AJSLP

Technical Report

Read, Understand, Learn, & Excel: Development and Testing of an Automated Reading Strategy Detection Algorithm for Postsecondary Students

Priya Kucheria,[a] McKay Moore Sohlberg,[a] Jason Prideaux,[b] and Stephen Fickas[b]

Purpose: An important predictor of postsecondary academic success is an individual's reading comprehension skills. Postsecondary readers apply a wide range of behavioral strategies to process text for learning purposes. Currently, no tools exist to detect a reader's use of strategies. The primary aim of this study was to develop Read, Understand, Learn, & Excel, an automated tool designed to detect reading strategy use and explore its accuracy in detecting strategies when students read digital, expository text.
Method: An iterative design was used to develop the computer algorithm for detecting 9 reading strategies. Twelve undergraduate students read 2 expository texts that were equated for length and complexity. A human observer documented the strategies employed by each reader, whereas the computer used digital sequences to detect the same strategies. Data were then coded and analyzed to determine agreement between the 2 sources of strategy detection (i.e., the computer and the observer).
Results: Agreement between the computer- and human-coded strategies was 75% or higher for 6 out of the 9 strategies. Only 3 out of the 9 strategies–previewing content, evaluating amount of remaining text, and periodic review and/or iterative summarizing–had less than 60% agreement.
Conclusion: Read, Understand, Learn, & Excel provides proof of concept that a reader's approach to engaging with academic text can be objectively and automatically captured. Clinical implications and suggestions to improve the sensitivity of the code are discussed.
Supplemental Material: https://doi.org/10.23641/asha.8204786

Increasing numbers of students enrolled in postsecondary institutions face challenges with reading comprehension (National Research Council, 2012). Only 37% of students completing Grade 12 in the United States display reading levels proficient enough to engage in college-level coursework (U.S. Department of Education, 2015). Students with developmental learning disorders, such as attention-deficit disorder and specific language impairment, as well as acquired disorders such as brain injury and concussion, frequently experience cognitive difficulties that impair their ability to recall, integrate, and understand complex academic texts (DuPaul, Weyandt, O'Dell, & Varejao, 2009; Reaser & Prevatt, 2007; Sohlberg, Griffiths, & Fickas, 2015; Wolf, 2001). They often demonstrate sufficient reading comprehension for functional reading tasks such as comprehending text in magazines or social media; however, they struggle when faced with reading comprehension demands for the purposes of learning complex, new academic content (Sohlberg et al., 2015; Zabrucky & Moore, 1999). Of concern, students with reading comprehension deficits at postsecondary levels tend to have higher dropout rates, lower participation in the workplace, and lower earnings (Dodge, 2012; Kennedy, Krause, & Turkstra, 2008; Meulenbroek, Bowers, & Turkstra, 2016; Wilkins & Huckabee, 2014). Professionals face an urgent need to address impairments in reading comprehension at the postsecondary level.

Reading-to-learn is a complex process that requires extraction of information while engaging in continual construction, integration, and updating of concepts (C. Biancarosa & Snow, 2006). Established reading comprehension models such as the landscape model (Yeari & van den Broek, 2011) and the structure building framework (Gernsbacher, 1997) describe the cognitive processes that allow readers to hold

[a]Department of Communication Disorders and Sciences, University of Oregon, Eugene
[b]Department of Computer and Information Science, University of Oregon, Eugene

Correspondence to Priya Kucheria: priyak@uoregon.edu

Editor-in-Chief: Julie Barkmeier-Kraemer
Editor: Kristie Spencer

Received August 17, 2018
Revision received November 26, 2018
Accepted January 10, 2019
https://doi.org/10.1044/2019_AJSLP-18-0181

Disclosure: The authors have declared that no competing interests existed at the time of publication.

on to relevant text as they are reading and integrate it with previously read text or with existing background information, in order to build units of understanding. For example, working memory allows the reader to maintain and update concepts during reading (Yeari & van den Broek, 2011). The executive processes of enhancement and suppression are also integral to connectionist theories of comprehension and explain how the process of building coherent mental structures while reading depends upon spreading and connecting previously read or known information as well as inhibiting activation when information is not needed (Gernsbacher, 1997). These activation processes are usually automatic in skilled comprehenders. Recently proposed models such as the integrated model of reading comprehension (van den Broek & Espin, 2012) and the previously well-accepted landscape model of reading (van den Broek, 1995) also emphasize the role of conscious, strategic reading behaviors that enhance a reader's ability to create a coherent understanding of textual content. Behaviors such as looking back at preceding text, using text headers to create a mental outline of textual content, are all representative of strategies that the reader consciously initiates to build textual coherence (van den Broek & Espin, 2012). The appeal of these models is that they have been validated in both good and poor comprehenders (van den Broek & Helder, 2017). They describe the comprehension process in a way that suggests points of entry for intervention, particularly for the training of reading comprehension strategies (Griffiths, Sohlberg, Kirk, Fickas, & Biancarosa, 2016).

Individuals who display strong reading comprehension skills tend to have better working memory and executive functions than those who struggle with reading comprehension (Hannon, 2012). They also actively employ strategies that allow them to update and integrate information as they read in order to construct and maintain a coherent understanding across the text (Poole, 2014; Wigent, 2013). For example, proficient readers will use a strategy of attending to a table of contents or section headings as a way to orient themselves to the text organization and facilitate activation of schemas and construction of meaning (Graesser, 2007; Long & Chong, 2001). It follows that teaching strategies to compensate for weak working memory and executive functions and to enhance the ability to construct meaning from text is the strongest evidence-based approach for addressing reading comprehension deficits (Griffiths et al., 2016; Watter, Copley, & Finch, 2017). There are numerous studies showing that instructing struggling students in techniques to intentionally activate concepts and help them assess and monitor their understanding during reading facilitates comprehension and retention of content (Hall, 2004; B. E. Johnson & Zabrucky, 2011; Lei, Rhinehart, Howard, & Cho, 2009; Watter et al., 2017). Although the literature supports the training of reading comprehension strategies, there are no protocols describing which strategies will be optimal for which learners. We currently do not have methods to select strategies matched to the specific needs of postsecondary readers with comprehension challenges. Individual strategy selection would require the clinician or educator to detect strategies that a reader is and is not using. The purpose of this technical report is to explore methods that would allow interventionists to personalize reading strategy selection.

Commonly used tests of comprehension focus on product-oriented measures without relating it to a reader's process. Products of comprehension refer to the units of understanding constructed after a reader finishes reading a text (Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007). Typical product-based measurements of comprehension tend to include recognition tasks such as multiple-choice tests, open- or close-ended questions, and maze completion tasks found in standardized assessments such as the Nelson–Denny Reading Test (NDRT; Brown, Fishco, & Hanna, 1993) and Gray Oral Reading Test–Fifth Edition (Wiederholt & Bryant, 2012) as well as comprehensive batteries such as the Woodcock-Johnson IV Tests of Achievement (McGrew, LaForte, & Schrank, 2014). These tools serve as quick, easy-to-administer, norm-referenced screeners but are not useful for identifying techniques to improve reading comprehension, which requires measuring the reading process, in addition to the reading product (Kucheria, Sohlberg, Yoon, Fickas, & Prideaux, 2018; National Research Council, 2012).

Currently, clinicians and educators do not employ direct measures of the reading process as that would require observing and querying students while they are engaged in reading-to-learn tasks, which is not a feasible or objective clinical assessment activity. The alternative to direct observation of reading behavior is self-report via interviews and questionnaires such as the Learning and Study Strategies Inventory (LASSI; Weinstein, Palmer, & Acee, 2016) that ask students to describe their reading behaviors and rate their strengths and challenges. A limitation of this approach is that students often have limited insight into their own reading behavior that would allow the identification of strategies or interventions to improve comprehension (Bergey, Deacon, & Parrila, 2017).

Use of Digital Technologies to Measure Reading Comprehension Processes

Digital technology offers a potential solution to promote the detection of reading comprehension strategies. In postsecondary educational contexts, within the domain of reading comprehension, digital technologies have been primarily used as (a) electronic aids/supports (G. Biancarosa & Griffiths, 2012), (b) platforms to deliver instruction on strategy use (Griffiths et al., 2016; Magliano, Millis, The RSAT Development Team, Levinstein, & Boonthum 2011), and (c) product measures of comprehension (Graesser & McNamara, 2012). Automated procedures result in replicable protocols and algorithms that allow precision and efficiency when delivering content and analyzing performance. The bias and errors introduced by human observation are typically reduced, allowing for greater reliability.

Direct, digital measurement of a reader's process to create a consolidated understanding of the text in the college population has not yet emerged. However, a nascent body of research suggests that detection of overt strategic behaviors can be automated. For instance, Hyönä, Lorch, and Kaakinen (2002) focused on detecting strategic behavior in order to build cognitive models of reading and assess the actual text content for readability. Similar to our work, their interest was with adults reading expository text. They used eye-tracking technology to refine existing models of reading by measuring topic recognition within the text. Using eye tracking, they were able to capture four types of eye-gaze (fixation) patterns at the sentence level and support their hypothesis that readers employ different text reprocessing strategies. For example, some readers tended to employ frequent and longer "lookback" gaze patterns to previously read sentences, indicating a nonselective reprocessing strategy, whereas others were noted to have longer and frequent lookbacks toward the paragraph heading, indicating a topic heading reprocessing strategy. This electronically detected lookback strategy maps onto a behavioral strategy of "periodic review of material," which would be difficult to detect through human observation. Although the author's goal of developing a one-size-fits-all model of reading was not possible, eye tracking did shed light on reading behavior. Similarly, the work of Johnson-Glenberg (2005) used computer behavior to assess a reading strategy of reviewing previously read text. She used a web-based application to teach middle schoolers with poor reading comprehension to use metacognitive strategies, specifically rereading strategies. She hypothesized that training would increase the amount of "ScrollBacks" through text as measured by the number of mouse clicks, indicating that a reader's use of the rereading strategy had increased. Findings suggested that strategy training improved comprehension as evidenced by the significantly higher scores on a reading comprehension measure, and of interest to our work, readers in the experimental condition used significantly more ScrollBacks compared to readers in the control condition. ScrollBacks appeared to be a sensitive measure of reading strategy use.

Existing research suggests that eye tracking provides a potentially sensitive and direct estimate of the reading comprehension behavior that is linked to reading processes. However, the equipment demands limit the feasibility, accessibility, and affordability for implementation in clinical and educational settings; Hyönä et al. (2002) required sophisticated instruments and a multistep protocol to set up the equipment for each participant. In contrast, the work by Johnson-Glenberg suggests that it may be possible to leverage digital technology and use human–computer interaction (HCI) to create feasible measures of strategy use. We built on this work in our use of HCI to detect reading strategy use in postsecondary learners.

Mapping Reader Strategies to Computer Behavior

Our first development goal was to generate a theoretically grounded reading strategy framework that categorized the primary reading comprehension strategies important for postsecondary readers. This work has been summarized elsewhere in the development of a reading strategy training package (Griffiths et al., 2016). Our second development goal was to evaluate whether each of the identified strategies could be validly mapped to specific computer behaviors initiated by the reader. Once these two goals were achieved, we initiated the current study comparing computer-detected strategy use to human observer strategy identification.

Reading Strategy Framework

We conducted a literature search to identify the range of strategies reported to improve reading comprehension. Search terms included combinations of the following terms: *reading comprehension*, *reading strategies*, *strategy use*, *expository*, and *exposition*. Searches were done across five databases: PsycNET, Academic Search Premier, ERIC, Medline, and PubMed. Our search was limited to studies evaluating adult readers of expository text applicable to secondary and/or postsecondary settings and included 18 studies. Each of the studies was reviewed, and the individual strategies were extracted resulting in a total of 24 strategies that readers use to monitor their understanding, increase their ability to consolidate, and integrate large chunks of information across expository texts. Consistent with the constructionist theories of reading, these reader behaviors heighten activation and integration of concepts and therefore assist with the construction and retention of meaning from the expository text.

Strategies most commonly supported across studies included the following: previewing content prior to reading text, taking notes, conducting periodic reviews, summarizing content at the end of a reading session, and engaging in self-testing (Hall, 2004; Kobayashi, 2007; Lei et al., 2009; Ramsay, Sperling, & Dornisch, 2010; Reid & Morrison, 2014; Sohlberg, Griffiths, & Fickas, 2014; van den Bos, Nakken, Nicolay, & Van Houten, 2007; Ward-Lonergan & Duthie, 2016; Watter et al., 2017; Zabrucky & Moore, 1999). Less common strategies included teaching readers to plan and monitor their reading session (e.g., set a goal for the duration of reading time; state the purpose of the reading and evaluate amount of remaining text) and to focus their attention during reading (e.g., highlighting; Anmarkrud & Bråten, 2012; Bergey et al., 2017; Watter et al., 2017). A number of studies evaluated strategy packages with combinations of strategies (e.g., Think before reading, While reading, and After reading; State, Question, Read, Review, Recite; ; B. E. Johnson & Zabrucky, 2011; J. W. Johnson, Reid, & Mason, 2012: Linderholm, Therriault, & Kwon, 2014; McNamara, Boonthum, Levenstein, & Millis, 2007; O'Reilly & Sabatini, 2013; Ward-Lonergan & Duthie, 2016; Watter et al., 2017). Comparative analysis of the most effective group of strategies has not been completed and is not likely to be fruitful considering the heterogeneity of reader behavior (Hyönä et al., 2002). In general, however, the literature supports the use of strategies that help readers (a) activate prior knowledge and establish a schema for new information (e.g., previewing content, stating purpose of reading), (b) consolidate information (e.g., periodic review, note-taking in one's own words), and (c) encourage monitoring one's learning and retention (e.g., self-testing; Griffiths et al., 2016). As shown in Table 1, we distilled what appeared to be nine key evidence-based strategies into a chronological format implemented "before," "during," and "after" reading (Griffiths et al., 2016; Sohlberg et al., 2014).

Mapping Strategies to Computer Behavior

For the current study, we took each of the strategies in our framework and attempted to map them to the corresponding HCI behaviors. We observed readers and queried them in order to formulate a list of overt HCI behaviors, such as scrolling, mouse movement/clicking, typing, pausing, and highlighting, which corresponded to the target strategy. Some strategies, such as note-taking and highlighting, were direct and easy to capture. Other strategies, such as previewing content, had to be inferred. For instance, when we observed a reader scrolling and pausing at learning objectives, we assumed the reader was using a previewing content strategy. The previewing strategy was then coded as "scrolling until designated text is viewable in the scroll window and then pausing for greater than 5 seconds." We used pixels to track a reader's reading location and time stamps to map computer user behavior onto a strategy. For instance, if a reader scrolled to the end of the chapter and returned to his or her reading location within 5 s of being exposed to the chapter, then this rapid scrolling behavior was tagged as a "previewing" strategy. If the same scrolling pattern occurred when the reader was in the middle of the chapter, it was tagged by the computer as "evaluating amount of remaining text." Each strategy was similarly matched to a set of HCI behaviors and assigned an alphanumeric code. A singular code or a sequence of codes could represent a strategy. For instance,

Table 1. Reading strategy taxonomy.

Inferred strategy	Purpose	Computer behavior associated with strategy
Before reading		
Previewing scope	Plan and monitor	Rapid scroll (5 s or less) to the end of chapter, followed by return to the beginning
Previewing content	Activate background knowledge to establish a schema for new information	Mouse cursor hovers over learning objectives Scrolling to learning objectives section, followed by pause (10 s or more) Scrolling to the first heading or first sentence of the chapter, followed by pause (10 s or more) Click on notebook button and type notes Highlight learning objectives
During reading		
Highlighting key information	Focus attention during reading	Highlighting words/phrases/sentences
Evaluating amount of remaining text	Plan and monitor, self-regulate progress	Rapid scroll (3 s or less) to the end of chapter, followed by return to original reading position. Scrolling to unread sections, followed by quick (3 s or less) return to original reading position
Note-taking or annotating key information	Consolidate information	Taking notes while reading sections/subsections
Periodic review and/or iterative summarizing	Monitor learning and retention	Open and scrolling through notes before proceeding to unread sections Revising/editing previously taken notes Scrolling up to learning objectives and pausing (5 s or more), followed by returning to original reading position Scrolling up to previously read section (with or without pausing), followed by returning to original reading position Mouse hovering over previously read headings Creating a summary of notes taken for a section before continuing reading Scrolling through previously written notes after taking notes on current section Notes are opened and scrolled, followed by long pause (10 s or greater) inferred as verbally or mentally rehearsing notes while looking away from screen before proceeding to reading.
After reading		
Reviewing content	All the after-reading strategies serve to monitor learning and retention	"Consistent" scrolling rate (1 scroll event every 2 s or less) from end to beginning of chapter Generating notes Scrolling through notes
Summarizing and/or refining notes		Generating a grand summary of notes taken Opening notebook to edit notes
Self-testing		Open notebook function and scroll through notes with occasional pauses (5 s or more).

highlighting learning objectives and pausing for a few seconds in the "learning objectives" section when the chapter was presented before scrolling to the first heading of the chapter were all behaviors coded as the "previewing scope" strategy. A visual schemata illustrating steps from identifying to coding and matching behaviors to strategies is provided in Supplemental Material S2.

Preliminary validation of the HCI behaviors was conducted using an iterative design process that consisted of running small numbers of individual participants consecutively to test the code against human observation paired with querying readers: "Did you do anything special while reading to help you remember or understand the information better?" If the computer omitted a strategy that the human observer identified, we discussed whether the human had mistakenly identified a strategy or whether the computer had omitted it. If it was determined that the computer omitted it, we continued to refine the coding. If the computer identified a strategy that the human did not detect, we looked at other exemplars to determine whether it was a false positive or whether the human had missed the strategy. We observed readers until subjectively we had sufficient confidence that the strategies detected digitally matched the ones detected by the human evaluator. The final set of events captured by the computer corresponding to each strategy is presented in Table 1. These comprised the automated reading strategy detection algorithms evaluated in our study.

Purpose of Study

This exploratory study describes the use of a digital tool designed to detect reading strategies used by postsecondary readers when reading expository content. The purpose was to conduct a preliminary evaluation of a clinically feasible method for detecting the use of reading comprehension strategies. The ultimate goal of this work is to evaluate readers' strategy use in order to make intervention recommendations. Our exploratory research questions included the following:

1. Could we design and program a reading strategy detection tool based on HCI behavior?
2. How would results of computer-detected strategies compare to detection via direct observation by a human evaluator?

Method

The study was undertaken as part of a larger project, the Reading for Understanding and Learning to Excel (RULE), aimed at developing a digital reading comprehension tool to assess comprehension in postsecondary students. An extensive description of the recruitment procedures, administration protocol, and other components of the RULE measure has been outlined elsewhere (Kucheria et al., 2018). Below, we describe the participants and study procedures.

Participants

We recruited and consented typical, undergraduate readers in accordance with the regulations of our institutional review board protocol. Each participant was compensated with $30.00. Individuals were recruited if they met the following criteria: (a) in age range of 18–30 years, (b) self-identified as fluent speakers of English who had acquired the language before 7 years of age, (c) currently enrolled as full-time undergraduate students in a university or college, (d) able to read for daily needs (e.g., street signs, menus, and bills), (e) able to comprehend three- to four-paragraph length material at the 10th grade level, and (f) familiar with using a laptop for reading text, typing, and clicking and scrolling functions of a mouse. Exclusion criteria included (a) diagnosis of a disability or condition that affected basic reading abilities, (b) admission to a hospital or outpatient program in the last 12 months for substance abuse or psychiatric issues, and (c) inability to accurately demonstrate use of the interface features. There were 11 participants from the original sample that met criteria.

Tables 2 and 3 present participant characteristics. Participants were administered three different measures to gather descriptive data on their typical reading performance, study and learning habits, and comprehension skills. These measures included (a) the LASSI (Weinstein et al., 2016), (b) a demographic questionnaire designed by the authors, and (c) the NDRT (Brown et al., 1993). Of note, 55% of the sample reported that it was somewhat typical of them to forget information that they had just read. Mean scores fell below the 50th percentile on the following LASSI scales: information processing, time management, self-testing, attitude, motivation, concentration, and using academic resources scale. According to interpretation guidelines for this measure, the sample would benefit from training on learning and study strategies. Mean percentile score (M = 53) on the Comprehension subtest of the NDRT suggested that the sample consisted of "average" readers.

The range and frequency of strategies based on reader self-report were captured using a "think-aloud approach" (Schmitter-Edgecombe & Bales, 2005) and are shown in Table 4. Readers were asked "Did you do anything special while reading to help you remember or understand the information better?" after they completed a chapter. Readers most frequently reported using highlighting, note-taking or annotating key info, periodic review, and integrating learning by reviewing content.

Procedure

Participants came to a university clinic room and were seen individually. They were oriented to features of the digital RULE interface, including highlighting, accessing the note-taking feature, and scrolling bars by an evaluator. Participants read passages on a MacBook. The evaluator remained in the room and took notes on a laptop recording any observed digital behaviors executed by the reader using a formatted Excel document that was categorized in a

Table 2. Demographic questionnaire responses.

Items	I can read for extended periods of time	I forget information I have just read	I forget information after a delay
Not at all typical of me	0%	9%	0%
Not very typical of me	9%	27%	27%
Somewhat typical of me	0%	55%	64%
Fairly typical of me	55%	0%	9%
Very much typical of me	36%	9%	0%

Note. n = 11 readers.

before–during–after reading structure. The readers were videotaped with a camera capturing the computer screen. Evaluators included the first author and four undergraduate research assistants who were trained to observe and take notes on reading strategy behavior.

The reading stimuli were expository texts and consisted of two opening chapters extracted from introductory-level textbooks on public speaking and social psychology. Two different chapters were selected to control for specific reading strategy behavior biases that might arise from reader's background knowledge and topic interest. Reading stimuli in the two chapters were equated for length, semantic and syntactic complexity, text cohesiveness, and reading level across a variety of indices using latent semantic analysis (Kucheria et al., 2018). Each chapter was between 2,100 and 2,400 words in length and designed to be read within 30 min.

Study data consisted of the codes captured by the computer, video observation, and human observation notes. The computer captured the identified computer behavior sequences and generated a corresponding list of strategies associated with alphanumeric codes each time a reader completed reading a chapter. The human evaluators were trained to take notes in the form of a narrative. Training comprised two steps: (a) introducing evaluators to sample strategies that could be employed during these reading phases and (b) practice using the observation protocol and generating a narrative of reader behavior on practice participants. Details on the instruction provided to evaluators for accurately discerning strategies are provided in Supplemental Material S1. The first author provided feedback on the level of detail and gave tips for facilitating brevity in note-taking. After the evaluators completed running participants, their identified strategies were assigned the corresponding alphanumeric code, so that human- detected strategies could be compared to the HCI sequences generated by the computer. When notes were unclear or confusing, the human evaluator responsible for generating the notes was consulted to resolve doubts.

We used descriptive statistics to compare computer-generated and human-identified strategy codes. Percent agreement between the computer and humans was calculated for the initial detection of each strategy. The computer detection of a strategy was correlated to a precise time stamp, and the human observation of a strategy used was in rough time increments as noted by the observer. To calculate percent agreement, we used dummy codes to convert the raw data (i.e., alphanumeric codes) from each source into a common metric indicating agreement or disagreement. Thus, each strategy was assigned a code of either 1 (indicating that both sources agreed that the strategy was used by a reader in the relative time frame) or 0 (indicating that only one of the sources, i.e., only the computer or human, detected the strategy). These dummy codes were added across chapters and readers and divided by the total number of opportunities for strategy use (n = 24 instances of strategy use) to produce the percentage of agreement for that strategy. A higher percentage of agreement suggested that the computer and human agreed on the use of a particular strategy. When there was disagreement between the human and computer, we went back to review videos for 30% (n = 4 instances of strategy use) of the sample chapters to determine the source of the error. We were interested in whether the error was a false positive or a missed strategy and whether the error was made by the computer or the human observer. The first author analyzed videos for 72% of the sample (n = 18 instances

Table 3. Means and standard deviations of scores on the Learning and Study Strategies Inventory (LASSI) and the Nelson–Denny Reading Test (NDRT).

	LASSI							NDRT
Variable	Information processing	Time management	Self-testing	Attitude	Concentration	Motivation	Using academic resources	Comprehension subtest
M	42	25	29	48	37	45	39	53
SD	28	28	23	20	21	29	26	29

Note. n = 11 readers. Only scales where the mean performance was below the 50th percentile on the LASSI are displayed.

Table 4. Comparison of strategy detection by computer to human observer.

Reading strategy	Percent agreement	Human error	Computer error
Before			
Previewing scope	84%	6%	9%
	81%	**6%**	**13%**
Previewing content[a]	63%	25%	13%
	38%	**44%**	**6%**
Highlighting key info[a]	100%	0%	0%
	94%	**6%**	**0%**
During			
Evaluating amount of remaining text	31%	66%	3%
	19%	**25%**	**56%**
Note-taking or annotating key info	97%	3%	0%
	94%	**0%**	**6%**
Periodic review and/or iterative summarizing	53%	44%	3%
	50%	**19%**	**31%**
After			
Reviewing content	78%	9%	13%
	69%	**12%**	**19%**
Summarizing and/or refining notes	88%	3%	9%
	88%	**6%**	**6%**
Self-testing	97%	0%	3%
	100%	**0%**	**0%**

Note. n = 22 readers. The sample size refers to the number of opportunities for detecting a strategy. In our study, we had 11 participants. Each participant read two chapters, resulting in a total of 22 opportunities for detecting the presence or absence of each strategy. Computer error shows the percentage of opportunities where the computer missed a strategy. Human error shows the percentage of opportunities where the human evaluator missed a strategy. Results in bold display the agreement obtained when a third rater looked at 72% of the videos (n = 16). Computer error refers to instances when a computer misidentified or missed a strategy, and human error refers to instances when there was disagreement between human raters about the presence or absence of a strategy.

[a]Strategies that the computer was more likely to accurately detect and humans were likely to miss.

of strategy use) to identify reader strategies using a checklist that contained all possible behaviors. Strategy codes from the second human rater were used to determine if differences in strategy detection were more likely attributed to errors of the human evaluator versus computer.

Results

Table 4 presents the results of the strategy detection comparison. Agreement between the computer- and research assistant–coded strategies was 75% or higher for six out of the nine strategies. Only three out of nine strategies—previewing content, evaluating amount of remaining text, and periodic review and/or iterative summarizing—had less than 70% agreement. Examination of a subset of the video samples by a second human observer suggested that, for two of the three strategies (evaluating amount of remaining text and periodic review), the source contributing to disagreement stemmed primarily from computer error, that is, computer missing strategies that are both identified by all human observers. On further inspection, one of the digital strategy monitors was identifying scrolling patterns associated with a reader adjusting their reading position incorrectly as the strategy evaluating the amount of text that needed to be read. In essence, the sensitivity of the monitor was too high. For the third strategy, previewing content, video review revealed that disagreement stemmed from human error, that is, human missing the strategy that was accurately detected by the computer. This pattern of humans missing the strategy but the computer accurately detecting it was also noted for the highlighting strategy.

We also calculated Cohen's kappa to evaluate the degree of interrater between human observers and the computer algorithm. Data were aggregated across chapters and strategies, to include all instances of detection (n = 198 instances of detection). Cohen's kappa between the computer and undergraduate research assistants for the entire sample was .44, suggesting fair to weak agreement (McHugh, 2012). Kappa values were also calculated between various human observers and computer algorithm for a subset of the sample (n = 144 instances of detection). A similar trend of weak agreement between the computer and research assistants was noted for this sample as well, with a kappa value of .39. Kappa values between the first author and the computer for the same sample, however, was noted to be .63, indicative of moderately strong agreement. Post hoc power analysis in *G*power* indicated a 66% chance of detecting a small effect size (defined by Cohen, 1992, as a 0.20 difference between the means of two dependent groups when running t tests), when using a sample size of 144 and a significance level of .05 (two-tailed).

For each strategy, we also calculated the percentage of readers who implemented that strategy as measured by computer, human observation, and self-report. Table 5 presents these results. Data showed that regardless of the detection modality or chapter, four out of nine strategies were used

Table 5. Percent strategy use by readers.

Strategy	Computer detection: Public speaking chapter	Computer detection: Social psychology chapter	Human observation: Public speaking chapter	Human observation: Social psychology chapter	Self-report: Public speaking chapter	Self-report: Social psychology chapter
Previewing scope[a]	17%	17%	25%	25%	0%	0%
Previewing content	50%	58%	42%	17%	0%	0%
Highlighting key info	25%	33%	25%	33%	9%	9%
Evaluating amount of remaining text	92%	83%	17%	17%	0%	0%
Note-taking or annotating key info[a]	17%	8%	8%	17%	18%	18%
Periodic review and/or iterative summarizing	100%	67%	50%	50%	36%	9%
Reviewing content	75%	67%	83%	58%	18%	45%
Summarizing and/or iterative refining notes[a]	0%	0%	0%	17%	9%	9%
Self-testing[a]	0%	0%	8%	0%	0%	9%

Note. $n = 11$ readers. Numbers represent percentage of readers who used a strategy.

[a]Strategies that are least frequently used (i.e., used by less than 30% of readers) across chapters and measures.

by less than 30% ($n = 12$) of readers: previewing scope, note-taking or annotating key info, integrating learning by summarizing and/or iterative note-taking, and self-testing. The frequency of strategy use when measured via reader self-report was lower compared to measurement by human evaluator or the computer. This is consistent with the unreliability of self-reporting (Hux et al., 2010).

Discussion

The goal of this study was to explore the use of a digital tool to detect reading strategies used by postsecondary students reading expository text. Findings suggest that it is possible to capture computer behavior that corresponds to the implementation of specific reading strategies. We were encouraged by (a) the relatively high agreement between computer detection and human observation for the majority of the strategies evaluated and (b) the ease of implementation. Whereas some strategies, such as highlighting and note-taking, were easy to detect and capture because of the concrete digital imprints created while using them (typing, text formatting), other strategies required inference for detection. For example, computer detection of a reader evaluating the amount of remaining text involved correlating complex digital behavior (rapid sequences of scrolling to different areas of the screen) with time stamps of when the behavior occurred (after the reader had finished reading vs. while the reader was in the middle of the chapter). The iterative process of running small sample sizes and comparing computer-generated results against the human evaluator was integral in evolving the sensitivity and inferencing capacity of the code in detecting such strategies. Digital implementation made it convenient and practical to improvise this tool and improve its sensitivity.

Of surprise was that the computer appeared to be more accurate than the human observer at identifying two key strategies, previewing content and highlighting. We had anticipated that human observation would be the gold standard for strategy detection. The reason for wanting to automate the observation process was due to the impracticality of clinicians or educators observing students reading long passages, the inaccuracy of self-reporting, and the intrusion into the reading process of think-aloud procedures. Interestingly, we learned that there may be some strategies that are more accurately identified when automated.

Although strategies were used across readers and chapters, there was low agreement between the self-reported frequency of strategy use and frequency detected by the human evaluator or computer. The direct measure of strategy use thus seems to be an important evaluation component for planning supports that is not currently available in reading comprehension assessments. This finding also has important implications for clinicians relying on self-reporting with individuals with conditions associated with diminished self-awareness, affecting their ability to accurately report on their study habits and strategy use (Mealings, Douglas, & Olver, 2012; Sohlberg et al., 2015; Ward-Lonergan & Duthie, 2016). If RULE or related programs were shown to be reliable and sensitive, they could eliminate the need to rely on self-assessment.

This was an exploratory, developmental study and needs to be interpreted with caution. A major limitation of the study was the small sample size that may have been somewhat skewed in reading performance. Participant response to reading questionnaires suggested that the study sample consisted of relatively low readers with more than expected reported difficulty reading for long durations of time and remembering what they read. This could explain the somewhat reduced frequency of strategy use for a majority of the strategies. Readers of higher ability levels may have very different patterns of behavior, making it difficult to generalize our findings. Our own personal clinical experience suggests that skilled readers may not always employ overt strategies, which could also attribute to low strategy use. Next steps would be to collect data on a larger and more heterogenous sample in order to evaluate the impact of skill level on strategy detection.

Another explanation for the low agreement between computer and human observers for certain strategies could be attributed to the limited experience and strategy detection technique used by the research assistants. The higher kappa values between the first author and the computer algorithm compared to the undergraduate research assistants for the same set of data suggest that training humans to use a narrative description of strategies might undermine the validity and reliability of detection. It is also important to note that the human observers differed in their educational background and clinical experience (undergraduate vs. graduate observers), which could have impacted the outcome.

Lack of strategy use could also have affected the accuracy and sensitivity of the computer algorithm. For instance, even though percent agreement between the computer and human evaluator for the self-testing strategy was high, the finding can be misconstrued as accurate computer detection. High agreement could be attributed to readers not using the strategy frequently. For instance, the self-testing strategy was used by only one reader in the sample, a sample size that is insufficient to judge the accuracy of computer detection. In some cases, the element of digital versus nondigital strategy use could explain variations in computer versus human agreement. Not all strategic reading behaviors were digital, and therefore could not be captured. For instance, if a client rehearsed information verbally, this would be detected by humans but not the computer. Generating accurate and sensitive computer sequences is dependent on numerous instances of the same event that create a large enough data set for the computer to recognize replicable patterns.

Given this preliminary proof of concept, future studies are warranted. Such studies should employ large sample sizes to further test the validity and accuracy of the computer sequence in detecting strategies. Studies should measure all instances of strategy use in a sample. This implies that all behaviors that could be associated with a strategy need to be consistently captured. Another step that could improve the validity of these results is to determine agreement between the computer and the human evaluator on the types and frequency of associated computer behaviors that led to identification of the corresponding strategy. Using organized checklists to train the human observers would ensure accuracy in detecting behaviors and matching them to strategies. Using think-aloud or retrospective review with readers could help resolve discrepancies in what the computer versus human evaluator detect and improve reliability of the results.

Overall, the findings suggest that automated detection of reading strategy use is possible and feasible. The relatively high power associated with the agreement between digital and manual detection suggests that this work warrants continued research. The potential to identify reading process behaviors would be valuable for clinicians to plan reading supports for struggling readers and look at the efficacy of strategy training programs. It is hoped the findings of this early study lay the groundwork for future studies.

Acknowledgments

This study was supported by the National Science Foundation PFI-AIR TT Grant 1640492, awarded to McKay Moore Sohlberg, at the University of Oregon. The authors would like to thank all the participants who contributed to the study as well as Kayla Davis, Haley Hash, Elise Heater, Rachel Lacewell, Kelsey Lewey, and Garrett Porter who acted as research assistants.

References

Anmarkrud, Ø., & Bråten, I. (2012). Naturally-occurring comprehension strategies instruction in 9th-grade language arts classrooms. *Scandinavian Journal of Educational Research, 56*(6), 591–623. https://doi.org/10.1080/00313831.2011.621134

Bergey, B. W., Deacon, S. H., & Parrila, R. K. (2017). Metacognitive reading and study strategies and academic achievement of university students with and without a history of reading difficulties. *Journal of Learning Disabilities, 50*(1), 81–94. https://doi.org/10.1177/0022219415597020

Biancarosa, C., & Snow, C. E. (2016). *Reading next—A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

Biancarosa, G., & Griffiths, G. G. (2012). Technology tools to support reading in the digital age. *The Future of Children, 22*(2), 139–160. https://doi.org/10.1353/foc.2012.0014

Brown, J. I., Fishco, V. V., & Hanna, G. S. (1993). *Nelson–Denny Reading Test*. Rolling Meadows, IL: Riverside Publishing.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155.

Dodge, K. M. (2012). *Examining the lived experience of students with reading comprehension learning disabilities and the perceived value of the accommodations received* (Doctoral dissertation). Retrieved from ProQuest LLC database.

DuPaul, G. J., Weyandt, L. L., O'Dell, S. M., & Varejao, M. (2009). College students with ADHD: Current status and future directions. *Journal of Attention Disorders, 13*(3), 234–250. https://doi.org/10.1177/1087054709340650

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes, 23*(3), 265–304.

Graesser, A. C. (2007). An introduction to strategic reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies* (pp. 3–26). New York, NY: Erlbaum.

Graesser, A. C., & McNamara, D. S. (2012). Automated analysis of essays and open-ended verbal responses. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 1: Foundations, planning, measures, and psychometrics* (Vol. 1, pp. 307–325). Washington, DC: American Psychological Association.

Griffiths, G. G., Sohlberg, M. M., Kirk, C., Fickas, S., & Biancarosa, G. (2016). Evaluation of use of reading comprehension strategies to improve reading comprehension of adult college students with acquired brain injury. *Neuropsychological Rehabilitation, 26*(2), 161–190. https://doi.org/10.1080/09602011.2015.1007878

Hall, L. A. (2004). Comprehending expository text: Promising strategies for struggling readers and students with reading disabilities? *Reading Research and Instruction, 44*(2), 75–95. https://doi.org/10.1080/19388070409558427

Hannon, B. (2012). Understanding the relative contributions of lower-level word processes, higher-level processes, and working memory to reading comprehension performance in proficient adult readers. *Reading Research Quarterly, 47*(2), 125–152. https://doi.org/10.1002/RRQ.013

Hux, K., Bush, E., Zickefoose, S., Holmberg, M., Henderson, A., & Simanek, G. (2010). Exploring the study skills and accommodations used by college student survivors of traumatic brain injury. *Brain Injury, 24*(1), 13–26. https://doi.org/10.3109/02699050903446823

Hyönä, J., Lorch, R. F., Jr., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94*(1), 44–55. https://doi.org/10.1037//0022-0663.94.1.44

Johnson, B. E., & Zabrucky, K. M. (2011). Improving middle and high school students' comprehension of science texts. *International Electronic Journal of Elementary Education, 4*(1), 19–31. Retrieved from https://www.iejee.com/index.php/IEJEE/article/view/211

Johnson, J. W., Reid, R., & Mason, L. H. (2012). Improving the reading recall of high school students with ADHD. *Remedial and Special Education, 33*(4), 258–268.

Johnson-Glenberg, M. C. (2005). Web-based training of metacognitive strategies for text comprehension: Focus on poor comprehenders. *Reading and Writing, 18*(7–9), 755–786. https://doi.org/10.1007/s11145-005-0956-5

Kennedy, M. R. T., Krause, M. O., & Turkstra, L. S. (2008). An electronic survey about college experiences after traumatic brain injury. *Neurorehabilitation, 23,* 511–520.

Kobayashi, K. (2007). The influence of critical reading orientation on external strategy use during expository text reading. *Educational Psychology, 27*(3), 363–375. https://doi.org/10.1080/01443410601104171

Kucheria, P., Sohlberg, M. M., Yoon, H., Fickas, S., & Prideaux, J. (2018). Read, understand, learn, & excel (RULE): Development and feasibility of a reading comprehension measure for postsecondary learners. *American Journal of Speech-Language Pathology, 27*(4), 1363–1374.

Lei, S. A., Rhinehart, P. J., Howard, H. A., & Cho, J. K. (2009). Strategies for improving reading comprehension among college students. *Reading Improvement, 47*(1), 30–42.

Linderholm, T., Therriault, D. J., & Kwon, H. (2014). Multiple science text processing: Building comprehension skills for college student readers. *Reading Psychology, 35*(4), 332–356. https://doi.org/10.1080/02702711.2012.726696

Long, D. L., & Chong, J. L. (2001). Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(6), 1424–1429. https://doi.org/10.1037/0278-7393.27.6.1424

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*(2), 131–154. https://doi.org/10.1007/s11409-010-9064-2

McGrew, K. S., LaForte, E. M., & Schrank, F. A. (2014). *Technical manual. Woodcock-Johnson IV.* Rolling Meadows, IL: Riverside.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282.

McNamara, D. S., Boonthum, C., Levinstein, I., & Millis, K. (2007). Evaluating self-explanations in iSTART: Comparing word-based and LSA algorithms. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 227–241). Mahwah, NJ: Erlbaum.

Mealings, M., Douglas, J., & Olver, J. (2012). Considering the student perspective in returning to school after TBI: A literature review. *Brain Injury, 26*(10), 1165–1176.

Meulenbroek, P., Bowers, B., & Turkstra, L. S. (2016). Characterizing common workplace communication skills for disorders associated with traumatic brain injury: A qualitative study. *Journal of Vocational Rehabilitation, 44*(1), 15–31. https://doi.org/10.3233/JVR-150777

National Research Council. (2012). Improving adult literacy instruction: Options for practice and research. In A. M. Lesgold & M. Welch-Ross (Eds.), *Committee on learning sciences: Foundations and applications to adolescent and adult literacy, division of behavioral and social sciences and education.* Washington, DC: The National Academies Press.

O'Reilly, T., & Sabatini, J. (2013). Reading for understanding: How performance moderators and scenarios impact assessment design. *ETS Research Report Series, 2013*(2), i–47. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/j.2333-8504.2013.tb02338.x/abstract

Poole, A. (2014). Successful and struggling students' use of reading strategies: The case of upperclassmen. *Learning Assistance Review (TLAR), 19*(2), 59–80.

Ramsay, C. M., Sperling, R. A., & Dornisch, M. M. (2010). A comparison of the effects of students' expository text comprehension strategies. *Instructional Science, 38*(6), 551–570. https://doi.org/10.1007/S11251-008-9081-6

Rapp, D. N., van den Broek, P., McMaster, K. L., Kendeou, P., & Espin, C. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading, 11*(4), 289–312. https://doi.org/10.1080/10888430701530417

Reaser, A., & Prevatt, F. (2007). The learning and study strategies of college students with ADHD. *Psychology in the Schools, 44*(6), 627–638. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/pits.20252/abstract

Reid, A. J., & Morrison, G. R. (2014). Generative learning strategy use and self-regulatory prompting in digital text. *Journal of Information Technology Education: Research, 13,* 49–72. Retrieved from http://jite.org/documents/Vol13/JITEv13ResearchP049-072Reid0549.pdf

Schmitter-Edgecombe, M., & Bales, W. B. (2005). Understanding text after severe closed-head injury: Assessing inferences and memory operations with a think-aloud procedure. *Brain and Language, 94,* 331–346.

Sohlberg, M. M., Griffiths, G. G., & Fickas, S. (2014). An evaluation of reading comprehension of expository text in adults with traumatic brain injury. *American Journal of Speech-Language Pathology, 23,* 160–175. https://doi.org/10.1044/2013_AJSLP-12-0005

Sohlberg, M. M., Griffiths, G. G., & Fickas, S. (2015). An exploratory study of reading comprehension in college students after acquired brain injury. *American Journal of Speech-Language Pathology, 24*(3), 358. https://doi.org/10.1044/2015_AJSLP-14-0033

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP). (1992–2015). *Reading assessments.* Retrieved from https://www.nationsreportcard.gov/reading_math_g12_2015/#reading

Van den Bos, K. P., Nakken, H., Nicolay, P. G., & Van Houten, E. J. (2007). Adults with mild intellectual disabilities: Can their reading comprehension ability be improved? *Journal of Intellectual Disability Research, 51*(11), 835–849. https://doi.org/10.1111/j.1365-2788.2006.00921.x

Van den Broek, P. (1995). A "landscape" model of reading comprehension—Inferential processes and the construction of a stable memory representation. *Canadian Psychology-Psychologie Canadienne, 36*(1), 53–54.

Van den Broek, P., & Espin, C. A. (2012). Connecting cognitive theory and assessment: Measuring individual differences in reading comprehension. *School Psychology Review, 41*(3), 315–325.

Van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes, 54*(5–6), 360–372.

Ward-Lonergan, J. M., & Duthie, J. K. (2016). Intervention to improve expository reading comprehension skills in older children and adolescents with language disorders. *Topics in Language Disorders, 36*(1), 52–64. https://doi.org/10.1097/TLD.0000000000000079

Watter, K., Copley, A., & Finch, E. (2017). Discourse level reading comprehension interventions following acquired brain injury: A systematic review. *Disability and Rehabilitation, 39*(4), 315–337. https://doi.org/10.3109/09638288.2016.1141241

Weinstein, C. E., Palmer, D. R., & Acee, T. W. (2016). *LASSI user's manual: Learning and study strategies* (3rd ed.). Clearwater, FL: H&H.

Wiederholt, J. L., & Bryant, B. R. (2012). *Gray Oral Reading Test–Fifth Edition: Examiner's record booklet form-A*. Austin, TX: Pro-Ed.

Wigent, C. A. (2013). High school readers: A profile of above average readers and readers with learning disabilities reading expository text. *Learning and Individual Differences, 25,* 134–140. https://doi.org/10.1016/j.lindif.2013.03.011

Wilkins, J., & Huckabee, S. (2014). A literature map of dropout prevention interventions for students with disabilities. Clemson, SC: National Dropout Prevention Center for Students with Disabilities, Clemson University.

Wolf, L. E. (2001). College students with ADHD and other hidden disabilities. *Annals of the New York Academy of Sciences, 931*(1), 385–395. https://doi.org/10.1111/j.1749-6632.2001.tb05792.x

Yeari, M., & van den Broek, P. (2011). A cognitive account of discourse understanding and discourse interpretation: The landscape model of reading. *Discourse Studies, 13*(5), 635–643.

Zabrucky, K. M., & Moore, D. W. (1999). Influence of text genre on adults' monitoring of understanding and recall. *Educational Gerontology, 25*(8), 691–710. https://doi.org/10.1080/036012799267440