# Detecting reading strategies during task-oriented reading: Building an automated classifier

George Kachergis[*]
Donders Institute for Brain, Cognition, and Behaviour
Department of Artificial Intelligence, Radboud University
G.Kachergis@donders.ru.nl

Jolique Kielstra[*]
Behavioural Science Institute
Department of Educational Sciences, Radboud University
Nijmegen, the Netherlands
J.Kielstra@pwo.ru.nl

Lars Bokkers
Department of Artificial Intelligence, Radboud University
Nijmegen, the Netherlands

Bellamie Persad
Department of Artificial Intelligence, Radboud University
Nijmegen, the Netherlands

Inge Molenaar
Behavioural Science Institute
Department of Educational Sciences, Radboud University

## ABSTRACT

Secondary school students are continuously asked to read texts and execute tasks related to these texts. Many students experience difficulties in reading to successfully perform these tasks. Task-oriented reading is conceptualized as an adaptive problem-solving process in which readers engage with the text selectively based on the task. Usage of appropriate reading strategies leads to more efficient and accurate task performance. Our aim is to provide students with personalized support to effectively select reading strategies. A first step towards personalized support is the automatic detection of students' reading strategies. This study describes the development of a supervised machine learning classifier to detect students' reading strategies. Human raters classified 1,091 graphs of students' behavior recorded on tablets as they engaged in task-oriented reading. These ratings were used to train a classifier on 13 features extracted from the students' reading behavior. The overall accuracy for classifying reading strategies was 0.74, significantly greater than chance. Searched reading strategies were the easiest to identify, with a balanced accuracy of 0.84, followed by intensive (0.81) and targeted reading strategies (0.69). The most important features in the classifier were the ratio of sentences that readers skimmed too quickly, the number of unique sentences read, and the variance of time spent reading each sentence. These features are quite different from typical process variables used to study task-oriented reading, yet are easy to automatically extract in tablet-based reading. This classifier is a first step in the development of personalized support based on students' use of reading strategies.

## KEYWORDS

reading strategies, task-oriented reading, automatic classifier

[*]Equal contribution.


*PALE'2018, June 2018, London, UK*


## 1 INTRODUCTION

Task-oriented reading involves reading with the purpose of processing information for the execution of a specific task [2, 27]. During task-oriented reading, students must do more than read and comprehend the text: the focus is on executing the task connected to the text. Thus, students need to appropriately understand the task and select a reading strategy aligned with the task [27, 28]. The RESOLV model [21] puts students' task *representation* in a central position to explain strategy selections students make. Prior to reading, a task can act to signal relevance, allowing readers to understand which sections of the text are relevant for the task [14]. This allows students to select a reading strategy that supports successful accomplishment of the task [21, 22, 29]. Differences in tasks induce different reading strategies. For example, simple tasks ask students to locate the needed information in the text with a *search* strategy, whereas complex tasks, that require integration of different parts of the text, elicit *intensive* reading [5, 22]. Reading strategies are methods students use to understand and process information presented in a text. Decision-making regarding which reading strategy to apply depends on student characteristics, task objectives and text objectives [1]. In summary, task-oriented reading is the ability to successfully create a task representation in order to select the most appropriate reading strategy for a given task [9, 10].

Task-oriented reading research shows that task performance is influenced by how students apply reading strategies [2, 7, 13]. For example, Rouet and colleagues [22] showed that differences in students' reading comprehension can be explained by their application of reading strategies during task execution. Previous research indicates that proficient readers tend to use more diverse reading strategies compared to less proficient readers [23], and that they apply these reading strategies more frequently [15]. Less proficient readers, such as students in vocational secondary education, experience difficulty in applying reading strategies that support successful task execution [4]. They often resort to two different compensatory approaches during task-oriented reading. Either they start reading a text before developing a task representation or they immediately start with task execution by scanning the text before reading [15]. Moreover, less proficient readers tend not to change their approach and reading strategies even when they prove to be unsuccessful, whereas more proficient readers adapt their approach to changing perceptions of task execution [15].

Thus, in order to enhance successful task-oriented reading, less proficient students need help in task understanding, strategy selection and applying reading strategies to fulfill their tasks [3]. In order to support students selection of reading strategies, it is of great importance to detect the actual reading strategy students use. The automatic detection of reading strategies can guide the development of interventions to improve students' task-oriented reading. Although ample research provides evidence for the importance of reading strategies during task-oriented reading, it does not inform us how students' apply these reading strategies. An increased understanding of how students' select and apply reading actions and how these form reading strategies is needed to further understand reading as an *"adaptive, problem-solving process whereby readers engage with text selectively based on their self-generated goals and plans"* as stated by Rouet [21]. Automatic detection of reading strategy used during task execution could help develop advanced forms of support for students. Below we discuss how reading strategies are currently measured.

## 1.1 Measurement of reading strategies

Generally we see three different approaches to the measurement of students' reading strategies. First, researchers have used computers to measure students reading time. Students click to receive the consecutive sentence, allowing researchers to record the reading time per sentence [14]. These measurements are typically used to assess how relevant instructions affect the reading time of relative sentences. Second, think-aloud procedures are used, in which students are asked to read aloud and also verbalize their thoughts while reading [6]. This method is used to understand how students regulate during reading. A special approach of thinking aloud is the bridging method, in which students are asked at a regular interval (e.g., every 2 minutes) to indicate what they are thinking. Based on these utterances, researchers analyze how and which inferences students make during reading [12]. Third, reading research has been using masking as a means to follow how student read a text [27]. Task-oriented reading processes are measured with a tool called *Read&Answer*, by masking the text, with the exception of a text part deblurred by the student. This procedure allows Read&Answer to capture how students read a text. So far these data have been used to derive a number of process variables which are informative of the student's reading process. Both regulation and reading indices are derived. Examples of the indices are time for initial reading of the text, number of rereadings during initial reading, and number of search decisions.

The above-described online measurements of the reading process can all be classified as variable-based approaches, which uses online data to construct variables as indicators of particular reading actions such as the number of search decisions or the time spend on relevant text. These variable-based approaches focus on the analysis of variance between independent and dependent variable(s), such as the association between search decisions and task performance [13]. Variable-based approaches mainly focus on reading actions, such as rereading, time spend on task and monitoring decisions by checking the text [7, 13]. In contrast, event-based approaches analyze the (dynamic) relations between events [19]. Researching the nature of relations between events and their development over time is central in this approach. For example, such an approach can indicate how different reading actions together form a reading strategy. This allows for insights into temporal characteristics of reading actions, as well as how different actions interplay over time and form a reading strategy. Consistency and change in reading actions can be investigated by specifying these temporal characteristics [24]. It is important to emphasize that event-based approaches are a deviation from the traditional research paradigm used in prior reading research [19, 24].

A clear distinction can be made between two dimensions of time that are useful for the field of reading research, i.e. focusing on individual events within the continuous flow of events or relative arrangements of multiple events (Molenaar & Wise, in prep). So far, when using online data, mainly frequency analysis indicating the number of occurrences of a variable during a particular time window are used. This provides insights into the prevalence of a particular reading actions during learning. For example, Rouet et al. [22] found that tasks that demand explicit information from a text elicit more rereading among students. This analysis showed the significance of rereading, however we do not know whether the position of these actions in the reading process matters, nor if the duration of these actions or the rate at which these rereading actions occur during reading matter. Thus frequency analyses treat the reading process as one holistic unit, ignoring the individual time-related characteristics of variables. However, the first dimension of time, 'individual events within the continuous flow of events', captures how variables behave by examining the individual time-related characteristics of events within the flow. These individual time-related characteristics can illustrate how events occur within the flow of continuous events in a particular time window by analyzing the significance of the position of events, the duration and rate at which particular events occur within the reading process. For example, poor readers reread the task more often when they are reading the text, while more proficient readers spend more time initially reading the question before reading the text and less time rereading the question when reading the text [7].

In contrast the second dimension of time, 'the relative arrangements of multiple events in time', focuses on how events are organized among each other. An example is how the combination of multiple reading actions form an arrangement that can be recognized as a reading strategy over different settings. The second dimension of time conceptualizes how reading actions behave in relative arrangements of multiple events by examining the organization of these actions. For example 'Low-level' questions that focus on retrieving a single concept and demand for little inferencing often initiate a search strategy referred to as "locate-and-memorize" [20]. During this searched reading strategy students browse several paragraphs quickly and select a small number of elements in those paragraphs to answer the question. On the other hand, 'high-level' questions that comprise multiple concepts and require students to integrate multiple elements of the text are referred to as "review-and-integrate" [20]. These questions elicit a targeted reading strategy, in which students search the text to read larger text parts intensively to process information relevant to the task. Whereas more elaborate tasks, such as making a summary, require students to fully read the text and therefore elicits an intensive reading strategy. In these elaborate tasks, students infer the task solution from the text as a whole.

In order to push the conceptual understanding of reading as an adaptive problem solving process, it is important to understand how students' reading actions evolve into reading strategies. This supports conceptual clarity among researchers engaging in process analysis and deepens the scientific debate around reading as an adaptive problem solving process. Furthermore, it can be used to unravel characteristics of the reading processes and how these interrelate with student, task and context characteristics.

### 1.2 The Present Study

In this study, an event-based approach is taken to detect students' reading strategy based on the pattern of reading actions. We use a method inspired by Read&Answer in which students need to unmask (i.e. deblur) individual sentences during reading in order to read the text. This allows us to identify students' reading actions, such as sentences opened, read, time spent on relevant sentences, checking, rereading and task monitoring actions. These reading actions are visualized to understand how reading actions are informative of students' reading strategies use. When students read intensively, we expect they will deblur every sentence linearly and open sentence long enough to read the sentence. In case of a search reading strategy, sentences are opened only shortly to identify the information and whether this is relevant for the current task. Finally, in a target reading strategy the student looks for the relevant section of the text which is then opened long enough to read. Accordingly, we aim to create a machine learning classifier to detect students' reading strategies from reading actions in the trace data.

We followed a two-tiered process to fulfill this aim. We visualized students' reading behavior in graphs, and interactively refined them to support human coding of the three reading strategies (i.e. searched, targeted and intensive). In the process additional reading actions (i.e. checking, re-reading, tab switched to task and stopped reading after relevant sentence) were derived, to see how they are related to one of the three reading strategies. Next, all graphs were classified by humans and used as a supervisory signal to train the machine learning classifier. Thus, we aim to address two research questions: 1) Can we (automatically) detect students reading strategies from the pattern of reading behavior over time? and 2) How do reading actions relate to students' reading strategies?

## 2 METHOD

### 2.1 Participants

In this study, 44 fourth-year vocational secondary training students (20 female; 24 male) participated and collected data are part of a design study into the effects of reciprocal peer tutoring on task-oriented reading. Students had an average age of 15.56 years (SD=0.65), and gave active consent after parental consent was given. Students participated in task-oriented reading during three economics classes.

### 2.2 Measurements

Each student completed three task-oriented reading sessions at four-week intervals. All sessions were similar in set-up, with the only difference being an added example task in the first session. Each session consisted of 9 text-task pairs appropriate for vocational secondary students.

### 2.3 Text

The texts were all informative texts with an average length of 21 sentences per text, and a range of 208 to 516 words. Each text had three paragraphs discussing environmental sustainability topics, at a complexity and difficulty level evaluated by two experts on vocational secondary education.

### 2.4 Task

Tasks were developed as text-based, bridging or elaboration tasks [18]. This categorization was based on two scales namely the number of relevant sentences needed to execute the task and the required level of inferencing between these elements [12, 22, 25]. In text-based tasks, answers can be found within one sentence, bridging tasks require making inferences about relations between two or more sentences, and elaboration tasks require a great deal of inferencing across the text and sometimes making connections with students' own prior knowledge [18]. These different task types are commonly found in reading comprehension measurement instruments, such as PISA [17] and NT2 [22, 25]. Executing reading tasks implies a process of categorization that involves identifying the type of task and relevant information sources in a text [5, 8, 11, 16, 22, 26]. The instruments all consisted of 3 text-based, 3 bridging and 3 elaboration tasks.

### 2.5 Procedure per session

At the beginning of each session the students were given ten statements which they had to answer as being correct or incorrect. The statements were masked and before answering students had to tap on the sentence to make them visible. These taps were timestamped to estimate a reading calibration time per student. Next, in the first session students were given an example text and task to familiarize and practice the procedure. Entering the full task, students are given the first out of nine tasks. Students first saw the planning page, where they answered questions about the task (e.g., difficulty, level of inference, strategy decision and number of relevant sentences). During the planning phase students could move freely between the text and the task, upon answering the questions on the planning page. Students were able to see the complete task, but apart from the headers, the text was masked and could not be deblurred. After answering the planning questions, students would enter the execution phase. During this phase students were transferred to the text but could, at any time, switch to the task tab. After task completion, students saw the reflection page asking which reading strategy they used and which would have been best for the task. The planning, execution and reflection cycle would start again for the following eight tasks and texts of the test.

### 2.6 Graphs

The visualizations showed the index of the sentence opened in the text on the y-axis and time spent on task-oriented reading on the x-axis. An example graph is shown in Figure 1. Since data was collected in a classroom setting we used masking to gain information about students reading process. To read the masked text, students had to tap on a sentence. Hence providing information about the order in which sentences were opened, the relevance of the sentence opened to complete the task and the time a sentence was opened. The estimated calibration time gained from the ten statements at the beginning of each test, provided information about whether a sentence was opened long enough to read. In addition, the time spend on text and task was logged. The time spend on text and task, the average calibration time, order of sentences opened, relevancy of sentences and reading time was then used to visualize students reading process in graphs.

Each deblurred sentence was visualized either as a triangle, indicating that the opening time of the sentence was too short to read the sentence or as a circle when opened long enough to read (based on the estimated calibration time). The size of the plot character indicated the deblurring time of the sentences, with larger plot characters representing a longer opening time. Blue and red plot characters were used to indicate sentences with relevant and irrelevant information for the task, respectively. Vertical lines indicated a switch from the text to the task.

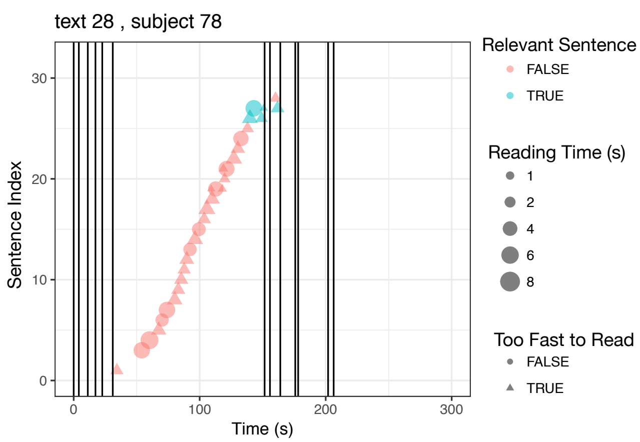


**Figure 1: Example of the graphs used by human raters to detect a reader's strategy and reading actions for each text.**

## 2.7 Human Classification

Based on visualizations of students reading behavior, human coders classified students' reading strategies as searched, targeted, or intensive, as described below. A total of 1091 graphs were classified. The graphed features used to code these reading strategies were opening time of sentences (e.g. triangle/circle), relevance of sentence opened (e.g. blue/red), and the number and order of sentences opened.

In a *searched* reading strategy, students browse several paragraphs quickly and select a small number of elements in those paragraphs to answer the question. Thus, sentences might be opened linearly or haphazardly, but most sentences are not opened long enough to read (i.e. triangle). When students find a relevant sentence (blue), it should be opened long enough to read (circles). During searched reading this usually applies for only a few sentences.

In a *targeted* reading strategy, students search for keywords and then read larger sections intensively to gather task-relevant information. Reading often starts from the second or third paragraph, which is then read intensively. This is seen as a cluster of relevant and irrelevant sentences opened long enough to read (i.e. circle). Students may instead quickly go through each sentence from the start of the text before focusing on a cluster of relevant sentences.

An *intensive* reading strategy involves reading the text carefully. This is seen in the graphs as a linear opening of sentences, viewed long enough to have been read (i.e. circles). Students may either read the text completely or stop after the relevant text parts to complete the task.

Additionally a number of reading actions were coded based on the graphs (see Table 1). Seventy-three graphs showed that students read no sentences, so these were excluded from human and automatic classification. Ten percent of the graphs were coded by two coders to establish inter-rater reliability.

**Table 1: Coding features for reading actions**

| Variables | Description of student behavior |
|---|---|
| Checking | Re-opens sentences too short to read. |
| Re-reading | Re-opens sentences long enough to read. |
| Switched tabs | Switches between the task and text. |
| Relevant stop | Stops reading after finding task-relevant information. |

## 2.8 Automatic Classification

To determine what behavioral features define a reading style–and how accurately such classifications can be made–we calculated a collection of features from the logged behaviors of each participant reading each text and trained a machine learning classifier to predict the 1018 human-classified reading strategies of students.

## 2.9 Features

Fifteen features were extracted from participants' recorded reading behavior on each text, including total number of sentences read, the number of unique sentences read, the number of re-read sentences, and the number of times the participant switched tabs while reading. Also included were the location in the text of the first and last sentence indices that were read (*First Sentence Read* and *Last Sentence Read*), scaled to $[0, 1]$. A linear regression was fit to each text's sentence indices and each reader's tapped indices, and this model's *slope* and $R^2$ were included. The reader's total time spent reading the text (*Time-on-Task*), and the standard deviation of the time they spent reading each sentence (*sd(Reading Time)*) were included, as well as their median per-sentence calibration time (*Subject Calibration RT*). *Proportion Fast Reading* is the proportion of read sentences that the user spent less time on than the mean of all subjects' median calibration times (3210 ms).[1] Three features relating to the text were included: the unique ID of the text, and both *Number/Proportion of Relevant Sentences Read* measured the reader's taps that were on sentences relevant to the answer.

## 2.10 Classifier

A tree-based XGBoost (eXtreme Gradient Boosting) classifier was built in R to classify reading strategy (searched, targeted, or intensive) based on the 15 features calculated above for each participant and text. This algorithm was chosen because it learns simple, interpretable tree models that when boosted (i.e., combined), offer

[1]Using *Subject Calibration RT* instead of the group average calibration time significantly reduced the classifier's accuracy.

**Table 2: Confusion matrix for reading strategy classifier.**

| | Reference | | |
|---|---|---|---|
| **Predicted** | Searched | Targeted | Intensive |
| Searched | 485 | 42 | 23 |
| Targeted | 81 | 107 | 63 |
| Intensive | 22 | 32 | 214 |

**Table 3: Reading strategy classifier statistics.**

| | Searched | Targeted | Intensive |
|---|---|---|---|
| Precision | 0.88 | 0.43 | 0.80 |
| Recall | 0.82 | 0.59 | 0.71 |
| F1 | 0.85 | 0.50 | 0.75 |
| Prevalence | 0.55 | 0.17 | 0.28 |
| Detection Rate | 0.45 | 0.10 | 0.20 |
| Detection Prevalence | 0.51 | 0.23 | 0.25 |
| Balanced Accuracy | 0.84 | 0.71 | 0.82 |

excellent fit balanced with good generalization, achieving state-of-the-art performance in many domains. The classifier was trained with the multiclass softprob objective function, evaluated with multiclass logloss, and limited to a maximum depth of 3. Parameters were chosen after examination of 10-fold cross-validation results.

A second, text-naive classifier was constructed using the subset of 12 features that depended on only reading behavior, and not on details of the texts. Using these same 12 behavioral features, binary classifiers were built for each reading action (checking, rereading, tab-switching, and relevant stopping), using a binary logistic objective function and with error as the evaluation metric.

# 3 RESULTS

## 3.1 Reading Strategy Classifiers

The reading strategy classifier achieved an overall accuracy of 0.75 (95% CI = [0.73, 0.78], kappa = 0.59), significantly greater than the no information rate (0.55; i.e., guessing the most frequent class). A confusion matrix is shown in Table 2. Table 3 shows performance statistics of the classifier on each reading strategy. A searched reading strategy was the easiest to identify, with a balanced accuracy of 0.84, followed by intensive (0.82) and targeted (0.71).

The most important features for the classifier were *Proportion Fast Reading* (gain = 0.50), *Unique Sentences Read* (0.11), and *sd(Reading Time)* (0.06). *First Sentence Read* and *text ID* yielded gains of 0.05, while *Time-on-Task* and *Linear Model* $R^2$ each offered 0.04. *Total Sentences Read*, *Number of Tab Switches*, and *Linear Model Slope* had gains of 0.03, and *Subject Calibration RT* and *Proportion of Relevant Sentences Read* each offered 0.02. *Number of Relevant Sentences Read* gave a gain of .01, while *Number of Reread Sentences* and *Last Sentence Read* yielded gains < .01.

To determine how well reading strategy might be determined without specific knowledge of the text, a second classifier was created using only the 12 features related to readers' behavior (i.e., without *text ID, Number/Proportion of Relevant Sentences Read*. This text-naive classifier achieved an overall accuracy of 0.73 (95% CI = [0.70, 0.76], kappa = 0.55), with balanced accuracies for each style quite similar to those in the 15-feature classifier (searched = 0.82, targeted = 0.68, and intensive = 0.81).

As shown in Figure 2, the most important features for this text-naive strategy classifier were largely the same as for the classifier

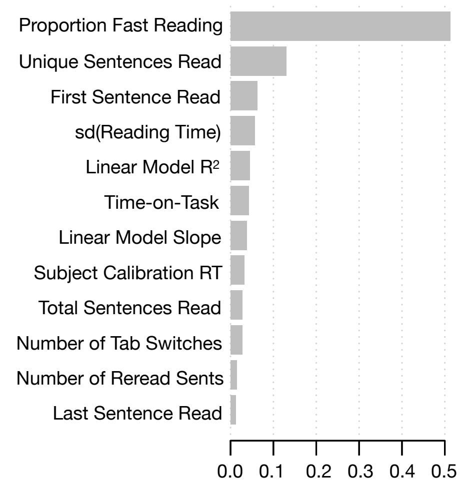


**Figure 2: Feature importance for the text-naive reading strategy classifier.**

with text-based features. The most important features for the classifier were *Proportion Fast Reading* (gain = 0.51), *Unique Sentences Read* (0.13), followed by *First Sentence Read* (0.06) and *sd(Reading Time)* (0.06). *Linear Model Slope,* $R^2$, and *Time-on-Task*, all yielded gains of 0.04, while *Subject Calibration RT*, *Total Sentences Read*, and *Number of Tab Switches* each offered 0.03. Finally, *Number of Re-read Sentences* and *Last Sentence Read* offered gains of .01.

## 3.2 Reading Action Classifiers

The checking classifier achieved a test accuracy of 0.9, relying mostly on *numRelevantRead* (gain=.18) and *numReread* (.09). The rereading classifier achieved a test accuracy of 0.77, relying on *Rsq* (gain=.49), *numReread* (.14) and *timeOnTask* (.09). The tab-switching classifier reached a test accuracy of 0.83, with important features being *numSwitchTabs* (.43) and *Rsq* (.25). The relevant stop classifier had a test accuracy of 0.73, relying on *numSwitchTabs* (.26), *propRelevant* (.20), *uniqueSentRead* (.12), and *lastIndexProp* (.11).

Using the human-classified reading actions as additional features alongside the original 12 calculated behavioral features, we trained a final classifier to see if these actions contribute to the reading strategy decision. With an overall accuracy of 0.75 (95% CI = [.72,.77], kappa = .58), the results were quite similar. Checking, rereading, and relevant stopping were the three least important features (gains < .01), and tab-switching was the 11th-most-important feature (gain = .02). In short, the four hand-coded reading actions yielded little benefit beyond the 12 automatically-extracted features.

# 4 DISCUSSION

The automated classifier achieved reasonably high accuracy on the two more common reading strategies, searched and intensive, while targeted reading presented some difficulty. Despite using different features of reading behavior than the automated classifier, human coders were also easily able to recognize searched and intensive reading strategies. However, targeted reading strategies also proved difficult to identify for human coders. The features that were most important to the human coders were the linear or non-linear opening of sentences, duration of opened sentence (too fast to read or opened long enough to read), the starting point of reading and whether sentences provided relevant information to students to solve the task. While the automated classifiers relied on the proportion of read sentences that were viewed too quickly (*Proportion of Fast Reading*), the number of unique sentences that were read, the index of the first sentence read in the text, the variance of the per-sentence reading time, and the $R^2$ of the linear regression (sentence index read vs. time). Using only these top five behavioral measures as features already yields a classifier with an accuracy of 0.73. At least three of these features are not used by nor easily discernable to human raters: the variance of the per-sentence reading time, the number of unique sentences read, and *Proportion of Fast Reading* are all difficult to visually estimate with much precision. This suggests that reading strategies such as searched, targeted and intensive reading can be measured automatically, and perhaps more accurately than by human coding, since the classifier can use more features of greater complexity to determine a reading strategy. However, both human coders and the classifier had difficulty identifying targeted reading, suggesting a need for further research.

Surprisingly reading actions, such as checking, rereading, tabs switched and relevant stop (see Table 1), did not help to classify the reading strategies. Rather, the automated classifier was able to classify the reading actions separately from the reading strategies. Thus it appears that even though these actions appear during reading, they are not necessarily representative of particular reading strategies.

Importantly, the classifier did not perform much better with the addition of either the features requiring specific knowledge of the text (e.g., the text ID or which sentences are read) nor with the addition of human-coded reading actions. Thus, we suggest it would be fruitful for researchers to first use a classifier trained on automatically extracted behavior-based features. These machine classifications can then be inspected by human raters, and corrected if need be. Such human-in-the-loop classification can yield lower error rates than machine- or human-only systems. Future work should investigate whether defining more complex behavioral features can further improve classification, especially for the difficult and relatively rare targeted reading strategy. Moreover, the relation between task performance, reading actions and strategies will be explored to assess the efficacy and efficiency of different strategies for task-oriented reading. This work will inform development of an intervention to support students' reading strategies during task-oriented reading.

## ACKNOWLEDGMENTS

Thanks to NVIDIA for a grant of GPU hardware to G. Kachergis.

## REFERENCES

[1] Cor Aarnoutse and Gonny Schellings. 2003. Learning reading strategies by triggering reading motivation. *Educational Studies* 29, 4 (2003), 387–409.
[2] Øistein Anmarkrud, Matthew T McCrudden, Ivar Bråten, and Helge I Strømsø. 2013. Task-oriented reading of multiple documents: Online comprehension processes and offline products. *Instructional Science* 41, 5 (2013), 873–894.
[3] Sheri Berkeley, Margo A Mastropieri, and Thomas E Scruggs. 2011. Reading comprehension strategy instruction and attribution retraining for secondary students with learning and other mild disabilities. *Journal of Learning Disabilities* 44, 1 (2011), 18–32.
[4] Egbert Helge Bonset, Hans de Vries, and Edith Alladin. 2009. *Talige startcompetenties hoger onderwijs.* SLO, nationaal expertisecentrum leerplanontwikkeling.
[5] Raquel Cerdán, Eduardo Vidal-Abarca, Tomas Martinez, Ramiro Gilabert, and Laura Gil. 2009. Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction* 19, 1 (2009), 13–27.
[6] Karl Anders Ericsson and Herbert Alexander Simon. 1993. *Protocol analysis: Verbal reports as data, Rev.* the MIT Press.
[7] Laura Gil, Tomas Martinez, and Eduardo Vidal-Abarca. 2015. Online assessment of strategic reading literacy skills. *Computers & Education* 82 (2015), 50–59.
[8] Susan R Goldman and Richard P Durán. 1988. Answering questions from oceanography texts: Learner, task, and text characteristics. *Discourse Processes* 11, 4 (1988), 373–412.
[9] William Grabe. 2008. C11. Becoming a Strategic reader. In *Reading in a second language. Moving from theory to practice.* 220–242. https://doi.org/dx.doi.org/10.1017/CBO9781139150484
[10] William Peter Grabe and Fredricka L Stoller. 2013. *Teaching and researching: Reading.* Routledge.
[11] Arthur C Graesser, Sallie E Gordon, and Lawrence E Brainerd. 1992. QUEST: A model of question answering. *Computers & Mathematics with Applications* 23, 6-9 (1992), 733–745.
[12] Karyn Higgs, Joseph P Magliano, Eduardo Vidal-Abarca, Tomas Martínez, and Danielle S McNamara. 2017. Bridging skill and task-oriented reading. *Discourse Processes* 54, 1 (2017), 19–39.
[13] Amelia Mañá, Eduardo Vidal-Abarca, and Ladislao Salmerón. 2016. Effect of delay on search decisions in a task-oriented reading environment. *Metacognition and Learning* (2016). https://doi.org/10.1007/s11409-016-9162-x
[14] Matthew T McCrudden, Gregory Schraw, and Gretchen Kambe. 2005. The Effect of Relevance Instructions on Reading Time and Learning. *Journal of Educational Psychology* 97, 1 (2005), 88.
[15] Ilona Milliano, Amos Gelderen, and Peter Sleegers. 2016. Types and sequences of self-regulated reading of low-achieving adolescents in relation to reading task achievement. *Journal of research in reading* 39, 2 (2016), 229–252.
[16] Peter B Mosenthal. 1996. Understanding the strategies of document literacy and their conditions of use. *Journal of Educational Psychology* 88, 2 (1996), 314.
[17] OECD. 2004. *The PISA 2003 Assessment Framework: Mathematics, reading, science and problem solving knowledge and skills.* Number 1. Paris. 194 pages. https://doi.org/10.1787/9789264101739-en
[18] Cecile A Perret, Amy M Johnson, Kathryn S McCarthy, Tricia A Guerrero, Jianmin Dai, and Danielle S McNamara. 2017. StairStepper: An Adaptive Remedial iSTART Module. In *Int'l. Conf. on Artificial Intelligence in Education.* Springer, 557–560.
[19] Peter Reimann. 2009. Time is precious: Variable-and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning* 4, 3 (2009), 239–257.
[20] Jean-Francois Rouet and M Anne Britt. 2011. Relevance processes in multiple document comprehension. *Text relevance and learning from text* (2011), 19–52.
[21] Jean-François Rouet, M Anne Britt, and Amanda M Durik. 2017. RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist* 52, 3 (2017), 200–215.
[22] Jean-François Rouet, Eduardo Vidal-Abarca, Alain Bert Erboul, and Victor Millogo. 2001. Effects of information search tasks on the comprehension of instructional text. *Discourse Processes* 31, 2 (2001), 163–186.
[23] Ladislao Salmerón, Johannes Naumann, Victoria García, and Inmaculada Fajardo. 2017. Scanning and deep processing of information in hypertext: an eye tracking and cued retrospective think-aloud study. *Journal of Computer Assisted Learning* 33, 3 (2017), 222–233.
[24] Bernhard Schmitz. 2006. Advantages of studying processes in educational research. *Learning and Instruction* 16 (2006), 433–449. https://doi.org/10.1016/j.learninstruc.2006.09.004
[25] Roel van Steensel, Ron Oostdam, and Amos van Gelderen. 2013. Assessing reading comprehension in adolescent low achievers: Subskills identification and task specificity. *Language testing* 30, 1 (2013), 3–21.
[26] E Vidal-Abarca, R Gilabert, and JF Rouet. 1998. The role of question type on learning from scientific text. In *Meeting on Comprehension and Production of Scientific Texts, Aveiro, Portugal.*
[27] Eduardo Vidal-Abarca, Amelia Mañá, and Laura Gil. 2010. Individual differences for self-regulating task-oriented reading activities. *Journal of Educational Psychology* 102, 4 (2010), 817.
[28] Eduardo Vidal-Abarca, Tomás Martinez, Ladislao Salmerón, Raquel Cerdán, Ramiro Gilabert, Laura Gil, Amelia Mañá, Ana C Llorens, and Ricardo Ferris. 2011. Recording online processes in task-oriented reading with Read&Answer. *Behavior research methods* 43, 1 (2011), 179–192.
[29] Philip H Winne, Allyson F Hadwin, et al. 1998. Studying as self-regulated learning. *Metacognition in educational theory and practice* 93 (1998), 27–30.