White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/149770/

Version: Published Version

Article:
Cerdan, Raquel, Candel, Carmen and Leppink, Jimmie (2018) Cognitive Load and Learning in the Study of Multiple Documents. Frontiers in Psychology. ISSN: 1664-1078

https://doi.org/10.3389/feduc.2018.00059

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

frontiers in Education

EMPIRICAL STUDY
published: 24 July 2018
doi: 10.3389/feduc.2018.00059

Cognitive Load and Learning in the Study of Multiple Documents

Raquel Cerdan[1]*, *Carmen Candel*[1] and *Jimmie Leppink*[2]

[1] ERI Lectura, Department of Developmental and Educational Psychology, Universitat de València, València, Spain,
[2] Onderwijs Instituut FHML, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands

This study had two main purposes. First, to test how the availability of documents in multiple document reading might affect students' levels of cognitive load. Secondly, to develop an instrument that captures the different sources of load when working with multiple documents. A total of 125 secondary school students read four short texts on transgenic foods and subsequently responded to an open-ended question that required them to write an essay expressing their personal stance toward the topic. Participants in the experimental treatment condition ($n = 54$) were allowed to go back to the texts any time during the essay task, whereas their peers in the control condition ($n = 71$) were not allowed to do so. As hypothesized through the lens of cognitive load theory, the cognitive load arising from cognitive processes that in themselves do not contribute to learning (i.e., extraneous cognitive load) was somewhat lower in the experimental treatment condition, probably due to split attention effects in the control condition. However, no statistically significant differences were found in perceived task complexity or learning task performance. A reliable instrument to measure different sources of intrinsic and extraneous load in multiple document reading is provided. Implications of these findings for future research are discussed.

Keywords: multiple document reading, cognitive load theory, split attention, extraneous cognitive load, learning outcomes

OPEN ACCESS

Edited by:
Anastasiya A. Lipnevich,
The City University of New York, United States

Reviewed by:
Eric C. K. Cheng,
The Education University of Hong Kong, Hong Kong
Rosemary Hipkins,
New Zealand Council for Educational Research, New Zealand

***Correspondence:**
Raquel Cerdan
raquel.cerdan@uv.es

Specialty section:
This article was submitted to Assessment, Testing and Applied Measurement, a section of the journal Frontiers in Education

Received: *11 April 2018*
Accepted: *02 July 2018*
Published: *24 July 2018*

Citation:
Cerdan R, Candel C and Leppink J (2018) Cognitive Load and Learning in the Study of Multiple Documents. Front. Educ. 3:59. doi: 10.3389/feduc.2018.00059

INTRODUCTION

Reading and multiple document handling skills are a very important asset in our information society where we are drowning in information coming from all kinds of channels. Imagine a group of high school students surfing on the internet with the goal of writing a short synthesis about a recent trip, for instance as a part of school homework. The students may encounter different types of documents varying in, amongst others, topical relevance and source characteristics. To be efficient and appropriately select and read the set of documents that might help students accomplish their goals, strategic decisions have to be made such as selecting some documents and discarding others. They will also have to engage in reading processes to understand the single documents and to make interconnections among the selected documents. Moreover, some school tasks or assessment situations also require readers to first select and read documents and then perform a specific task with or without access to the texts. In short, solving tasks based on multiple documents presents considerable challenges to an individual's text processing skills, and appropriate training of these skills can be of use in formal education as well as in everyday life and lifelong learning.

Processing Demands in Multiple Document Reading

Ample research has been conducted in the field of learning with multiple documents (e.g., Braasch et al., 2016; Rouet et al., 2016; Scharrer and Salmerón, 2016). Reading multiple documents implies that students need to process single texts (Kintsch, 1998) as well as connect information across texts (Rouet and Britt, 2011). Given that the different documents stem from different authors or sources, linking content to sources is also an essential process students need to undertake when processing multiple documents. Moreover, students are frequently asked to solve tasks of a different type based on documents that might or might not be available for task completion, with varying effects on learning from text and task performance (i.e., Cerdán and Vidal-Abarca, 2008). Although the impact of text availability has been a topic of study in the context of single texts (Vidal-Abarca et al., 2010; Schroeder, 2011), it has to our knowledge not been studied in the context of dealing with multiple documents.

Dealing with more than one document with the purpose of comprehension or to solve comprehension tasks poses important challenges on the learner, beyond those at the single text level. In this context, Rouet and Britt (2011) developed a theoretical model that accounts for the processing demands involved in the comprehension of multiple documents (i.e., the multiple-document task-based relevance assessment and content extraction or MD-TRACE). This model identifies dimensions of both external resources (i.e., documents) and internal resources (i.e., knowledge, skills, and beliefs) which may affect how information is processed. Remarkably, this model identifies a set of iterative steps in the processing of multiple documents. The first step refers to the construction of a task model which would represent readers' goals (i.e., locating a piece of information, understanding a conflict among texts). The second step implies a reader's decision of achieving the reading goal by retrieving information from memory, or searching in external resources. The third step of this model involves reading and learning from the textual material. Related to the latter, the documents model framework (Perfetti et al., 1999; Rouet, 2006; Britt and Rouet, 2012) and recent frameworks of purposeful reading (Rouet et al., 2017) specify the types of representations needed when reading multiple documents. In accordance with these frameworks, readers can represent information about the document as an entity, which is referred to as a document node. These may include information about authorship and the document. This source information can be connected to information from other sources through intertext links (i.e., intertext model, or the identification of sources and source-to-content links) and the integrated mental model, which refers to the model readers would build as a result of the integration of the ideas and sources included in the different documents. As a fourth and final step, MD-TRACE proposes that readers come up with a task product to meet the goals of the task model. It may involve either a single piece of information or a complete argument. As they write, students need to monitor whether the product satisfies the goals in the task model. If not, they should update their task model, read additional information (e.g., additional documents) and update their product until satisfactory.

In sum, MD-TRACE and recent developments of a purposeful reading framework (i.e., RESOLV, Rouet et al., 2017) provide a thinking framework of how students deal with and learn from multiple texts. Readers, text, and contextual factors influence how reading and learning from multiple documents takes place (i.e., Strømsø et al., 2011; Macedo-Rouet et al., 2013; McCrudden et al., 2016). In all, when working with multiple documents, readers need to face the cognitive demands of processing both single and multiple texts, plus dealing with the task/s at hand. A systematic identification of sources of processing effort when dealing with multiple documents in task-oriented reading situations may allow to better identify the varying demands involved in learning from multiple documents. For this purpose, cognitive load theory (for recent developments, see: Leppink et al., 2015) may provide a useful framework.

Cognitive Load Theory as a Framework for the Study of Multiple Document Reading

In cognitive load theory, learning is the gradual development of cognitive schemas (Leppink et al., 2015). New content elements—which have yet to be integrated into these cognitive schemas—constitute an *intrinsic cognitive load*, whereas cognitive processes that do not contribute to this integration of content elements into cognitive schemas are commonly referred to as *extraneous cognitive load* (Leppink and Van den Heuvel, 2015). The sum of the intrinsic and extraneous cognitive load constitutes the total cognitive load and needs to operate within the narrow limits of working memory (Leppink et al., 2015). As the extraneous cognitive load does not contribute to learning (i.e., schema development), we should design learning materials and instruction around them such that this ineffective cognitive load is minimized (Leppink and Van den Heuvel, 2015). Under that condition, we may stimulate learners to allocate their remaining working memory resources to dealing with the intrinsic cognitive load (Leppink, 2014; Leppink and Duvivier, 2016). In fact, when extraneous cognitive load is kept to a minimum, a somewhat higher intrinsic cognitive load can result in more learning (Lafleur et al., 2015).

Research inspired by cognitive load theory has resulted in a wide variety of instructional design principles and guidelines (Van Merriënboer and Sweller, 2010; Sweller et al., 2011). These principles and guidelines have resulted in concrete models for the design of curricula, coursework, and single documents and learning tasks (Van Merriënboer and Kirschner, 2012; Leppink and Van den Heuvel, 2015; Leppink and Duvivier, 2016). Moreover, measures of intrinsic and extraneous cognitive load have been developed (Leppink et al., 2013, 2014). When used along with learning and/or performance measures, cognitive load measures can help to gain insight into factors that may facilitate or hinder learning (Leppink, 2016).

One instructional design principle that may well be present in a multiple document reading context is that of *split attention* (Van Merriënboer and Sweller, 2010), that is: a reader has to divide attention between instruction and the content itself. For example, in the context of online learning, we should minimize *spatial split attention* by avoiding whenever possible situations

when learners have to scroll back and forth between instruction on one (part of a) page and the content (e.g., documents to be read) on another (part of a) page. Analogously, withholding instruction about a learning task or relevant task content itself (e.g., the documents that need to be incorporated in an essay) while learners are supposed to perform that task may create *temporal split attention*. Both spatial and temporal split attention contribute to extraneous cognitive load (Leppink and Van den Heuvel, 2015) and should therefore be minimized in order for learners to be able to optimally allocate their mental resources for dealing with the intrinsic cognitive load that is a function of the complexity of the content to be learned.

Using cognitive load measurements in experiments that manipulate extraneous cognitive load (e.g., more or less split attention) can help us understand to what extent split attention is indeed an issue in multiple document reading.

The Current Study

The purposes of this study are two-fold. First, in the light of the aforementioned split attention effect in the context of multiple document reading, the current study examines to what extent allowing students to go back to the texts while writing an essay incorporating information from these texts can help reduce extraneous cognitive load among secondary school learners. As intrinsic and extraneous cognitive load can vary independently and split attention should affect only extraneous cognitive load (i.e., not the complexity of the texts itself), we hypothesized that allowing students to back to the texts would reduce extraneous cognitive load but we had no reason to believe that the intrinsic cognitive load would be reduced as well. Finally, given that extraneous cognitive load in itself does not benefit learning, it would be odd to observe a superior essay performance in the condition where students are not allowed to go back to the texts. Second, in order to appropriately capture cognitive load effects when processing multiple documents, an instrument has been developed to measure the different sources of cognitive load when dealing with this learning scenario. This paper seeks thus also to provide an appropriate and valid instrument to measure cognitive load when dealing with multiple documents of study.

To cover both goals, an experimental study was designed that involved a group of secondary school students reading a set of conflicting documents on a biology topic and subsequently performing and essay task based on the texts. One condition would have access to the documents while performing the essay, while the rest would not. Students completed a cognitive load measurement instrument to capture sources of processing demands in these learning scenarios. This study is presented next.

METHODS

Participants and Design

A total of 125 secondary school students read four short texts on transgenic foods—two in favor and two against—and subsequently responded to an open-ended question that required them to write an essay expressing their personal stance toward the topic. Participants in the experimental treatment condition ($n = 54$) were allowed to go back to the texts any time during the essay task, whereas their peers in the control condition ($n = 71$) were not allowed to do so.

Students from four equivalent classes from the same school were included in the study. Within each class, random assignment to each of the experimental conditions was performed. Teachers reported a lack of students with reading difficulties within the sample. The study was conducted by the same research assistants in different days. This way, we would control that instructions and procedure were exactly the same for all participants.

The study was approved by the school principal. Both the school and parents provided informed written consent to participate in the study. This study was carried out in accordance with the recommendations of the Research Ethics Committee of the University from the first author. No approval was required as per institutional as well as national guidelines and regulations. In order to guarantee the privacy of participants, no personal data was provided. Code identifiers were used in order to gather the different learning materials from the students.

Materials

Text Materials

Materials were adapted from those used in a previous study (Cerdán et al., 2013). Four texts of ∼300 words each dealt with the question whether transgenic foods should be cultivated and distributed, and differed in trustworthiness (i.e., two more and two less trustworthy, as rated by expert biology teachers) as well as in the level of agreement toward the topic (i.e., two in favor and two against). A difference with the previous study (Cerdán et al., 2013) is that in the current study, materials were presented in paper-and-pencil model instead of electronically.

Cognitive Load

Using the principles from previous work in cognitive load measurement (Leppink et al., 2013, 2014), MD-TRACE (Rouet and Britt, 2011), and recent developments of purposeful reading frameworks (Rouet et al., 2017), we developed an instrument to capture sources of intrinsic and extraneous cognitive load experienced by students while performing the task. Items were formulated such that they cover the main processes involved when solving a task based on the reading of multiple documents—that is, reading and comprehending the single texts and creating the interconnected model from the multiple texts, as well as dealing with instructions and task demands—and as such reflect different sources of cognitive load. The resulting Cognitive Load questionnaire for Multiple Document Reading (CL-MDR) is presented in **Table 1**.

Items 1–8 represent three subtypes of intrinsic cognitive load: *single text processing* (items 6 and 7); *multiple texts processing* (items 1–4); and developing *an intertext model* (Rouet and Britt, 2011) of what are the main ideas expressed by the texts and how authors of the different texts relate to these main ideas (items 5 and 8). Whereas, items referring to *multiple documents processing* would emphasize students' effort to integrate information from the texts, items in the category of *intertext model* would reflect the process of constructing mental representations of sources and source-to-content links. Items 9–12 represent extraneous

cognitive load related to task demands. As explained in the introduction, when readers encounter situations of multiple document reading, they need to activate processes to understand the texts individually (*single text processing items*), how specific features such as source characteristics connect to each other (*intertext model items*) and, finally, how the different ideas present in the documents can be integrated (*multiple texts processing items*). In addition, processing demands derived from the task/s at hand might increase students' load (*task completion items*).

The difference between the CL-MDR and cognitive load questionnaires used in other studies (e.g., Leppink et al., 2013, 2014; Lafleur et al., 2015) is that the CL-MDR is the first questionnaire that conceptualizes intrinsic cognitive load as consisting of different subtypes. Reason for that difference is that research inspired by cognitive load theory has thus far largely focused on learning from single documents and hence two of the three subtypes of intrinsic cognitive load in the CL-MDR simply do not apply there.

Procedure

Participants individually performed the task in a single session. They were presented with a booklet containing the aforementioned condition-specific instructions and the four texts. Students were allowed to read the whole set of texts (i.e., around 1,200 words) for a period of 15 min. Subsequently, they were asked to write an essay expressing their personal stance toward the topic of transgenic foods and were given 20 min to do so. Students in the experimental treatment condition had the documents available while writing the essay, whereas students in the control condition did not have access to these documents. Immediately after completion of the essay task, students completed the cognitive load questionnaire (i.e., another ∼10 min).

TABLE 1 | Cognitive Load questionnaire for Multiple Documents Reading (CL-MDR), organized by cognitive process and type of cognitive load: items 1–8 represent three subtypes of intrinsic cognitive load, whereas items 9–12 represent extraneous cognitive load.

Item wording	Process	Load
[1] I have invested mental effort in locating important information in the texts	Multiple texts	Intrinsic-2
[2] I have invested mental effort in detecting contradictions among the texts	Multiple texts	Intrinsic-2
[3] I have invested mental effort in integrating information among texts	Multiple texts	Intrinsic-2
[4] I have invested mental effort in integrating complementary information among texts	Multiple texts	Intrinsic-2
[5] I have made a big effort to construct a general idea of what the texts say	Intertext model	Intrinsic-3
[6] I have had difficulties in understanding some ideas in the texts	Single text	Intrinsic-1
[7] I have had difficulties in understanding some words in the texts	Single text	Intrinsic-1
[8] I have dedicated effort to relating the authors to the ideas they expressed in the different texts	Intertext model	Intrinsic-3
[9] I have had difficulties in understanding the instructions for the task	Task completion	Extraneous
[10] I have had difficulties in understanding the wording of the task I had to solve	Task completion	Extraneous
[11] I have had difficulties in understanding what the task asked me to do with the texts	Task completion	Extraneous
[12] The explanations on how to read the texts have been confusing	Task completion	Extraneous

Data Analysis

Essay Task Coding Scheme

In order to capture multiple text comprehension and document integration in the essay task participants had to complete, a set of categories following similar coding schemas as in previous studies (Cerdán and Vidal-Abarca, 2008; Cerdán et al., 2013) were used and rated by two independent raters: the number of ideas literally extracted from the text (interrater reliability: $r = 0.925$); the number of inferences (i.e., reflecting the quality of a participant's mental model; $r = 0.890$); the number of trustworthy and untrustworthy ideas from the text (i.e., both reflect a participant's awareness of the quality of information; interrater reliability for both trustworthy and untrustworthy: $r = 1.000$); and integration (i.e., should reflect if a participant has done the effort to include information from all documents, which is what is expected in multiple document processing; $r = 0.966$). This measure was calculated by identifying the number of texts students were extracting their ideas from in their essays (ranging from 1 to 4). Given the high interrater reliabilities, scores of the two raters were averaged for each of these measures for each participant.

Cognitive Load (CL-MDR)

Confirmatory factor analysis (M*plus* 7.3, Muthén and Muthén, 2014) on the 12 items of the CL-MDR revealed a four-factor solution in line with **Table 1**: single text processing (items 6 and 7); multiple text processing (items 1–4); developing an intertext model (items 5 and 8); and extraneous cognitive load (items 9–12). Comparisons of treatment conditions on cognitive load were therefore based on these four factors.

Treatment Effects

Differences between the experimental treatment and control condition in terms of the three subtypes of intrinsic cognitive load, extraneous cognitive load, and essay performance were analyzed using Frequentist and Bayesian *t*-tests (JASP; JASP

TABLE 2 | A four-factor model for the CL-MDR yielding good fit.

Factor	Items	Loading each item	*R*-squared each item
Single text	6 and 7	0.861	0.748
Multiple texts	1–4	0.744	0.558
Intertext model	5 and 8	0.615	0.381
Extraneous	9–12	0.777	0.608

Team, 2016). Although Frequentist *p*-values may provide some evidence against a null hypothesis of "no difference," Bayes factors—such as obtained through Bayesian *t*-tests (Rouder et al., 2009)—can enable the study of both evidence against and in favor of a null hypothesis (Jeffreys, 1961; Rouder et al., 2009). In the context of an experiment, for example, a statistically non-significant *p*-value cannot be interpreted as there being "no treatment effect." Bayes factors indicate under which of two competing hypotheses—the "no treatment effect" null hypothesis or the "treatment effect" alternative hypothesis—the findings observed are more likely to have occurred.

RESULTS

CL-MDR

A four-factor model with equal loadings for items loading on the same factor and zero correlation between extraneous cognitive load (i.e., items 9–12) and the three other intercorrelated factors yielded good fit (CFI = 0.934; TLI = 0.945; RMSEA = 0.066 with a 90% confidence interval from 0.042 to 0.089). **Table 2** summarizes the results.

In the confirmatory factor model summarized in **Table 2**, the correlation between *single text* (i.e., items 6 and 7) and *multiple texts* (i.e., items 1–4) processing was 0.749 ($p < 0.001$), the correlation between *single text* processing and *intertext model* (i.e., items 5 and 8) was 0.590 ($p < 0.001$), and the correlation between *multiple texts* processing and *intertext model* was 0.261 ($p = 0.037$). Zero correlations between these three intercorrelated subtypes of intrinsic cognitive load and the extraneous cognitive load factor (i.e., items 9–12) is in line with cognitive load theory, where intrinsic and extraneous cognitive load are conceptually independent additive types of cognitive load (Leppink et al., 2014, 2015).

TABLE 3 | Means (M) and standard deviations (SD) per condition for each measure of cognitive load and essay performance.

Measure	Control ($n = 71$): *M* (*SD*)	Treatment ($n = 54$): *M* (*SD*)
CL-MDR items 6 and 7	3.197 (2.136)	2.870 (2.070)
CL-MDR items 1–4	3.810 (1.620)	3.324 (1.630)
CL-MDR items 5 and 8	3.824 (1.821)	4.028 (2.228)
CL-MDR items 9–12	1.866 (1.579)	0.694 (0.773)
Literal ideas	3.317 (1.801)	3.241 (1.930)
Inference	2.401 (1.446)	2.074 (1.478)
Trustworthy ideas	1.070 (0.816)	0.833 (0.746)
Untrustworthy ideas	1.423 (0.625)	1.315 (0.609)
Integration	2.458 (0.925)	2.130 (0.802)

Treatment Effects on Cognitive Load and Essay Performance

Table 3 presents means and standard deviations per condition for each measure of cognitive load and essay performance.

Table 4 summarizes the outcomes of Frequentist and Bayesian *t*-tests for the differences between the two conditions in terms of cognitive load and essay performance. **Table 3** indicates that differences between conditions in terms of CL-MDR and essay performance were on the small side except for extraneous cognitive load (i.e., CL-MDR items 9–12) with the latter being in the direction that is in line with cognitive load theory.

Table 4 indicates that there is some preference for the null hypothesis of "no difference" between conditions with regard to the three subtypes of intrinsic cognitive load (i.e., CL-MDR items 1–8) as well as with regard to essay performance, but that the conditions clearly differ in terms of extraneous cognitive load.

DISCUSSION

The findings have value to the cognitive load and multiple documents reading communities for a number of reasons. Firstly,

TABLE 4 | Outcomes of Frequentist and Bayesian *t*-tests for the differences between the two conditions in terms of cognitive load and essay performance: *t*-value (df = 123), *p*-value, 95% confidence interval (CI), and Bayes factors for the alternative hypothesis (i.e., there is a difference) vs. the null hypothesis (i.e., there is no difference) (BF_{10}) and for the null vs. the alternative hypothesis (BF_{01}).

Measure	$t_{(123)}$	*p*-value	95% CI[b]	BF^c_{10}	BF^c_{01}
CL-MDR items 6 and 7	−0.859	0.392	−1.080; 0.426	0.269	3.721
CL-MDR items 1–4	−1.656	0.100	−1.066; 0.095	0.661	1.512
CL-MDR items 5 and 8	0.563	0.575	−0.513; 0.921	0.222	4.501
CL-MDR items 9–12[a]	−5.012	<0.001	−1.635; −0.709	8086.425	<0.001
Literal ideas	−0.227	0.821	−0.740; 0.588	0.197	5.075
Inference	−1.242	0.217	−0.849; 0.194	0.386	2.589
Trustworthy ideas	−1.669	0.098	−0.518; 0.044	0.674	1.483
Untrustworthy ideas	−0.965	0.336	−0.329; 0.113	0.293	3.409
Integration[a]	−2.079	0.040	−0.641; −0.016	1.337	0.748

[a] *Although the assumption of homogeneity of variances was violated here, Welch's and Mann-Whitney's test provided similar outcomes: for CL-MDR items 9–12, both Welch's and Mann-Whitney's $p < 0.001$; for integration, Welch's $p = 0.036$, Mann-Whitney's $p = 0.019$.*

[b] *Positive differences indicate higher mean scores in the experimental treatment condition.*

[c] *BF_{01} is the inverse of BF_{10} (i.e., $BF_{01} = 1 / BF_{10}$). The presentation of both BF_{01} and BF_{10} makes it easier to evaluate the strength of evidence in favor of either of the two hypotheses under consideration. BFs and strength of evidence (e.g., Jeffreys, 1961): 1–3.2 = not worth more than a bare mention; 3.2–10 = substantial; 10–32 = strong; 32–100 = very strong; > 100 = decisive.*

the CL-MDR provides a measurement instrument that is well anchored in cognitive load theory (Sweller et al., 2011; Leppink and Van den Heuvel, 2015; Leppink et al., 2015) and previous research on the measurement of cognitive load (Leppink et al., 2013, 2014; Lafleur et al., 2015) as well as in contemporary theory on multiple documents reading such as MD-TRACE (Rouet and Britt, 2011) and models of purposeful reading such as RESOLV (Rouet et al., 2017).

Secondly, the current study provides empirical support for a four-factor model in a carefully designed randomized controlled experiment: not providing students with the opportunity to go back to the texts when writing an essay about these texts clearly increases extraneous cognitive load (i.e., split attention) but not intrinsic cognitive load. Although in the current study, the increase in extraneous cognitive load did not result in deteriorated essay performance, this is probably due to the fact that the combination of intrinsic and extraneous cognitive load was not such that it approached overload. Moreover, not allowing students to go back to texts appeared to favor the level of integration reflected in the essay text. This result could be interpreted in the light of the desirable difficulties framework proposed by Bjork et al. (2013). Creating some difficulties in the learning process might facilitate deep learning, but not superficial acquisition of information, a pattern we find in our data.

Thirdly, CL-MDR is the first cognitive load measurement instrument that unravels different sources of intrinsic cognitive load. Since cognitive load research has thus far more or less exclusively been carried out in a context of single document processing, other sources of intrinsic cognitive load identified in the current study have not yet emerged. The current study indicates that the study of cognitive load in a context of a more natural learning context—that of multiple document reading, which is of key importance in nearly all learning in secondary and higher education—is needed to unravel different subtypes of intrinsic cognitive load and as such to move toward a better understanding of intrinsic cognitive load.

In line with the finding of different subtypes of intrinsic cognitive load, future research might as well consider different subtypes of extraneous cognitive load. For instance, if stress or other emotions consume mental resources that could otherwise be used for the task at hand, that constitutes a source of extraneous cognitive load that may not or only partly be covered by the instructions (Leppink and Duvivier, 2016; Tremblay et al., 2016). Such emotions may be induced when students have to learn from multiple documents within a very limited period of time and/or when the documents concern a very complex topic or a topic that has emotional valence (e.g., certain political questions). In other words, future research should vary in topics as well as in time given to participants for task completion.

A second line of future research could focus on manipulations in one or more subtypes of intrinsic cognitive load while keeping extraneous cognitive load the same across conditions. An exemplar study is this context is that of Lafleur et al. (2015), who demonstrated that in the case of low extraneous cognitive load a modest increase in intrinsic cognitive load—for instance by requiring students to engage in reasoning more intensively—can actually stimulate learning. In such studies, one should find that the conditions in which one or more subtypes of intrinsic cognitive load is increased learning outcomes or performance measures are elevated as well.

In sum, the current study provides the cognitive load and multiple documents reading communities with a new instrument for the measurement of cognitive load that could be used in future experiments on multiple documents reading. We hope that this article provides researchers with various ideas for how to design such experiments.

ETHICS STATEMENT

Participants were secondary school students. Consent was provided by both the participant school and parents.

AUTHOR CONTRIBUTIONS

RC and CC, from ERI Lectura, were responsible of conducting the empirical study and the coding of data. RC and JL worked jointly in the design of the measurement instrument and empirical design. JL contributed very significantly with his expertise in cognitive load theory and data analysis. The three authors worked collaborately in writing up the empirical report.

REFERENCES

Bjork, R. A., Dunlosky, J., and Kornell, N. (2013). Self-regulated learning: beliefs, techniques, and illusions. *Annu. Rev. Psychol.* 64, 417–444. doi: 10.1146/annurev-psych-113011-143823

Braasch, J. L. G., McCabe, R. M., and Daniel, F. (2016). Content integration across multiple documents reduces memory for sources. *Read. Writ.* 29, 1571–1598. doi: 10.1007/s11145-015-9609-5

Britt, M. A., and Rouet, J. F. (2012). "Learning with multiple documents: component skills and their acquisition," in *Enhancing the Quality of Learning: Dispositions, Instruction, and Learning Processes,* eds M. J. Lawson and J. R. Kirby (Cambridge: Cambridge University Press), 276–314.

Cerdán, R., Marín, M. C., and Candel, C. (2013). The role of perspective on students' use of multiple documents to solve an open-ended task. *Psicol. Educ.* 19, 89–94. doi: 10.1016/S1135-755X(13)70015-0

Cerdán, R., and Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *J. Educ. Psychol.* 100, 209–222. doi: 10.1037/0022-0663.100.1.209

JASP Team (2016). *JASP (Version 0.7.5.6) [Computer software]*. Available online at: https://jasp-stats.org/ (Accessed December 5, 2016).

Jeffreys, H. (1961). *Theory of Probability, 3rd Edn*. Oxford: Oxford University Press.

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge: Cambridge University Press.

Lafleur, A., Côté, L., and Leppink, J. (2015). Influences of OSCE design on students' diagnostic reasoning. *Med. Educ.* 49, 203–214. doi: 10.1111/medu.12635

Leppink, J. (2014). Managing the load on a reader's mind. *Perspect. Med. Educ.*3, 327–328. doi: 10.1007/s40037-014-0144-x

Leppink, J. (2016). Cognitive load measures mainly have meaning when they are combined with learning outcome measures. *Med. Educ.* 50:979. doi: 10.1111/medu.13126

Leppink, J., and Duvivier, R. (2016). Twelve tips for medical curriculum design from a cognitive load theory perspective. *Med. Teach.* 38, 669–674. doi: 10.3109/0142159X.2015.1132829

Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., and Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behav. Res. Methods* 45, 1058–1072. doi: 10.3758/s13428-013-0334-1

Leppink, J., Paas, F., Van Gog, T., Van der Vleuten, C. P. M., and Van Merriënboer, J. J. G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learn. Instruct.* 30, 32–42. doi: 10.1016/j.learninstruc.2013.12.001

Leppink, J., and van den Heuvel, A. (2015). The evolution of cognitive load theory and its application to medical education. *Perspect. Med. Educ.*4, 119–127. doi: 10.1007/s40037-015-0192-x

Leppink, J., Van Gog, T., Paas, F., and Sweller, J. (2015). "Cognitive load theory: researching and planning teaching to maximise learning," in *Researching Medical Education,* eds J. Cleland and S. J. Durning (Chichester: John Wiley & Sons), 207–218.

Macedo-Rouet, M., Braasch, J. L. G., Britt, M. A., and Rouet, J. F. (2013). Teaching fourth and fifth graders to evaluate information sources during text comprehension. *Cogn. Instr.* 31, 204–226. doi: 10.1080/07370008.2013.769995

McCrudden, M. T., Tonje, S., Bråten, I., and Strømsø, H. I. (2016). The effects of topic familiarity, author expertise, and content relevance on Norwegian students' document selection: a mixed methods study. *J. Educ. Psychol.* 108, 147–162. doi: 10.1037/edu0000057

Muthén, L. K., and Muthén, B. O. (2014). *Mplus User's Guide.* Los Angeles, CA: Muthén & Muthén.

Perfetti, C. A., Rouet, J. F., and Britt, M. A. (1999). "Toward a theory of documents representation," in *The Construction of Mental Representations During Reading,* eds H. Van Oostendorp and S. R. Goldman (London: Lawrence Erlbaum Associates), 88–108.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* 16, 225–237. doi: 10.3758/PBR.16.2.225

Rouet, J. F. (2006). *The Skills of Document Use: From Text Comprehension to Web-Based Learning*. Mahwah, NJ: Erlbaum.

Rouet, J. G., and Britt, M. A. (2011). "Relevance processes in multiple document processing," in *Text Relevance and Learning from Text,* eds M. T. McCrudden, J. P. Magliano, and G. Schraw (Greenwich, CT: Information Age), 19–52.

Rouet, J. F., Britt, M. A., and Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educ. Psychol.* 52, 200–215. doi: 10.1080/00461520.2017.1329015

Rouet, J. F., Le Bigot, L., De Pereyra, G., and Britt, M. A. (2016). Whose story is this? Discrepancy triggers readers' attention to source information in short narratives. Read. Writ. 29, 1549–1570. doi: 10.1007/s11145-016-9625-0

Scharrer, L., and Salmerón, L. (2016). Sourcing in the reading process: introduction to the special issue. *Read. Writ.* 29, 1539–1548. doi: 10.1007/s11145-016-9676-2

Schroeder, S. (2011). What readers have and do: effects of students' verbal ability and reading time components on comprehension with and without text availability. *J. Educ. Psychol.* 103, 877–896. doi: 10.1037/a0023731

Strømsø, H. I., Bråten, I., and Britt, M. A. (2011). Do students' beliefs about knowledge and knowing predict their judgement of texts' trustworthiness? *Educ. Psychol.* 31, 177–206. doi: 10.1080/01443410.2010.538039

Sweller, J., Ayres, P., and Kalyuga, S. (2011). *Cognitive Load Theory*. New York, NY: Springer.

Tremblay, M. L., Lafleur, A., Leppink, J., and Dolmans, D. H. (2016). The simulated clinical environment: cognitive and emotional impact among undergraduates. *Med. Teach.* 39, 181–187. doi: 10.1080/0142159X.2016.1246710

Van Merriënboer, J. J. G., and Kirschner, P. A., (2012). *Ten Steps to Complex Learning: A Systematic Approach to Four-Component Instructional Design, 2nd Edn*. New York, NY: Taylor & Francis Group.

Van Merriënboer, J. J. G., and Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Med. Educ.* 44, 85–93, doi: 10.1111/j.1365-2923.2009.03498.x

Vidal-Abarca, E., Mañá, A., and Gil, L. (2010). Individual differences for self-regulating task-oriented reading activities. *J. Educ. Psychol.* 102, 817–826. doi: 10.1037/a0020062

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Cerdan, Candel and Leppink. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.