

Read&Answer, A Tool to Capture on-Line Processing of Electronic Texts

Raquel Cerdán*, Eduardo Vidal-Abarca, Ladislao Salmerón, Tomás Martínez and Ramiro Gilabert

Department of Developmental and Educational Psychology, University of Valencia, Avda. Blasco Ibáñez 21, 46010 Valencia, Spain

Abstract: This paper is aimed at presenting *Read&Answer*, a tool that records reading times, one of the main on-line methods employed in text processing research. *Read&Answer* allows the recording, analysis and interpretation of the learner processing in order to test specific hypotheses and explain final comprehension results. First, we will describe the tool, and then we will briefly explain some research studies using the tool. We will show how *Read&Answer* can be used in combination with another on-line method extensively employed in text processing research, i.e., verbal protocols, and we will also compare *Read&Answer* with eye movement tracking, a widely accepted on-line reading times technique.

Keywords: Online processing measures, electronic text reading, think aloud, eye movements.

1. ON-LINE METHODS TO STUDY ELECTRONIC TEXT COMPREHENSION

Learning from current text electronic systems such the Web require readers to engage in several strategic tasks such as searching for information, evaluating the quality of the accessed content, and integrating information from different texts [1, 2]. This complex picture of learning from text differs from traditional views that depict the reader of print text as a passive actor whose main role is to infer the meaning of a unique text. According to the complexity of learning from current text electronic systems, recent research efforts have been devoted to develop research tools to evaluate the strategic behaviors used by readers learning in electronic environments [3]. Most of these methods rely on on-line measurements of the cognitive and behavioral activity concurrent with learning, because they allow tracing comprehension processes as they unfold [3]. On-line methods enable researchers to test theoretical hypotheses from reading models [4, 5], to study individual differences on reading [6] or to test the effects of specific reading conditions [7, 8], among other purposes.

In the current paper we will describe a software tool called *Read&Answer* developed in our research lab to capture on-line processing of electronic texts. First we will briefly explain some research studies conducted in our lab with the tool. Next we will show how *Read&Answer* can be used in combination with another on-line method extensively employed in text processing research, i.e., verbal protocols. Finally, we will compare reader's behavior using a hypertext in *Read&Answer* with that of a second group using the Eye movement method, a widely accepted on-line reading times technique.

1.1. Characteristics of the *Read&Answer* Software

Broadly speaking, two main types of online methods can be distinguished, i. e., reading time methods (e.g. moving

window, eye movements) and verbal protocols. Whereas reading times methods focus on quantitative data (e. g., time spent on a piece of text), verbal protocols aim at providing a qualitative description of the cognitive processes at work during comprehension (e. g., self-explanations when reading a sentence). *Read&Answer* is a software tool that allows the recording of reading time when readers read an electronic text and answer questions from it, and that can also be easily combined with verbal protocols, in order to study reading processes and strategies when participants perform complex comprehension tasks. *Read&Answer* presents readers with a full screen of text in which all information except the unit currently read is masked. A unit can be a sentence, a paragraph, an image or any piece of information as it has been previously defined by the researcher. Readers unmask a unit by clicking on it with the mouse and they can reread them in any order they choose (see Fig. 1). Once a reader clicks on a different section it becomes readable, and the previous one is masked again. In addition, *Read&Answer* allows the inclusion of questions, prompts or a notepad during the reading task, and also records readers' movements from the text to the additional screens and vice versa. Readers can access the 'question screen', which presents the upper part for the question and the lower part for the answer. A simple interface allows the reader to move from one question to another and from the question screen to the text screen, and vice versa (see Fig. 1).

Read&Answer automatically generates a list of all units active at any given moment throughout the reader's processing, as well as the length of time each unit was active. This output also includes the writings from students (e.g. answers to questions, notes taken in the notepad), not only in its final version, but each of the modifications students include, so that the writing process can easily be traced and later analyzed. From this general output, the researcher may request *Read&Answer* to compute automatically a number of indices (e.g., time reading specific text units, number of visits to specific questions, or pre-defined self-regulation measurements). The first output generated by *Read&Answer* is a list of every activity the reader undertakes, that is, selecting or reading a text

*Address correspondence to this author at the Department of Developmental and Educational Psychology, University of Valencia, Avda. Blasco Ibáñez 21, 46010 Valencia, Spain; E-mail: Raquel.Cerdan@uv.es

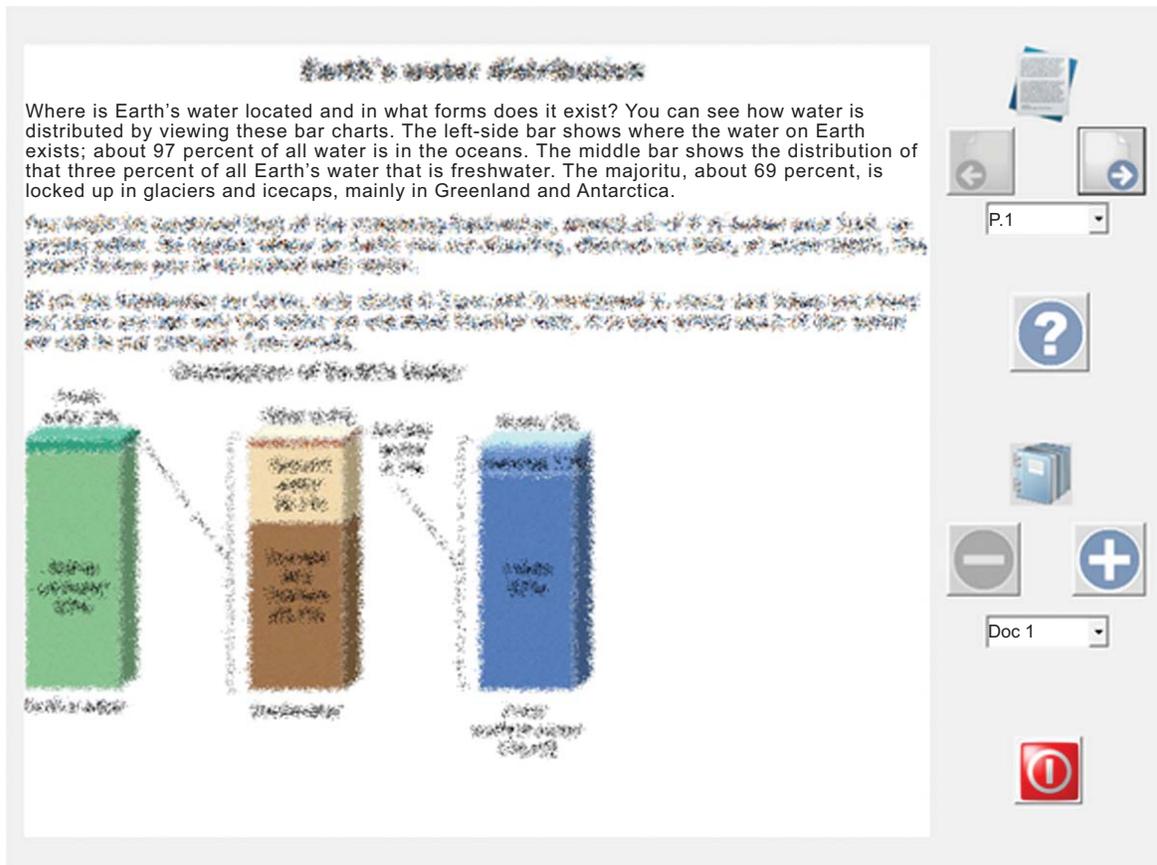


Fig. (1). Text screen from *Read&Answer* software showing the unmasked region.

paragraph, selecting or reading a specific question, writing the answer to every question, and the order of activities students undertake, which is a list sequentially ordered to follow the student's activities in the experimental session. The list also records the time each activity lasted. The second output is a summary of the reader's behavior on the task, that is, when the participant reads the text and answers the question, including the number and percentage of relevant paragraphs for the task a student has visited to answer a question. The third output is the record of the reader's answers to each question. Although these three outputs provide very valuable data to contrast our hypothesis, having only one output from which the researcher may predefine specific measurements to be computed automatically has obvious theoretical and practical advantages. Theoretical advantages because the tool becomes more flexible to adjust to the different study hypothesis. Practical advantages because it considerably lightens the amount of data analysis workload. Currently we are developing a new software that will allow to filter participant's data called *R.A.D.A.R.* that will be able to automatically generate highly specific indexes of reading behavior based on experimenter requests (e.g. percentage of times that a participant reads a relevant section for the task immediately before she correctly answers a question).

Read&Answer has some features that make it different from other reading time techniques and tools. First, traditional methods such the Moving window or the recording of eye movements are mainly focused on the

reader's behaviour during the reading of a text. *Read&Answer*, by contrast, has been designed to record the students' behaviour while they read a text and also while they perform a complex learning task such as answering questions, writing essays or taking notes. Second, *Read&Answer* has been designed to be as close to natural reading conditions, such as that permitted by the Eye movement technique. However, contrary to the Eye movement technique it is far easier to use from the experimenter perspective as there is no need of calibration and complex data filtering, which is the case when using any eye tracker machine. Moreover, the apparatus is simpler, since *Read&Answer* can be easily implemented on a basic PC. This also means that several participants can be tested simultaneously with *Read&Answer* (we have conducted experiments with up to 30 participants simultaneously, in both adolescents and adult population) whereas only one participant at a time can be tested with the eye movement technique. Third, *Read&Answer* can present both texts and pictorial information (e.g., images and texts), which is a drawback of the Moving Window technique that only deals with text. Finally, recording reading time and verbal protocols at the same time is easy using *Read&Answer*, as the student only needs to talk aloud by means of a microphone connected to the same computer running *Read&Answer*. With the Eye Movement technique, an extra device is needed to record speech due to the high computing demands required for the eye tracker. In addition, the recording of the eye gaze can be distorted if participants talk during the recording of the eye movements [9].

1.2. Studies of Electronic Text Literacy Using Read&Answer

Read&Answer has been successfully used to investigate different cognitive aspects of electronic reading literacy in adolescents and undergraduate students. Some of these topics include the study of the metacognitive strategies involved in answering questions from a text [10]; the study of the impact of specific task conditions (e.g., answering either high- vs low-level questions) on search processes and reading comprehension [8]; the study of the integration of information from multiple documents under specific task conditions [7, 11]; or the study of the role of graphical overviews in hypertext comprehension [12]

To briefly illustrate the possibilities of *Read&Answer*, we will describe a recent study by Cerdán and cols [8]. In an electronic reading study participants were asked to read a multiple-page text on Atomic Models to answer questions demanding either a deep processing of the text information (i.e. high level questions) or a shallow processing of the text (i.e. low level questions). Using the on-line measurements provided by *Read&Answer* we identified the question-answering patterns associated to both type of question and success in answering. All participants had the text available to search for the answer and they could move from the text screen to questions screen at will so that on-line behavior measurements could be recorded on-line. Apart from obtaining several post-learning measures, we specifically defined the following on-line measurements to track the question-answering pattern. First, a set of global time measurements which would highlight the general behavior in the question-answering task: (a) *time in experiment session*; (b) *time reading questions*; (c) *time reading text*, and (d) *time answering*. And, to account for the strategic behavior when answering the different types of questions we established the following units based on theoretical assumptions on the stages that a reader undergoes when answering questions from a text [13]. First, understanding the question demands (i.e., task-specifications phase) by computing the mean number of seconds per question when reading each question. Second, to capture the stage of *selection* and the *processing* of relevant units of information we computed the percentage of relevant paragraphs in comparison to the total number of paragraphs visited per question, and the percentage of time reading relevant paragraphs in comparison to the total time spent reading per question, respectively. Third, to capture the *answer-composition* stage we computed the mean number of seconds per question when answering and the number of QTW cycles, that is, reading-the-question (Q), reading the text (T) and writing an answer (W) as these cycles reflect search iterations undertaken to satisfy goals established in the task-specifications phase [13]. Finally, we also counted the number of times students *visited the table of contents*.

These on-line measures could be defined and captured with the software *Read&Answer* in the context of a question-answering task and could be connected to several post-learning measures to better define the successful strategic behavior when answering questions from a text [8]. In fact, main results indicated that students who answered high-level questions ($M = 38.02$, $SD = 14.96$) outperformed students answering low-level questions ($M = 26.23$, $SD = 16.96$) in a deep comprehension post-learning measure $F(1, 33) =$

5.031 , $p < .05$, but not on recall of the text. This superiority of high-level questions over low-level questions in fostering comprehension from text was associated with a strategic pattern that *Read&Answer* on-line data was able to reveal. Most significantly, high-level questions produced a higher number of QTW cycles ($M = 4.60$, $SD = 2.23$) than low level questions ($M = 3.25$, $SD = .66$), $F(1, 33) = 8.72$, $p < .01$, which suggested that students were in the main engaged in relational processing through various iterations in order to connect several pieces of information and produce a coherent answer when answering this type of questions. The higher connectivity of the relevant pieces of information through repeated iterations allowing inference drawing could explain the superiority of high-level questions on deep comprehension.

Although the previous studies clearly evidence that *Read&Answer* provide valuable information of the reading process of electronic texts, the assumption that the tool mimics real reading situations is still an open question. The fact that readers have to use continuously the mouse to be able to advance in their reading may interfere with the natural reading process. In order to test this assumption, we present two studies in which *Read&Answer* was used with other two classical methods for the recording of on-line processes: verbal protocols and eye movements.

2. CONVERGENT EVIDENCE FROM READ&ANSWER AND THINK-ALOUD METHOD

2.1. Convergent Evidence of the Readers' Behavior Using On-Line Time Data and Think-Alouds

There is a wide range of empirical studies that have successfully used the Think-aloud methodology in different fields of interest in Psychology. The think-aloud method has been demonstrated to reflect what is available in working memory, accessible to consciousness and codable in language [14, 15] and would therefore be indicative of which mental processes and contents are responsible for how students perform and learn from an specific task. Consequently, by making students verbalize what they are thinking simultaneously to the task (i.e., *concurrent verbal reports*) or after performing the task (i.e., *retrospective verbal reports*) we can analyze in greater detail on-line processing of text, strategies and problem solving.

To our interest is the growing use of this methodology to study general comprehension processes and strategies that students apply when reading texts, validating that the content of think-aloud protocols is linked to comprehension [16-19] Moreover, not only has this methodology been used to study on-line processing of one text, but it has also been used in classical and recent studies of multiple text comprehension and integration processes from multiple perspectives [20-22] Using a think-aloud procedure would provide very valuable data on processes and strategies that occur during comprehension, and the validity of these data would be increased if convergence with other kind of data were provided, such as on-line time measurements [23] Indeed, successful attempts have been done to provide converging evidence for the conclusions drawn from protocol data [24-26] Additionally, using the think-aloud methodology as only a part of a converging evidence strategy has been presented

in Magliano and Graesser's [27] three-pronged approach to comprehension research.

2.2. Integration of Information Across Documents (Experimental Study)

We designed a study in which *Read&Answer* and think aloud were used to provide convergent evidence on how students integrate information from multiple documents [7]. Undergraduate students read three scientific documents on *bacteria resistance to antibiotics* to describe and explain *how bacteria resist the effects of antibiotics and which biological mechanisms explain this phenomenon and its transmission to other bacteria*, for which students had to select and process a subset of the information in the documents. Half of the students wrote an inter-textual essay on the global question, whereas the other half answered shorter intra-text questions covering the same information that required students to integrate information within a single text. Based on a number of studies that found that making the process of learning more effortful might be helpful in terms of deep learning, which is not true in terms of superficial understanding [28, 29] it was hypothesized that the global inter-textual essay would be more effective at the deep level of understanding (i. e., problem solving) than shorter intra-text questions. However, no differences would be found at the more superficial level (i. e., sentence verification task). Behavioral on-line evidence for integration processes being induced by comprehension tasks were taken inspired on some previous research studies [4, 21].

Before reading the three texts, half of the students were instructed to write the global essay, whereas the other half answered four brief questions covering the same information. Half of the students in each condition thought-aloud when doing the task, while the rest did it silently. Students read the documents and wrote the essay on a computer screen using the application *Read&Answer*. After writing either the essay or answering the questions students were tested on: (a) Sentence Verification Task (SVT), which assessed the student's recall, and (b) problem solving task to assess their level of integration (i. e., deep comprehension).

2.3. On-Line Tracking of the Integration Behavior Using On-Line Time-Based Measurements and Think Alouds

To capture the integration process and the deeper processing of textual information, which should be more present in the essay task according to our hypothesis, we used the following main measures obtained from the three outputs in *Read&Answer*. We registered the students' *processing of both relevant and non-relevant units of information* by measuring the mean time per word spent when they read the two sorts of units of information. We also measured *integration of information* by counting non-consecutive readings of relevant units of information (i.e. *integration-processing behavior*), which indicates an effort to connect and integrate the two paragraphs, and the *single-unit pattern of processing* by counting the number of times a student, after reading a paragraph with relevant information, performed another action (e. g., reading a question, writing an answer, or going to a paragraph with non-relevant information), and then returned to the paragraph with previously read relevant information. This behavior is seen

to indicate the processing of text paragraphs in isolation from other paragraphs in the three texts.

To classify the think-aloud data, and make it reflect the integration and elaboration of information process, we established the following set of categories under which each of the students' statements were allocated: 1) *Task verbalizations* associated with the process of performing the task or the use of the software (e. g., *these are the instructions, aren't they?, I read them aloud. Okay, I will start by reading the questions*); 2) *Search*: whenever the student verbalized searching and locating information in the texts, regardless of whether it was relevant or irrelevant (e. g., *Okay, I think this was located in text 3. I will reread text 3 to find the answer*); 3) *Relevant-unit understanding*: verbalizations about relevant units of information that indicate deep comprehension, including inferences, summaries, elaborations, and explanations (4) *Non-relevant-unit understanding*: deep comprehension verbalizations, but referring to irrelevant units of information; (5) *Superficial understanding*: verbalizations indicating paraphrases and irrelevant associations, no matter if they referred to relevant or irrelevant units of information; (6) *Writing self-regulation*: they reflect self-regulation of the answering process (e. g., *I go to question 3, because I don't know what to answer in this one*), *I think this one is correct this way*"); (7) *Superficial writing*: they mainly manage orthography in the writing process (e.g., *I am not sure if I should add a period at the end of this sentence*"). Verbal protocols were coded by two experimenters according to the above categories. The two scorers coded approximately one third of the protocols separately, but checking the inter-rater agreement until reaching 95% agreement, then only one scorer coded the rest of the protocols.

2.4. Results

2.4.1. Product and Read&Answer Measurements

To analyze the learning outcomes, we performed one-way Anovas with independent variable Task (inter-text essay vs intratext questions). Results showed that students performing the essay task scored higher on the problem solving measure ($M = 48.22$, $SD = 16.87$) than those answering brief questions ($M = 37.33$, $SD = 12.16$), $F(1, 46) = 6.21$, $p < .05$, whereas no significant differences were found for SVT. When analyzing the *Read&Answer* text processing measures (i.e. *time per word in relevant and non-relevant information*), we performed one repeated-measures ANOVA with between-subjects variable Task and the repeated-measures variable Type of information (relevant vs non-relevant). A significant interaction between Task and Type of information indicated that inter-text task participants read textual information more slowly, that is, at a higher time per word rate ($M = 0.19$, $SD = 0.06$) than students answering intra-text questions ($M = 0.15$, $SD = 0.05$), $F(1, 46) = 4.99$, $p < .05$, but there were no differences in the non-relevant units of information (see Fig. 2). This finding means that the inter-text task induced a more detailed processing of documents than intra-text questions. Finally, to account for the integration process we had used two measures obtained from the analysis of the *Read&Answer* behavioral sequences: integration-processing behavior and single-unit processing. Essay students showed more integration-processing behavior

($M = 5.76$, $SD = 4.47$) than intra-text question students did ($M = 1.40$, $SD = 1.00$), $F(1, 46) = 21.48$, $p < .05$. This contributes to explaining why the former scored higher than the latter on the problem solving task. Contrarily, it was intra-text question students who scored higher in the single-unit processing measure ($M = 26.60$, $SD = 9.83$) than inter-text students ($M = 10.92$, $SD = 8.44$), $F(1, 46) = 34.45$, $p < .05$, as expected.

2.4.2. Think-Aloud Measurements

To examine the convergence between reading time and verbal protocols data, we also compared students' verbalizations from the essay inter-text task and the intra-text question groups. We performed one-way ANOVAs for each of the categories explained above. Results indicated that students performing the inter-text task ($M = 15.18$, $SD = 12.57$) significantly verbalized more on understanding relevant information than those who answered intra-text questions ($M = 6.25$, $SD = 6.67$), $F(1, 21) = 4.64$, $p < .05$. Second, students answering the inter-text task produced fewer superficial writing comments ($M = 2.72$, $SD = 2.96$) than students answering intra-text questions ($M = 8.75$, $SD = 6.21$), $F(1, 21) = 8.53$, $p < .05$. This two data, in addition to the relevant reading rate and the integration behavior, provided convergent evidence for our claim that the inter-text task promoted a deeper processing of the multiple sources, in comparison to intra-text questions, which was indeed the result we found in the problem solving measure.

2.5. Conclusions

Thus, this experiment showed the capability of *Read&Answer* to provide on-line evidence in terms of reading times for the integration processes, and its convergence with verbal protocol data. It also indicated the possibilities of the *Read&Answer* software to be easily combined with classical on-line methods. An important methodological issue to further analyze is the extent to which the combination of the two methods may be a source of cognitive overload and thus hinder student's performance, which is a result we partially found in this experiment. For further details, see Cerdán & Vidal-Abarca [7].

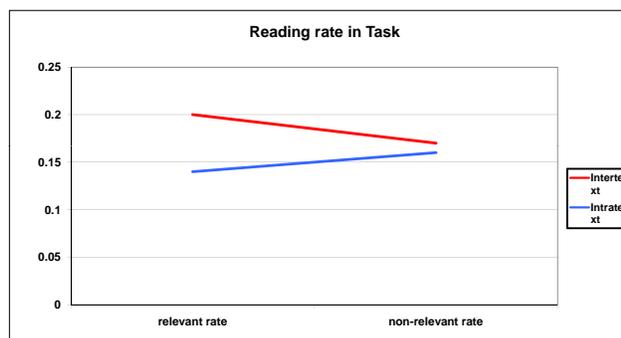


Fig. (2). Interaction effect between Task and Type of information. Note: Mean relevant and non-relevant time per word rate in inter-text task and intra-text questions.

3. COMPARISON OF READING BEHAVIOUR IN READ&ANSWER AND WITH THE EYE MOVEMENT TECHNIQUE

As mentioned above, readers using *Read&Answer* are required to use continuously the mouse to be able to advance

in their reading. To what extent this particular reading situation may interfere with the natural reading process? In order to test this assumption, we compared the reading behaviour in *Read&Answer* with that obtained with the Eye Movement technique, an on-line method that clearly mimics natural reading situations.

We compared reading times patterns from two experiments in hypertext comprehension, one of them conducted with the Eye Movements technique and the other run with *Read&Answer*. The two experiments were fairly similar: a set of undergraduate students read the same expository hypertext with the aim to prepare for a comprehension test after the reading. The text was provided with a graphical overview depicting the structure of the hypertext. In both experiments we tested the hypothesis that an overview might be useful only before reading most of the sections of a hypertext as it acts as a schema for the text organization [30].

3.1. Method

In the two experiments, 93 undergraduate students from the University of Valencia (40 in the EM study, 53 in the *R&A* study) volunteered for class credit. Participants read two expository hypertexts on the topic of Climate Change and the Mediterranean, although only the data from one common text to the two experiments is reported here. The text was divided into ten sections. An overview of the text content was available when reading each section (see Fig. 1). Thus, the text changed from one screen to the other, whereas the overview was always present.

In the EM study, eye movements were recorded by a EyeLink II headmounted eye-tracking system. The system registered data binocularly at a rate of 500 Hz. Participants seated approximately 60 cm from the presentation screen. Calibration of eye tracker was performed prior to the reading phase, and a drift-correction was done each time the participant moved to a new text section (i.e. page). For the *R&A* study, students were presented with both, the overview and the text masked. To read a piece of information (e. g., information from a particular box of the overview) students had to click on it, and then it unmasked.

First participants filled out a Prior Knowledge questionnaire, a multiple-choice test consisting of 17 items with five answer alternatives. The content of the items referred to concepts and information central to the issue of climate change. Students were instructed to read the text in order to learn as much as possible from the text because they had to answer some open questions after reading the text. They were also told that the overview could be helpful for comprehending the text. After having read the entire text, participants responded to two open questions that targeted four main ideas from the text (e.g. 'describe the process of how the reduction of the amount of organic carbon in the ground will impact the Mediterranean flora').

3.2. Results

We present reading times' patterns from the reading of the second text because it was identical in both experiments. We obtained the average mean of reading times during first reading and second reading, though we made a logarithmical transformation to normalize data. With the EM technique

(first experiment) first reading was calculated computing mean time of all fixations occurring in a text region (i.e. text sentence) that had not been seen before, which is usually referred in the literature as first pass fixation duration. Second reading was calculated computing mean time of all fixations occurring in a text region (i.e. text sentence or a box in the overview), provided that the saccade came from a different region and that the landing region had been fixated before, which is usually referred in the literature as second pass fixation duration. When using *Read&Answer* (second experiment), first reading was calculated averaging mean time unmasking a text region (i.e. sentence or box in the overview) for the first time, whereas second reading was obtained averaging times unmasking a text region that had been accessed before.

3.2.1. Reading Time Analyses

Reading time data were weighted on the basis of the number of characters in each text region. In addition, individual distributions were analyzed to detect outliers (fixation times 2 SD above or below the participant's mean). For the regions of the text, those values (between 1.4% and 2.6% of data) were replaced by the participant's mean fixation-reading time. For the overview regions, most of outliers corresponded to regions that had not received any fixation-were not unmasked (recall that reading the overview was not compulsory, and participants could read it just at the beginning and ignoring it at the following pages). Thus, for calculating the reading time of overview regions outliers were ignored, and we computed the sum of weighted fixation times on any of the ten sections corresponding to the graphical overview at each page.

Data were analysed with a repeated measures ANCOVA for first and second reading of the text and the overview as

dependent measures, page (1st to 10th) and condition (i.e., EM vs *Read&Answer*) as the independent variables, and participant's prior knowledge on the topic as covariate.

Regarding the text reading, no difference between condition was apparent either for the first reading, $F(1, 88) = 1.09, p = 0.29$, nor for the interaction between condition and page, $F(1, 88) = 1.59, p = 0.21$. Analyses for the second reading times revealed significant differences by condition, $F(1,88) = 69.64, p < .01$, as students using *Read&Answer* re-read the text for less time than those from the EM condition (see Fig. 3). No differences were found for the interaction between condition and page, $F(1, 88) = 2.01, p = 0.16$.

Regarding reading the graphical overview, differences between EM and *Read&Answer* were significant for first reading, $F(1,88) = 6.25, p < .05$, but not the interaction between condition and page, $F(1,88) = 0.98, p = .32$. Participants using *Read&Answer* checked the overview less often than those from the EM condition (see Fig. 4). Although the interaction with page was not significant, data clearly reveal that at the first page differences in reading time disappear. Finally, analyses for the second reading times showed significant differences by condition, $F(1, 88) = 27.64, p < .01$, and also a significant interaction between page and condition, $F(1, 88) = 3.99, p < .05$. Interestingly, planned comparisons of LS squares showed no differences between conditions in reading times for the first page, $F(1, 88) = 0.22, p = .64$.

3.2.2. Comprehension Analyses

Responses for each question were given credit of 0 (incorrect response), 1 (incomplete but correct response) or 2 (complete correct response). Two independent raters coded 20% of the data for each experiment for a reliability sample. Concordance on the sample data was 74% and 78% (EM and

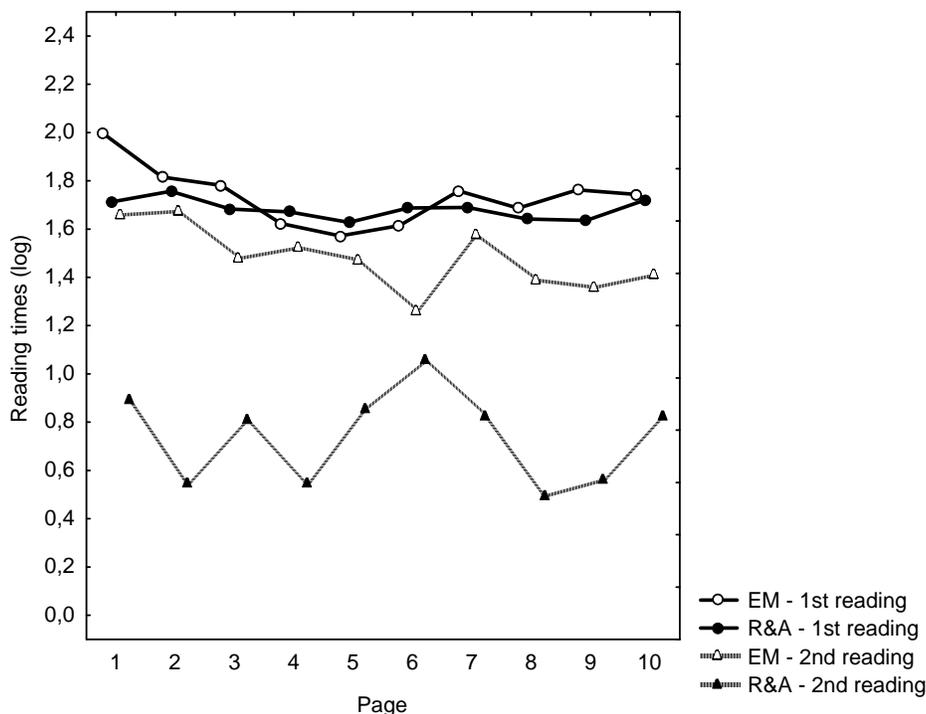


Fig. (3). Average reading times for the text by condition (eye movements and *Read&Answer*) and type of reading (first and second reading).

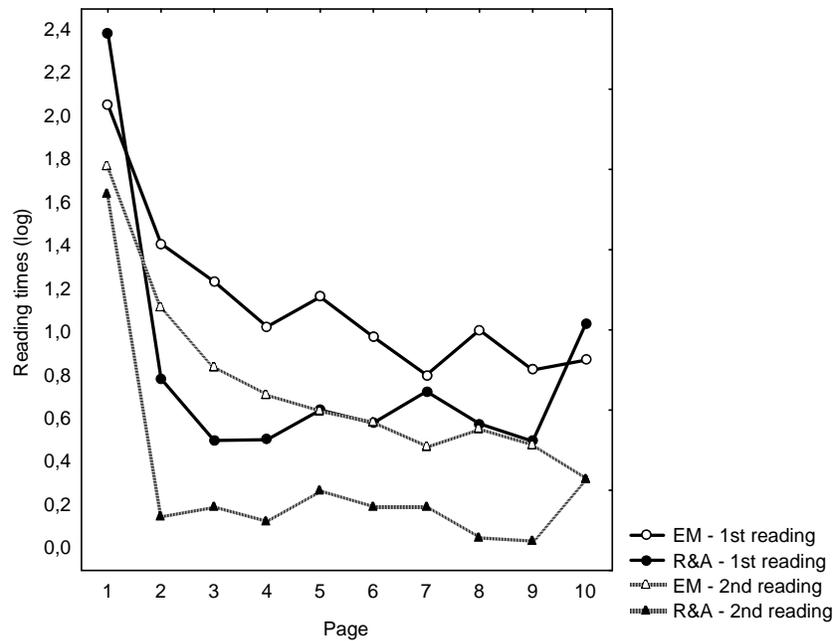


Fig. (4). Average reading times for the graphical overview by condition (eye movements and *Read&Answer*) and type of reading (first and second reading).

R&A study respectively). Disagreements were all resolved in discussion. One rater coded the remainder of the data. Data was averaged for each participant.

We run an ANCOVA with Prior Knowledge as covariate, condition (EM and R&A) as independent variable, and percentage of correct responses as dependent variable. Not surprisingly, Prior Knowledge was positively related to scores on the comprehension test, $F(1, 87) = 3.88, p = .05$ [31]. In addition, data revealed significant differences for condition, $F(1, 87) = 4.87, p < .05$. Participants on the EM condition scores less ($M = 0.32$ average comprehension score, $SD = 0.29$) than those from the R&A condition ($M = 0.47, SD = 0.31$).

3.3. Conclusions

Data from the two studies reveal that the pattern of first reading the text for the first time is very equivalent when using *Read&Answer* and EM technique, which is not true for second reading. When examining the overview, the pattern is different in both conditions. Participants read and reread the graphical overview for the same amount of time at the beginning of the reading (i.e. first page). This reading may correspond to the process of acquiring the representation of the general organization of the text, which has a strong impact in comprehension [5]. In the following pages (i.e. 2nd to 10th), participants using *Read&Answer* read and reread the overview for less time. A possible explanation for this behavior is that students had to click on every graphical overview box to unmask a piece of text. Maybe if the unit of reading would have been the whole overview instead of every single box the pattern of reading the overview with *Read&Answer* had been equivalent to that of eye movement. This possibility and some others (e. g., unmasking a unit

simply by passing the mouse over it, instead of having to click on it) will be checked in future studies.

Interestingly, comprehension scores revealed a surprising pattern: participants using *Read&Answer* score higher on the comprehension test than those using the EM technique. Although the results should be taken cautiously because we used a short comprehension test, in conjunction with the reading times data they seem to show that reading in *Read&Answer* may differ from natural reading situations. The pattern of results suggest that using the mouse to unmask the text may be more demanding for the readers, and that this situation may indeed be beneficial for the comprehension process. Similarly, we have found elsewhere [5] that hypertext readers who mainly read a graphical overview at the beginning of the reading and avoid consulting it afterwards –the same as our participants from the R&A study did– comprehend the text better. At the beginning of the text the overview may act as a schema for the text organization, allowing the reader to incorporate new information from the text into an existing representation. However, if readers continue reading during the rest of the hypertext the overview may inhibit the active processing required to integrate the information from different hypertext sections [5, 32]

4. GENERAL CONCLUSIONS

Read&Answer is a tool to record reading time measures when students have to read any sort of written document and do some of complex activity with the information in the documents (e. g., answering questions, writing an essay and so on). *Read&Answer* has some advantages over other reading times techniques, such as eye movements or moving window apart from the possibility of combining text reading and doing complex tasks. For example, *Read&Answer* is

easy to implement, a great number of participants can be tested simultaneously, and it can easily be used in combination with the think aloud method. *Read&Answer* produces similar patterns of processing when it is compared to eye movement technique. However, some patterns are different. More research is needed to check to what extent *Read&Answer* can be used instead of eye movements for some specific purposes.

ACKNOWLEDGEMENTS

This research was funded by the Spanish Minister of Education and Science (research Project SEJ2005-04500/EDUC).

REFERENCES

- [1] Rouet JF. The skills of document use. From text comprehension to web-based learning. Mahwah NJ: Lawrence Erlbaum Associates 2006.
- [2] Lazonder AW, Rouet JF. Information problem solving instruction: Some cognitive and metacognitive issues. *Comput Hum Behav* 2008; 24: 753-65.
- [3] Rouet J-F, Passerault J-M. Analyzing learner-hypermedia interaction: and overview of online methods. *Instrum Sci* 1999; 27: 201-19.
- [4] Rouet JF, Vidal-Abarca E, Bert-Erboul A, y Millogo V. Effects of information search tasks on the comprehension of instructional text. *Discourse Process* 2001; 31(2): 163-86.
- [5] Salmerón L, Baccino T, Cañas JJ, Madrid I, Fajardo I. Do graphical overviews facilitate or hinder comprehension in hypertext? *Comput Educ* 2009 ; [in press].
- [6] Hyönä J, Lorch RF, Kaakinen JK. Individual differences in reading to summarize expository text: evidence from eye fixation patterns. *J Educ Psychol* 2002; 94(1): 44-55.
- [7] Cerdán R, Vidal-Abarca E. The effects of tasks on integrating information from multiple documents. *J Educ Psychol* 2008; 100(1): 209-22.
- [8] Cerdán R, Vidal-Abarca E, Martínez T, Gilabert R, Gil L. Impact of question-answering tasks on search processes and reading comprehension. *Learn Instrum* 2009; 19 (1): 13-27.
- [9] Kaakinen, J, Hyöna J. Perspective effects on expository text comprehension: evidence from think-aloud protocols, eyetracking, and recall. *Discourse Process* 2005; 40(3): 239-57
- [10] Mañá A, Vidal-Abarca E, Gil L, Dominguez C, Martínez T, Cerdán R. Papel de los procesos metacognitivos en una tarea de pregunta-respuesta con textos escritos (The role of metacognitive processes in a question-answering task on written texts). *Infanc Aprendiz* 2009; [in press].
- [11] Gil L, Vidal-Abarca E, Martínez T. Eficacia de tomar notas para integrar información de varios textos (Efficacy of note-taking to integrate information from multiple documents). *Infanc Aprendiz* 2008; 31(2): 259-72.
- [12] Salmerón L. What eye movements say about how readers process graphical overviews in hypertext. *Workshop on Cognition and the Web: Information Processing, Comprehension and Learning*, Granada, Spain.
- [13] Rouet JF. Question answering and document search. In: Rouet J-F, Ed, The skills of document use. From text comprehension to web-based learning Mahwah NJ: Erlbaum 2006: pp. 93-121
- [14] Ericsson KA, Simon HA. Verbal reports as data. *Psychol Rev* 1980; 87(3): 215-51.
- [15] Ericsson KA, Simon HA. Protocol analysis. Verbal reports as data. Cambridge, MA: The MIT Press. 1993.
- [16] Coté N, Goldman SR. Building representations of informational text: Evidence from children's think-aloud protocols. In: Van Oostendorp H, Goldman SR, Eds. The construction of mental representations during reading. Mahwah, NJ: Lawrence Erlbaum Associates, Inc 1999; pp. 169-93
- [17] Magliano JP, Trabasso T, Graesser AC. Strategic processing during comprehension. *J Educ Psychol* 1999; 91: 615-29.
- [18] Pressley M, Afflerbach P. Verbal protocols of reading: the nature of constructively responsive reading. Hillsdale NJ: Lawrence Erlbaum Associates, Inc 1995.
- [19] Zwaan RA, Brown CM. The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Process* 1996; 21: 289-328.
- [20] Wineburg SS. Historical problem solving: a study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *J Educ Psychol* 1991; 83: 73-87.
- [21] Mannes S, Hoyes SM. Reinstating knowledge during reading: a strategic process. *Discourse Process* 1996; 21:105-30.
- [22] Strømsø HI, Bråten I, Samuelstuen MS. Student's strategic use of multiple sources during expository text reading: a longitudinal think-aloud study. *Cognit Instrum* 2003; 21 (2): 113-47.
- [23] Long DL, Bourg T. Thinking aloud: Telling a story about a story. *Discourse Process* 1996; 21: 329-39.
- [24] Trabasso T, Suh S. Understanding Text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Process* 1993; 16: 3-34.
- [25] Magliano JP, Millis KK. Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Discourse Process* 2003; 21(3): 251-83.
- [26] Suh S, Trabasso T. Inferences during reading: converging evidence from discourse analysis, talk-aloud protocols and recognition priming. *J Mem Lang* 1993; 32: 279-300.
- [27] Magliano JP, Graesser AC. A three-pronged method for studying inference generation in literary text. *Poetics* 1991; 20: 193-232.
- [28] Schmidt RA, Bjork RA. New conceptualizations of practice: common principles in three paradigms suggest new concepts for training. *Psychol Sci* 1992; 3: 207-17
- [29] Wiley J, Voss JF. Constructing arguments from multiple sources: tasks that promote understanding and not just memory for text. *J Educ Psychol* 1999; 91: 301-11
- [30] Kester L, Kirschner P, van Merriënboer J. Timing of information presentation in learning statistics. *Instr Sci* 2004; 32: 233-52
- [31] Chen SY, Fan J, Macredie RD. Navigation in hypermedia learning systems: Experts vs novices. *Comput Hum Behav* 2006; 22: 251-66.
- [32] Nilsson RM. The effects of graphic organizers that signal the structure of a hypertext document on user's navigation strategies and performance. *Proceedings of the conference on Human Factors in Computing Systems archive*. USA: Seattle, Washington 2001; pp. 421-2.

Received: November 22, 2008

Revised: February 2, 2009

Accepted: February 4, 2009

© Cerdán et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.