to administer and may have global diagnostic value, they tell little of students' strengths and weaknesses when it comes to comprehending texts. Practitioners need *formative* assessments that provide more direct evaluations of what students do as they read. Unfortunately, currently, there are few high quality formative assessments (Afflerbach, Cho, & Kim, 2015), and none that can assess the actual cognitive processes used during reading comprehension.

RSAT's assessment of comprehension "in the moment" is an important step in developing an assessment tool that fills this gap. Because it makes comprehension processes transparent, explicit, and tangible to school psychologists, teachers, students, and parents, all stakeholders can use RSAT to learn about important aspects of students' reading comprehension beyond a single score reflecting "comprehension." Additionally, McMasters and colleagues have shown that identifying different profiles of what students do when they think aloud while reading can provide a basis for tailoring interventions to the specific needs of students (McMasters et al., 2012). As stated earlier, the indirect questions in RSAT elicit think-aloud-like responses (Muñoz et al., 2006). Therefore, RSAT may provide an automated tool for generating student profiles, providing a basis for which to identify the most appropriate interventions.

In addition, the strategies measured by RSAT are similar to those emphasized in some reading interventions that focus on moment-to-moment reading, such as iSTART (see Chapter 7), which have been shown to improve both comprehension ability and class performance (e.g., McNamara & Scott, 1999). To the extent that there are multiple forms of RSAT (i.e., multiple versions of forms so that RSAT can be given multiple times to the same students), the tool has the potential to be used to monitor progress. For example, assume a student takes RSAT and essentially only paraphrases the sentences that immediately preceded an indirect prompt. This strategy has been shown to be detrimental for comprehension (Coté & Goldman, 1999; Magliano & Millis, 2003). However, if students were exposed to an intervention that promoted inference processes, then RSAT could be used to assess if students increase their inferencing skill over time.

Plans for Developing RSAT

Current efforts are focused on making RSAT a comprehension assessment that is appropriate and available for use in middle and high schools. Because it was originally developed to assess how college freshmen comprehend text, RSAT may not be suitable for younger readers. Additionally, future versions of RSAT will be developed to contain tools that help practitioners to interpret, evaluate,

and use RSAT scores. To date, RSAT's development has primarily centered on exploring the feasibility of using automated scoring of "think alouds" to assess comprehension. As a result, additional work is needed to explore the classroom applications of RSAT, which include larger tests of its reliability and validity. Finally, in order to support its use in schools, a user-friendly website will need to be developed so that practitioners and students can readily access RSAT and RSAT scores.

The current goal for RSAT development is to produce a version of RSAT that will help practitioners identify readers' initial strengths and weakness and measure how their comprehension changes over time. Because it is not a summative assessment, RSAT is likely not an appropriate replacement for traditional, high stakes tests. We believe that RSAT's greatest potential is as a tool to evaluate student progress toward developing basic comprehension proficiency.

Timeline for Developing RSAT

We will be developing and selecting new texts and exploring different types of computer-based algorithms that can be used to generate RSAT scores for middle school aged students. Importantly, we will develop tools that will help teachers use RSAT to evaluate students in meaningful ways, and help students learn more about what they do to understand what they read. Beta version of RSAT should be available by 2017, with the intent of a scalable version available by 2019.

References

- Afflerbach, P., Cho, B. Y., & Kim, J. Y. (2015). Conceptualizing and assessing higher order thinking in reading. *Theory Into Practice*, 54, 203–212
- Al Otaiba, S. & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning Disabilities*, 39, 414–431.
- Carnegie Council on Advancing Adolescent Literacy (2010). *A Time to Act*. New York: Carnegie Corporation of New York.
- Coté, N., & Goldman, S. R. (1999). Building representations of informational text: Evidence from children's think-aloud protocols. In H. Van Oostendorp & S. R. Goldman (Eds.), *The Construction of Mental Representations during Reading* (pp. 169–193). Mahwah, NJ: Erlbaum.
- Farr, R., Pritchard, R., & Smitten, B. (1990) A description of what happens when an examinee takes a multiple-choice reading comprehension test. *Journal of Educational Measurement*, 27, 209–226.
- Gilliam, S., Magliano, J. P., Millis, K. K., Levinstein, I., & Boonthum, C. (2007). Assessing the format of the presentation of text in developing a Reading Strategy Assessment Tool (R-SAT). *Behavior Research Methods, Instruments, & Computers*, 39, 199–204.
- Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying inference generation in literary text. *Poetics*, 20, 193–232.
- Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure. *Cognition and Instruction*, 21, 251–283.

- Magliano, J. P., Millis, K. K., The RSAT Development Team, Levinstein, I., & Boonthum, C. (2011). Assessing comprehension during reading with the reading strategy assessment Tool (RSAT). *Metacognition and Learning*, 6, 131–154.
- McMaster, K. L., Van den Broek, P., Espin, C. A., White, M. J., Kendeou, P., Rapp, D. N., . . . Carlson, S. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences*, 22, 100–111.
- McNamara, D. S., & Scott, J. L. (1999). Training reading strategies. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty-first Annual Meeting of the Cognitive Science Society* (pp. 387–392). Hillsdale, NJ: Erlbaum.
- Muñoz, B., Magliano, J. P., Sheridan, R., & McNamara, D. S. (2006). Typing versus thinking aloud when reading: Implications for computer-based assessment and training tools. *Behavior Research Methods, Instruments, & Computers*, 38, 211–217.
- Trabasso, T. & Magliano, J. P. (1996). Conscious understanding during text comprehension. *Discourse Processes*, 21, 255–288.