Journal of College Reading and Learning

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ucrl20

# Fostering University Students' Written Argumentation via Recursive Reading: A Randomized Controlled Trial

Christian Tarchi & Ruth Villalón

**To cite this article:** Christian Tarchi & Ruth Villalón (2022) Fostering University Students' Written Argumentation via Recursive Reading: A Randomized Controlled Trial, Journal of College Reading and Learning, 52:1, 42-63, DOI: 10.1080/10790195.2022.2021771

**To link to this article:** https://doi.org/10.1080/10790195.2022.2021771

Published online: 02 Feb 2022.

JOURNAL OF COLLEGE READING AND LEARNING
2022, VOL. 52, NO. 1, 42–63
https://doi.org/10.1080/10790195.2022.2021771

Routledge
Taylor & Francis Group



# Fostering University Students' Written Argumentation via Recursive Reading: A Randomized Controlled Trial

Christian Tarchi [a] and Ruth Villalón [b]

[a]Department of Education, Languages, Intercultures, Literatures and Psychology, University of Florence, (Florence, Italy); [b]Department of Education, Universidad de Cantabria, (Santander, Spain)

**ABSTRACT**

The present study analyzed the efficacy of a brief intervention aimed at scaffolding readers' "recursivity" (i.e., going back to the texts) while reading multiple texts and writing an argumentative essay. The participants were 151 university students, randomly assigned to two conditions: experimental (Recursivity-induced, RI) and active control (AC). We collected data about participants' thinking dispositions, perceived prior knowledge and perceived level of instruction in argumentative writing received, and prior beliefs. Then, students were assigned two texts about the evaluation of teachers, one pro and one against. RI students were prompted to compare the argumentation of each text with their own prior beliefs, whereas AC students were asked to write a summary of each text. Immediately after reading the texts and performing the accompanying tasks, RI and AC students were asked to write an argumentative essay to express their opinion on the topic. Process data were collected through the software Kidlogger. Results confirmed that the brief intervention improves students' analysis of the belief-inconsistent text, the overall argumentative quality of students' essay, and valid inferences made in a recall task one month after. The process analysis suggested that the intervention increases recursivity in at least a certain number of RI participants.

**KEYWORDS**

Written argumentation; multiple-texts comprehension; recall; thinking dispositions

## Introduction

In the present-day knowledge society, the evolution of the internet from web 1.0 to web 2.0 offered individuals tools to become content users and producers. As content users, people have easy access to a huge volume of information sources (Alexander & the Disciplined Reading and Learning Research Laboratory, 2012). As content producers, people have increasingly been called upon to understand and form opinions about complex and controversial topics, drawing on multiple, varied sources of information (List & Alexander, 2017). At the same time, the recent changes in the knowledge society have influenced higher education contexts, which introduce university students to learning tasks that are increasingly based on consulting,

**CONTACT** Ruth Villalón ruth.villalonr@unican.es Department of Education, (University of Cantabria), Avenida Los Castros, 5, Santander, ES, 39005, Spain

contrasting, and integrating information from multiple sources (Mateos et al., 2007). One example is the production of a written argumentative essay based on the reading of multiple sources, a complex task that requires organizing, selecting, and connecting information from several texts (i.e., multiple-documents comprehension; see, Barzilai et al., 2018). A challenge in this process is represented by the integration of conflicting information, both when reading and when writing. Indeed, the effortful processing of multiple documents is not always effective as it depends on readers' skills and dispositions (Bråten et al., 2013). Readers often struggle in integrating information from alternative perspectives, fail in evaluating the plausibility of arguments, and construct one-sided representations (Richter & Maier, 2017).

Several interventions have been designed to improve students' argumentative writing (Mateos et al., 2018), but they generally scaffolded either writing or reading. In the present study, we analyzed the efficacy of multiple-documents comprehension performances of a brief intervention aimed at scaffolding readers' "recursivity" (i.e., going back to the source texts and contrasting one's beliefs with them) while reading the texts and writing. The study contributes to the existing literature by focusing on recursivity, a neglected process of reading and writing about controversial topics, and it does so through a scaffold that can be easily implemented in multiple educational settings.

### *Multiple-texts Comprehension and Argumentative Writing*

Multiple-texts comprehension involves the "building of a coherent mental representation of an issue from the contents of multiple documents that deal with the same issue from different perspectives" (Bråten et al., 2013, pp. 322–23). When reading multiple texts, first readers need to represent the internal meaning of each text (i.e., textbase; Kintsch, 1998) and link it to relevant prior knowledge (i.e., situation model; Kintsch, 1998). Second, readers need to create an internal representation that integrates content across texts, including agreements and discrepancies in the accounts they read (i.e., situations model or integrated mental model) (see Documents Model: Britt et al., 1999; Perfetti et al., 1999).

The general organizational structure of the integrated model should be based on an argument schema (Bråten et al., 2011). Arguments are organized around a claim holding a central position, and they include supporting reasons, counter-arguments, and rebuttals for the main position (Toulmin, 1958). Thus, argumentative essay writing was used as a way to ascertain whether students construct an appropriate integrated mental model when reading from multiple sources (Primor & Katzir, 2018), besides being a construct directly targeted by research in the writing area. Writing an argumentative essay from sources is a complex process, with going beyond the perspectives presented in the texts and elaborating a coherent approach

(i.e., intertextual integration) being the core element (Kobayashi, 2015; De La Paz & Felton, 2010; Mateos et al., 2011, 2018). Intertextual integration can be achieved with a variety of strategies, from low-integration to high-integration. An example of a low-integration strategy is defending one perspective by identifying its supporting arguments and considering the arguments supporting the other perspective as wrong, irrelevant, or unsatisfactory (i.e., refutation). Examples of high-integration strategies are considering the pros and cons of both perspectives and identifying the stronger position (i.e., weighting); or identifying an intermediate position between the conflicting perspectives, which retains advantages and minimizes disadvantages (i.e., synthesizing).

In the last decade, research on multiple-document comprehension has examined the role of several cognitive factors, like prior knowledge (e.g., Strømsø et al., 2010) and critical thinking (Tarchi & Mason, 2020), as well as the role of motivational characteristics, like prior beliefs (Richter & Maier, 2017) and thinking dispositions (Bråten et al., 2014). The same processes are involved also in argumentative writing (Mason & Scirica, 2006; Nussbaum, 2005, 2008; Perkins et al., 1991; Tarchi & Villalón, 2021). Overall, people use their prior beliefs (as in the myside bias) or prior knowledge (as in the belief bias) to evaluate arguments and reasons (Macpherson & Stanovich, 2007), which leads to one-sided reasoning and poor intertextual integration in both multiple-documents comprehension and argumentative writing. It is important to improve students' biases and support their two-side reasoning, which in turn leads to deep learning (Felton et al., 2015; Kobayashi, 2015; Mateos et al., 2018; Voss & Van Dyke, 2001)

### *Improving Intertextual Integration*

Failure to integrate multiple information sources may result in biased understandings or conclusions based on unreliable data (Bråten et al., 2011), with negative consequences for personal decision-making and democratic deliberation on complex important problems (Barzilai et al., 2018). However, intertextual integration is a very complex and effortful process and interventions are needed to promote students' performances.

In their literature review on instructional approaches and practices promoting the integration of multiple texts, Barzilai et al. (2018) identified and analyzed 21 studies. Overall, effective interventions employed a variety of writing tasks (e.g., argument, inquiry, synthesis, and compare-and-contrast tasks). In all studies, students were taught integration processes. In a few studies, students were taught about text structures and the value of integration. Furthermore, as mentioned, most of the instruction programs focused on reading or writing. However, research has shown that providing instruction on reading can be a way to improve writing in upper elementary through high school (Graham et al., 2018). Teaching

reading may benefit writing in general terms but also multiple-texts tasks (De La Paz et al., 2017) and argumentative writing tasks (Harris et al., 2019). However, as far as we know, the reading and writing processes interplay has not been addressed when promoting effective argumentative essays from sources.

Nonetheless, several studies have highlighted that, when integrating multiple texts in a synthesis task, recursivity is a key element of successful students' performance. Recursivity is defined as the number of switches between writing and reading when processing source texts. Recursivity (i.e., going back to reading texts) can take place both while producing text and while revising the written text. Thus, students who follow a very linear process (reading first, writing second) produce low-quality texts. By contrast, those students who display more recursive patterns (going back and forth, contrasting source texts and their own text), produce better-written products and show better comprehension of the topic (Mateos & Solé, 2009; Solé et al., 2013). Recursivity may be an important approach for dealing with opposing positions since it may help to recognize the conflicting perspectives involved in the controversy. Interestingly, providing students with critical questions has been demonstrated to be as an effective tool to prompt argumentative essay writing (Mateos et al., 2018; Nussbaum & Edwards, 2011).

The notion that recursivity may be involved in text processing is supported by the levels-of-processing framework (Craik, 2002; Craik & Lockhart, 1972). In their seminal paper, Craik and Lockhart (1972) claimed that retention in memory depends on the depth of processing. Stimuli are processed at different levels: at a preliminary stage, physical and sensory features are analyzed, whereas at a later stage, readers engage in pattern recognition and extraction of meaning. Later stages correspond to a deeper processing of the stimuli, which in turn corresponds to a greater degree of semantic or cognitive analysis. Levels of processing are placed on a continuum of analysis: Rather than proceeding from one stage to another, readers can process stimuli by re-circulating information at one level of processing; however, mere re-circulation of information in the reader's memory system does not guarantee subsequent retention. Indeed, if stimuli are re-inserted in preliminary stages of processing, the trace is lost once attention is diverted. The repeated presentation of stimuli is beneficial to memory retention depending on whether the repetition is processed at the same level or encoded differently on its further presentation. Retention is not simply a function of longer processing time or repetitions but depends on how the individual processes the information each time.

### *The Present Study*

Considering these previous studies, and following Mateos et al. (2018), we designed a brief intervention based on critical questions aimed at guiding students' attention toward comparing two positions presented in two different

texts. The brief intervention consisted of scaffolds that were provided before and throughout the reading process, rather than just before writing. There are several evidence-based interventions supporting argumentation available in the literature (Barzilai et al., 2018; Brante & Strømsø, 2018); however, implementing an intervention could be costly in terms of time and effort. Providing a brief intervention mainly based on modifying instructions and providing scaffolds may be a more suitable alternative with large teaching groups, as are common in postsecondary settings. For that reason, in this study, we tested the use of a brief intervention to foster a central aspect of an integration process, recursivity, to support undergraduates' reading when writing an argumentation. The efficacy was tested on two outcomes, intertextual integration and deep comprehension. The former is defined as the ability to go beyond the perspectives presented in the texts and elaborate a coherent approach (Kobayashi, 2015; De La Paz & Felton, 2010; Mateos et al., 2011, 2018). The latter is defined as the ability to integrate information extracted from the text to prior knowledge to construct a situation model of the text through connecting inferences (Kintsch, 1998). Within this context, we have implemented an intervention based on explicit instruction and a scaffold. Its elements encourage students to come back and forth between the source texts and to contrast their prior ideas and the arguments presented to reach their own integrative conclusion. The intervention's design principles are explained in Table 1, following the reporting suggestions of Rijlaarsdam et al. (2017). Individuals in the active control condition were asked to perform a summary of both sources, which is a task that favors text deep processing (Gil et al., 2010).

Intertextual argumentation was measured through an argumentative task. Deep comprehension was measured through a text recall task. Past studies have found that these two constructs may be dissociated (Diakidoy et al., 2015, 2017). Prior topic beliefs, perceived topic knowledge, prior exposure to argumentative tasks, and thinking dispositions and were included as control variables. Prior topic beliefs are a fundamental control variable to assess as it represents the first cause of one-sided argumentation. Perceived topic knowledge can be considered as a proxy for topic knowledge (Andiliou et al., 2012; Stanovich & West, 2008) and has been used in prior studies on multiple-texts comprehension (Bråten et al., 2016). Perceived exposure to argumentative tasks was included as a measure of students' familiarity with the argumentative task. Finally, prior studies suggested that thinking dispositions are involved in both multiple-texts tasks. Specifically, Dai and Wang (2007) found that need for cognition (i.e., the disposition to engage in and enjoy activities that require thinking) should foster a deeper elaboration of texts and a more integrated situation model. In a study on multiple-texts comprehension conducted by the authors of the present study, need for cognition was indirectly associated with the argumentation quality of essays

**Table 1.** Intervention description.

| Design principle | Phase | Learning activity | | Instruction that leads to learning activity | |
|---|---|---|---|---|---|
| | | Description | Explanation | Description | Explanation |
| Representation of the task and attribution of meaning. | Problem centered. | Noticing the objective of the assignment. | This learning activity is aimed at motivating the student and focuses his/her attention on the task. | The student begins reading the instructions. The first paragraph briefly introduces the assignment. The second points out that students will face a real controversy in the field of Education and noticing that writing an essay is a common task in higher education. | This element in the instructions leads to an understanding and involvement in the task by providing a representation of the objective of the assignment. |
| Representation of the task and attribution of meaning. | Activation of existing knowledge. | Noticing how students usually deal with an essay. | This learning activity's goal is to provide a non-example for the concept of argumentative essay. | Reading of the second part of the instructions. The paragraph explains that a persuasion goal is the most frequent approach when dealing with an argumentation. | This element in the instruction activates prior knowledge on argumentative essays and explains to the student how he/she should not approach the task. |
| Meaningful verbal learning. | Demonstration of new knowledge. | Observing a new approach. | This learning activity aims to foster meaningful learning about how to deal with an argumentative essay to integrate. | The next part presents a desirable approach. It shares a new definition of the assignment as an "integration task" and explains the processes that should be developed in a recursive way (non-linear steps):<br>• comprehend each position: identifying and exploring different position on a topic<br>• Compare and contrast the positions, giving value to their supporting arguments.<br>• Draw conclusions, thinking about new positions that could overcome the limitations of the prior ones.<br>• Write the argumentative essay. | This element in the instruction aims to offer an explanation of what an integrative solution goal is and how to carry it out. |
| Learner's activity. Self-regulated learning. | Applying new knowledge. | Practicing answering critical questions. | This learning activity provides examples of questions which help to analyze a source text and demand the student to practice with them. | After reading text 1, the scaffold shows exactly what reflective questions on the text the student should make to deal with an integration task. | This element in the scaffold leads to compare and contrast student's prior beliefs and the position of text 1, then draw a conclusion. |

(*Continued*)

**Table 1.** (Continued).

| Design principle | Phase | Learning activity | | Instruction that leads to learning activity | |
|---|---|---|---|---|---|
| | | Description | Explanation | Description | Explanation |
| Learner's activity. Self-regulated learning. | Applying new knowledge. | Practicing answering critical questions. | This learning activity provides examples of questions which help to analyze two source texts and demand the student to practice with them. | Students are asked a new pack of reflective questions after reading text 2. | This element in the scaffold leads the student to compare and contrast text 1, text 2, and his/her prior beliefs and current position, then draw a conclusion. |

(Tarchi & Villalón, 2021). We hypothesized that the brief intervention would be effective in improving students' recursivity while writing, the quality of their argumentative essay, and the number of literal and valid inferences recalled from the texts.

## Method

### *Participants*

The participants were 151 university students (Age = 20.47 ± 2.16; 133 females, 18 males). They were enrolled in an Educational Psychology course, offered within the School of Education curriculum, in a large university in Italy. Students volunteered to participate and received a bonus for their collaboration. All participants were Italian and spoke Italian as their primary language. None of the participants was diagnosed with a learning disorder. The sample was relatively homogeneous (i.e., middle class) regarding socioeconomic status, as assessed through parents' occupation. Informed consent was obtained from all participants. The study followed all the indications of the Declaration of Helsinki (World Medical Association, 2013) and was approved by the Ethics Committee of the University of Florence (Italy).

### *Research Design and Procedure*

Research hypotheses were investigated with a randomized controlled trial research design, with pretest and posttest assessment. Students worked individually on a computer to complete all the tasks. In the pretest stage, we collected data about participants' prior topic beliefs, perceived topic knowledge, perceived expertise in argumentative tasks, and need for cognition in one single session. Then, in the following session, participants were randomly assigned to a condition, experimental (Recursivity-induced, RI, $n$ = 76) or active control (AC, $n$ = 75). In both conditions, students were assigned two

texts about the evaluation of teachers, one pro and one against. The two texts used in this study were derived from Mateos et al. (2018) and translated from Spanish into Italian by a native speaker. The texts were similar in length (640 words for the pro-evaluation text and 769 words for the against-evaluation text). The texts were balanced for reading difficulty, with a Gulpease index (Lucisano & Piemontese, 1988) of 42 for the pro-evaluation text and 46 for the against-evaluation text (within junior secondary school level: 11 to 13 age range); thus both were appropriate for undergraduate students. Both texts included nine arguments, and some of the supporting reasons were interconnected, so most arguments could be considered as counterarguments for the other position. For example, the pro-evaluation text claimed that parents should participate in the evaluation of teachers, and the against-evaluation text claimed that families are not reliable sources to be used to evaluate teachers because they are not professionals. To control for confounding variables, we removed indications about the source and presented texts as trustworthy. Students received the following instructions: "*Please, read these two texts discussing two different positions on a controversial topic in Education. Your task is to write an argumentative essay in which you discuss your perspective, taking into consideration what you have read and your own ideas*." The texts' presentation order was counterbalanced within conditions.

At the end of the first text, RI students were asked to write down the text position, the supporting argumentations, whether the text position was in line or not with their own prior beliefs, and whether they changed their prior beliefs after reading it. At the end of the second text, the students were asked the same questions, plus whether there was a difference in position between texts. AC students were asked to write a summary at the end of each text. Students had unlimited time to read the texts and answer the questions. See Appendix A for a full description of the material.

Immediately after reading the texts, RI and AC students were asked to write an argumentative essay to express their opinion on the topic. While students were writing the essay, computer activity was captured with the Kidlogger (kidlogger.net) Windows freeware. All students had access to the texts while answering the questions and while writing the essay. The texts were available on a separate window on the screen. Thus, texts were not always visible, but participants had to purposely switch to them if they wanted to read them again while writing the answers or the essay.

One month after, students were asked to recall all that they could from the texts read.

From the initial sample, 14 students were excluded from data-analysis as they did not submit an argumentative essay or the recall production, or their process data were incomplete. The final sample included 137 students (70 in the RI condition and 67 in the AC condition).

## Measures

### Perceived Topic Knowledge

Students were asked to self-report their perceived prior knowledge on the topic of teachers' evaluation (1 item on a 1–6 scale, from 1 = extremely low to 6 = extremely high).

### Perceived Expertise in Argumentative Tasks

Students were asked to self-report their perceived level of instruction in argumentative writing through three items, one asking how much instruction in argumentative writing they had received, one asking them how confident they felt on performing well in an argumentative writing task, and one asking them how confident they felt on performing well in a task-based on reading argumentative texts (three items on a 1–6 scale, from 1 = extremely low to 6 = extremely high).

### Prior Topic Beliefs

Students' prior topic beliefs were collected through an 8-item self-report questionnaire on a 6-point Likert-scale (from 1 = completely disagree to 6 = completely agree; see Appendix B for the full questionnaire). An example of an item was: "Students' outcomes cannot be simply attributed to the influence of a single professor"). The scoring of items with negative valence was reversed, and a sum score was calculated. Scores could range between 8 and 48 ($\alpha = .71$).

### Need for Cognition

Need for cognition was assessed through an Italian translation of the original 18-item instrument developed by Cacioppo and Petty (1982). Each item (e.g., "I would prefer complex to simple problems") was rated on a 5-point Likert-scale (from 1 = absolutely false for me to 5 = absolutely true for me). The scoring of items with negative valence was reversed, and a sum score was calculated. Scores could range between 18 and 90 ($\alpha = .86$).

### Argumentative Quality

Students were asked to write an argumentative essay reporting their stance. A research assistant (expert in school psychology) coded the essays written by the students following the coding system developed by Mateos et al. (2018). The research assistant was specifically trained by one of the authors of the coding system. The essays were first analyzed to identify the number of claims from both texts, and then coded on six progressive levels of integration, taking into account the type of claims and the type of final conclusion: 0) non-source-based (opinion not based on the arguments presented on the source texts); 1) neutral position (a clear standpoint is not identified because arguments of the two positions are presented but not integrated); 2)

supporting a position (one of the positions is defended, basically considering its arguments and not the other view); 3) integration of two positions via refutations (reasons of both positions are considered, but the opposite view's arguments are just refuted; 4) minimal integration via weighting or synthesis (one position is defended, but at least two arguments of the other view are valued and the conclusion is partial); 5) partial integration via weighting or synthesis (one position or both are claimed, including two arguments of the both views integrated, but the final conclusion is missing or partial; 6) full integration via weighting or synthesis (the essay concludes with a real overall conclusion, considering several arguments of both positions integrated). The second author of the present study acted as an independent rater and coded 50% of the essays, which were randomly selected. The inter-rater agreement was $k = .78$. Disagreements were solved through discussion. From this task, the following measures were derived: essay length (i.e., number of words in the essay) and argumentation quality.

### *Recursivity*

The number of switches between the word processor software and the texts was counted as a measure of recursivity.

### *Recall Task*

One month after reading the texts, students were asked to recall what they had read (without access to the texts). A research assistant (expert in school psychology), specifically trained by one of the authors of the present study, coded the recall protocols written by the students following the coding system developed by Diakidoy et al. (2015). Recall protocols were parsed into clauses ($k = .93$), and each clause was identified either as an explicit idea recalled from the texts (verbatim or paraphrase) or as a valid inference. The first author of the present study acted as independent rater and coded 50% of the essays, randomly selected ($k = .79$). Disagreements were solved through discussion. From this task, the following measures were derived: the number of literal claims recalled from the texts and the number of valid inferences.

## Results

Descriptive results are showed in Table 2. Overall, students' prior topic beliefs were pro-evaluation (scores could range between 8 and 48, with a median of 31 and a mean of 31.44) and reported an average knowledge of the topic (score could range between 1 and 6, mean = 2.99). On average, essays written were 526 words long, with an overall medium level of integration (mean = 3.04, median = 2). In the recall task, the number of literal and inferential propositions recalled was overall low (means, respectively, of 3.87 and 2.14).

**Table 2.** Descriptive results for the total sample, and by condition, RI ($N = 70$) and Control ($N = 67$).

| | Total | | | RI | AC |
|---|---|---|---|---|---|
| | Min. | Max. | M± SD | M± SD | M± SD |
| Need for cognition | 37 | 82 | 60.42 ± 8.79 | 61.67 ± 8.08 | 59.12 ± 9.36 |
| Perceived topic knowledge | 1 | 5 | 2.99 ± 1.06 | 3.06 ± 1.05 | 2.91 ± 1.08 |
| PC in writing | 1 | 5 | 2.77 ± 1.04 | 2.76 ± 1.05 | 2.77 ± .98 |
| PC in reading | 1 | 6 | 3.91 ± .95 | 3.80 ± 1.00 | 4.02 ± .88 |
| PC instruction | 1 | 6 | 3.43 ± 1.17 | 3.59 ± 1.19 | 3.26 ± 1.15 |
| Prior topic beliefs | 23 | 39 | 31.44 ± 3.29 | 31.14 ± 3.42 | 31.76 ± 3.13 |
| Essay length | 129 | 1400 | 525.67 ± 221.43 | 548.76 ± 208.98 | 501.55 ± 232.85 |
| Recursivity | 0 | 5 | 1.15 ± 1.35 | 1.34 ± 1.53 | .94 ± 1.11 |
| Argumentation quality | 0 | 6 | 3.04 ± 1.61 | 3.32 ± 1.81 | 2.75 ± 1.32 |
| Recall of literal claims | 0 | 11 | 3.87 ± 2.78 | 3.58 ± 2.72 | 4.18 ± 2.83 |
| Recall of valid inferences | 0 | 9 | 2.14 ± 1.67 | 2.53 ± 1.79 | 1.72 ± 1.44 |

*Note.* PC = Perceived competence

The analysis of correlations suggests that perceived prior topic knowledge was negatively associated with the number of words included in the argumentative essays. Among prior knowledge variables, only prior exposure to instruction in argumentative writing was associated with the quality of argumentative essays. Need for cognition was positively associated with the length and the quality of argumentative essays and with recall of valid inferences. Finally, performances in the argumentative task were not associated with performances in the recall task (see, Table 3).

A composite score for the three measures of prior exposure to argumentative task (reading, writing, and instruction) was created through a principal component analysis [$KMO = .61$; Bartlett, $\chi^2 = 65.73$, $df = 3$, $p < .001$; variance explained = 59.14%]. The principal component extract was defined as perceived competence in argumentation. Participants in the two conditions did not significantly differ in any of the pretest measures, as assessed through independent-samples *t*-tests (see Table 4).

**Table 3.** Correlations among variables.

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Need for cognition | 1 | .04 | .17 | .10 | .20* | .14 | .19* | .17* | −.09 | .18* |
| 2 | Perceived topic knowledge | | 1 | .65** | .28** | .24** | −.02 | −.20* | −.08 | .02 | −.01 |
| 3 | PC in writing | | | 1 | .51** | .38** | −.02 | −.09 | .08 | .08 | .09 |
| 4 | PC in reading | | | | 1 | .30** | .28** | .01 | .14 | −.06 | .10 |
| 5 | PC instruction | | | | | 1 | .01 | .16 | .25** | −.01 | .04 |
| 6 | Prior topic beliefs | | | | | | 1 | .10 | .04 | .02 | .22** |
| 7 | Essay length | | | | | | | 1 | .38** | .25** | .20* |
| 8 | Argumentation quality | | | | | | | | 1 | .14 | .12 |
| 9 | Recall of literal claims | | | | | | | | | 1 | −.04 |
| 10 | Recall of valid inferences | | | | | | | | | | 1 |

*Note.* $^{**}p < .01$, $^{*}p < .05$. PC = Perceived competence

**Table 4.** Condition differences in pretest measures.

| | *t* | *df* | *p* | 95%CI |
|---|---|---|---|---|
| Need for cognition | −1.72 | 135 | .09 | −5.51; .39 |
| Perceived topic knowledge | −.812 | 135 | .42 | −.51; .21 |
| PC in argumentation | −.004 | 135 | .99 | −.34; .34 |
| Prior topic beliefs | 1.09 | 135 | .28 | −.50; 1.72 |

### Efficacy of the Intervention

The research questions were explored with a generalized linear model with the condition as independent factor; perceived topic knowledge, perceived competence in argumentation, and essay length as control variables, and argument quality of students' essays, the number of literal claims recalled, and the number of valid inferences recalled as dependent variables. The two conditions did not differ in time spent on the writing task, $t = .33$, $df = 133$, $p = .74$, 95% CI = [−43.48, 31.02].

Overall, the condition variable contributed to explaining variance in the set of dependent variables [$F = 4.89$, $p < .01$, ${}_{P}\eta^2 = .10$]; after the effects of perceived topic knowledge, perceived competence in argumentation and essay length were controlled (see Table 5). More specifically, the RI condition outperformed the AC condition in the quality of the argumentative essays [$F = 3.97$, $p < .05$, ${}_{P}\eta^2 = .03$] and in the number of valid inferences recalled [$F = 7.62$, $p < .05$, ${}_{P}\eta^2 = .06$] (see Table 6).

### Moderation Effect of Perceived Topic Knowledge and Perceived Competence in Argumentation

To control for aptitude per treatment effects (Cronbach & Snow, 1977), we conducted two moderation analyses with the PROCESS macro for SPSS (Hayes, 2012). The effects of the interactions between the condition and

**Table 5.** Generalized linear model: Results from the multivariate test on argument quality of students' essays, number of literal claims recalled, and number of valid inferences recalled as dependent variables.

| Source | Wilks' λ | F | df | p | ${}_{P}\eta^2$ |
|---|---|---|---|---|---|
| Perceived topic knowledge | .97 | 1.58 | 3, 130 | .20 | .04 |
| Perceived competence in argumentation | .92 | 3.56 | 3, 130 | .02 | .08 |
| Essay length | .82 | 9.60 | 3, 130 | <.01 | .18 |
| Condition | .90 | 4.89 | 3, 130 | <.01 | .10 |

**Table 6.** Generalized linear model: Results from the univariate tests.

| | Argument quality | | | Recall of literal claims | | | Recall of inferential claims | | |
|---|---|---|---|---|---|---|---|---|---|
| Source | *F* | *p* | ${}_{P}\eta^2$ | *F* | *p* | ${}_{P}\eta^2$ | *F* | *p* | ${}_{P}\eta^2$ |
| Model | 8.86 | <.01 | .21 | 3.20 | .02 | .09 | 3.73 | .01 | .10 |
| Perceived topic knowledge | 2.84 | .09 | .02 | 1.37 | .24 | .01 | .26 | .61 | <.01 |
| Perceived competence in argumentation | 8.50 | <.01 | .06 | .37 | .55 | <.01 | 1.59 | .21 | .01 |
| Essay length | 15.90 | <.01 | .11 | 11.00 | <.01 | .08 | 3.59 | .06 | .03 |
| Condition | 3.97 | .04 | .03 | 3.06 | .08 | .02 | 7.62 | .01 | .06 |

*Note.* Degrees of freedom = 4, 132

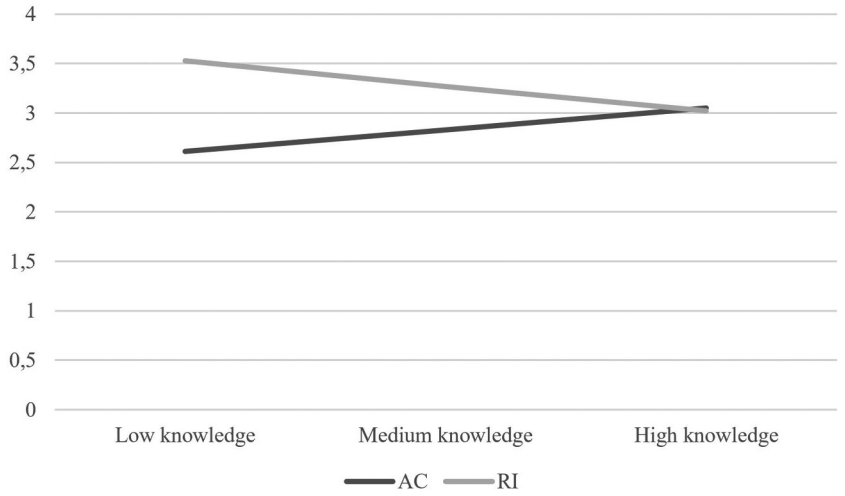


**Figure 1.** Graphical representation of the moderation effect of condition on the association between perceived topic knowledge, represented on the horizontal axis, and argumentation quality, represented on the vertical axis (AC = active control; RI = recursivity intervention).

perceived topic knowledge, and the effects of the interactions between the condition and perceived competence in argumentation on the quality of argumentative essays were tested, with essay length included as a control variable. The first moderation model was statistically significant, $R^2$ = .19, $F$ (4, 132) = 7.50, $p$ < .01, and so was the interaction between the condition and perceived topic knowledge, $R^2change$ = .02, $F$ (1, 132) = 3.94, $p$ < .05. Specifically, low knowledge students benefited from the intervention more than high knowledge students did (see Figure 1). It must be noted, however, that the effect size of this interaction was small.

The second moderation model was statistically significant, $R^2$ = .20, $F$ (4, 132) = 8.19, $p$ < .01, but the interaction between the condition and perceived competence in argumentation was not, $R^2change$ = .00, $F$ (1, 132) = .71, $p$ > .05.

### *Condition Differences in Recursivity Behavior*

The condition differences in frequency of recursive reading while writing the argumentative essay were investigated through a $t$-test for independent samples. RI students exhibited a tendency for more recursive behavior than AC students did [$t$ = 1.84, $p$ = .07, Cohen's $d$ = .30].[1] Specifically, a higher percentage of RI students went back to read the text from three to five times than AC students did (see Figure 2 and Table 7). However, it must be noted that the majority of students in both conditions displayed no or little recursive behavior.

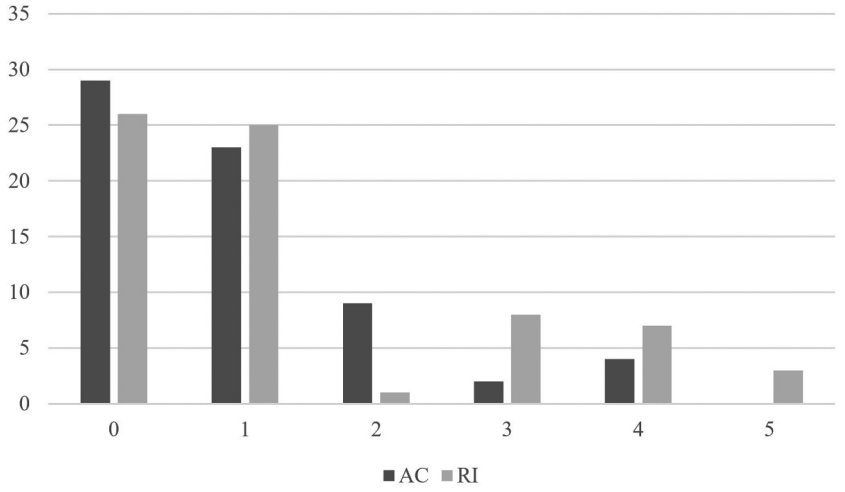


**Figure 2.** Graphical representation of frequencies of recursivity across conditions (AC = active control; RI = recursivity intervention).

**Table 7.** Contingency table condition X recursivity (raw scores and percentages).

| | | | Recursivity | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | 0 | 1 | 2 | 3 | 4 | 5 |
| **Condition** | **AC** | Raw score | 30 | 4 | 9 | 2 | 4 | 0 |
| | | Within-condition % | 61.2 | 8.2 | 18.4 | 4.1 | 8.2 | 0.0 |
| | | Within-recursivity % | 53.6 | 44.4 | 90.0 | 20.0 | 36.4 | 0.0 |
| | **RI** | Raw score | 26 | 5 | 1 | 8 | 7 | 3 |
| | | Within-condition % | 52.0 | 10.0 | 2.0 | 16.0 | 14.0 | 6.0 |
| | | Within-recursivity % | 46.4 | 55.6 | 10.0 | 80.0 | 63.6 | 100.0 |

## Discussion

The present study analyzed the efficacy of a brief intervention aimed at scaffolding readers' "recursivity" (i.e., going back to the texts) while comprehending multiple texts. Students' outcomes were assessed in terms of quality of an argumentative essay written immediately reading the texts and recall of literal and inferential information from texts one month after having read them.

Overall, results confirmed the efficacy of the brief intervention aimed at increasing students' recursivity when reading texts and when writing the essays. Recursivity-induced students wrote an argumentative essay of better quality, that is, with a higher level of intertextual integration, and recalled more valid inferences, i.e., achieved a deeper comprehension. These results are relevant for two reasons: they extend to multiple-texts comprehension the notion that writing processes may benefit from reading interventions (Graham et al., 2018), and they emphasize the contribution of recursivity in integrating multiple perspectives and depth of comprehension (Mateos & Solé, 2009; Solé et al., 2013). Most intervention studies on multiple-texts comprehension scaffold reading (e.g., sourcing interventions; see Brante & Strømsø, 2018) or

writing (e.g., integration interventions; see Barzilai et al., 2018, Luna et al., 2022), whereas results of this study represent a reminder of the highly interactive nature of these two processes, an aspect that research on multiple-texts comprehension should target more.

Although the scores were overall low, a score of three represents the first level of integration (even if through a low-integration strategy such as refutation). Thus, the average of 3.3 achieved by RI students can be considered (minimally) satisfying. It must be pointed out that the aim of the present study was to test the efficacy of a minimal intervention, but past studies have discussed other ways to instruct students about the argumentative task: in past studies about promoting intertextual integration, students were taught integration processes, or how to integrate; they learned about text structures; and they were made aware of the value of integration. The most frequent instructional practices were explicit instruction of integration, collaborative discussions and practice, graphic organizers or representations, modeling integration processes, and individual practice (see review by Barzilai et al., 2018). The efficacy of recursivity may increase if coupled with other evidence-based ways to promote integration, for instance, by encouraging learners to integrate their new knowledge about recursivity and integrative solutions into their everyday lives (Rijlaarsdam et al., 2017).

The process analysis revealed contradictory findings about whether the intervention was effective or not in increasing recursivity. Most of the students from both conditions displayed little to no recursivity even after the intervention. Notably, conditions varied across more dimensions than recursivity, as students in the RI condition were also provided with strategy instruction. Thus, the intervention may have induced an increase in strategic processing in some students, an increase of recursivity in other students, and an increase in both in a few students. The effectiveness of the intervention may depend on participants' individual differences, such as topic knowledge. Indeed, the moderation analyses showed that the intervention was effective in students with a low level of perceived knowledge. Probably, students who perceived themselves as knowledgeable are more resistant to external scaffolds.

According to the results of the correlational analysis, there was no association between argumentative essay quality and recalled valid inferences. Although this result may be counterintuitive, as it is reasonable to think that a higher integration between conflicting perspectives would lead to a deeper comprehension; past studies have found a similar pattern (Diakidoy et al., 2015, 2017). Such a dissociation may exist because in the essay task arguments are evaluated but information is not necessarily deeply elaborated. Indeed, integration can be achieved with different levels of elaboration (Mateos et al., 2018; Nussbaum, 2008; Nussbaum & Edwards, 2011). This aspect warrants more attention from research, and future studies should investigate the argumentative quality of essays by also including measures of depth of elaboration. The key variable of this study, recursivity, was positively associated with depth

of comprehension, as RI participants outperformed AC participants in recall of valid inferences. The dissociation between the two outcomes supports the hypothesis of an independent contribution of recursivity on depth of comprehension, refuting the hypothesis of a transfer of the effect from one outcome to the other one.

Results of the present study suggest that perceived topic knowledge was negatively associated with essay length, a variable that is considered as an index of engagement (Eisenberger et al., 1982; Latini et al., 2019) and a mediator of essay quality (MacArthur et al., 2019). This result is in line with studies on students' calibration, that is, the degree to which individuals' judgments about their knowledge corresponds to the knowledge they manifest (Alexander, 2013). Poor calibration can undermine performance, especially in complex tasks such as multiple-texts integration. Conversely, perceived competence in argumentative essay writing seems associated with better calibration, as it was positively associated with the quality of the argumentative essay. Overall, students seemed better at calibrating competence than knowledge.

### *Limitations and Directions for Future Research*

When interpreting the findings of the current study, some limitations should be taken into account. First, the effect sizes of the association between the condition and multiple-texts comprehension outcomes were small, but it must be noticed that the intervention was minimal, and the students' difficulties in writing an argumentative essay have been outlined (Mateos et al., 2018). The second limitation depends on the topic chosen for texts assigned to students. Socio-scientific topics may have triggered more biased beliefs, which could have influenced the efficacy of the brief intervention. Future studies should replicate the present research design by varying the controversial topic to assign to students. Third, the research design did not allow us to investigate the effects of the brief intervention on the actual reading and writing behavior. Moreover, the intervention included both a scaffold for recursivity and strategy instructions, making it difficult to attribute the increase of performances to a specific aspect of the intervention. The process analysis suggests that to a certain extent recursivity was successfully addressed; however, future studies should compare the performances of a full intervention condition with the performances of a condition with recursivity instructions only to determine the added value of strategy instruction. Furthermore, the last phase of Merrill's design principles was not addressed in our intervention, so future studies could improve the power of the scaffold, for example, by adding a practicing task (Rijlaarsdam et al., 2017). Thus, students could use the scaffold for another controversy and use their essay to perform a debate with their peers in the classroom, so they could consolidate their skills and have the opportunity to demonstrate them in public.

### *Conclusion*

Despite the limitations, the present study contributes to the literature on multiple-texts comprehension and argumentative writing by addressing a partially neglected factor of reading and writing behavior, recursivity. It also contributes to the literature by bridging reading and writing and demonstrating how writing outcomes can be improved by scaffolding reading even when engaged with multiple texts. A strength of the present research design is represented by the implementation of both a proximal and a distal measure of text processing, allowing investigation of the association between argumentation and depth of comprehension. On one hand, these two fundamental processes seem to be dissociated, probably because of the low level of integration that students can achieve. On the other hand, recursivity seems to be associated with both outcomes, suggesting that more importance should be given to this behavior in multiple-texts comprehension interventions.

## Note

1. Some scholars have considered .10 as an acceptable threshold for determining statistical significance, although the use of less conservative levels of alpha may increase the risk of type I error (see Schumm et al., 2013).

## Acknowledgments

This work was partially funded by COST ACTION IS1401ELN through two Short Term Scientific Missions grants.

## Disclosure Statement

No potential conflict of interest was reported by the author(s).

## ORCID

Christian Tarchi http://orcid.org/0000-0003-4013-4794
Ruth Villalón http://orcid.org/0000-0003-1600-8026

## References

Alexander, P. A., & the Disciplined Reading and Learning Research Laboratory. (2012). Reading into the future: Competence for the 21st century. *Educational Psychologist*, *47*(4), 259–280. https://doi.org/10.1080/00461520.2012.722511

Alexander, P. A. (2013). Calibration: What is it and why it matters? An introduction to the special issue on calibrating calibration. *Learning and Instruction*, *24*, 1–3. https://doi.org/10.1016/j.learninstruc.2012.10.003

Andiliou, A., Ramsay, C. M., Murphy, P. K., & Fast, J. (2012). Weighing opposing positions: Examining the effects of intratextual persuasive messages on students' knowledge and beliefs. *Contemporary Educational Psychology*, *37*(2), 113–127. https://doi.org/10.1016/J.CEDPSYCH.2011.10.001

Barzilai, S., Zohar, A. R., & Mor-Hagani, S. (2018). Promoting integration of multiple texts: A review of instructional approaches and practices. *Educational Psychology Review*, *30*(3), 973–999. https://doi.org/10.1007/s10648-018-9436-8

Brante, E. W., & Strømsø, H. I. (2018). Sourcing in text comprehension: A review of interventions targeting sourcing skills. *Educational Psychology Review*, *30*(3), 773–799. https://doi.org/10.1007/s10648-017-9421-7

Bråten, I., Anmarkrud, Ø., Brandmo, C., & Strømsø, H. I. (2014). Developing and testing a model of direct and indirect relationships between individual differences, processing, and multiple-text comprehension. *Learning and Instruction*, *30*, 9–24. https://doi.org/10.1016/j.learninstruc.2013.11.002

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist*, *46*(1), 48–70. https://doi.org/10.1080/00461520.2011.538647

Bråten, I., Ferguson, L. E., Anmarkrud, Ø., & Strømsø, H. I. (2013). Prediction of learning and comprehension when adolescents read multiple texts: The roles of word-level processing, strategic approach, and reading motivation. *Reading and Writing*, *26*(3), 321–348. https://doi.org/10.1007/s11145-012-9371-x

Bråten, I., Salmerón, L., & Strømsø, H. I. (2016). Who said that? Investigating the plausibility-induced source focusing assumption with Norwegian undergraduate readers. *Contemporary Educational Psychology*, *46*, 253–262. https://doi.org/10.1016/j.cedpsych.2016.07.004

Britt, M. A., Perfetti, C. A., Sandak, R., & Rouet, J.-F. (1999). Content integration and source separation in learning from multiple texts. In S. R. Goldman, A. C. Graesser, & P. van Den Broek (Eds.), *Narrative, comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 209–233). Erlbaum.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*, 116–131.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I. M. (2002). Levels of processing: Past, present . . . and future? *Memory*, *10*(5–6), 305–318. https://doi.org/10.1080/09658210244000135

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York, NY: Irvington.

Dai, D. Y., & Wang, X. (2007). The role of need for cognition and reader beliefs in text comprehension and interest development. *Contemporary Educational Psychology*, *32*(3), 332–347. https://doi.org/10.1016/j.cedpsych.2006.05.002

De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology*, *35*(3), 174–192. https://doi.org/10.1016/j.cedpsych.2010.03.001

De La Paz, S., Monte-Sano, C., Felton, M., Croninger, R., Jackson, C., & Worland Piantedosi, K. (2017). A historical writing apprenticeship for adolescents: Integrating learning with cognitive strategies. *Reading Research Quarterly*, *52*(1), 31–52. https://doi.org/10.1002/rrq.147

Diakidoy, I. A. N., Christodoulou, S. A., Floros, G., Iordanou, K., & Kargopoulos, P. V. (2015). Forming a belief: The contribution of comprehension to the evaluation and persuasive impact of argumentative text. *British Journal of Educational Psychology*, *85*(3), 300–315. https://doi.org/10.1111/bjep.12074

Diakidoy, I. A. N., Ioannou, M. C., & Christodoulou, S. A. (2017). Reading argumentative texts: Comprehension and evaluation goals and outcomes. *Reading and Writing: An Interdisciplinary Journal*, *30*(9), 1869–1890. https://doi.org/10.1007/s11145-017-9757-x

Eisenberger, R., Masterson, F. A., & McDermitt, M. (1982). Effects of task variety on generalized effort. *Journal of Educational Psychology*, *74*(4), 499–505. https://doi.org/10.1037/0022-0663.74.4.499

Felton, M., Crowell, A., & Liu, T. (2015). Arguing to agree: Mitigating my-side bias through consensus-seeking dialogue. *Written Communication*, *32*(3), 317–331. https://doi.org/10.1177/0741088315590788

Gil, L., Bråten, I., Vidal-Abarca, E., & Strømsø, H. I. (2010). Summary versus argument tasks when working with multiple documents: Which is better for whom? *Contemporary Educational Psychology*, *35*(3), 157–173. https://doi.org/10.1016/j.cedpsych.2009.11.002

Graham, S., Liu, X., Aitken, A., Ng, C., Bartlett, B., Harris, K. R., & Holzapfel, J. (2018). Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis. *Reading Research Quarterly*, *53*(3), 279–304. https://doi.org/10.1002/rrq.194

Harris, K. R., Ray, A., Graham, S., & Houston, J. (2019). Answering the challenge: SRSD instruction for close reading of text to write to persuade with 4th and 5th grade students experiencing writing difficulties. *Reading and Writing*, *32*(6), 1459–1482. https://doi.org/10.1007/s11145-018-9910-1

Hayes, A. F. (2012). *PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]*. http://www.afhayes.com/public/process2012.pdf

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Kobayashi, K. (2015). Learning from conflicting texts: The role of intertextual conflict resolution in between-text integration. *Reading Psychology*, *36*(6), 519–544. https://doi.org/10.1080/02702711.2014.926304

Latini, N., Bråten, I., Anmarkrud, Ø., & Salmerón, L. (2019). Investigating effects of reading medium and reading purpose on behavioral engagement and textual integration in a multiple text context. *Contemporary Educational Psychology*, *59*, 101797. https://doi.org/10.1016/j.cedpsych.2019.101797

List, A., & Alexander, P. A. (2017). Analyzing and integrating models of multiple text comprehension. *Educational Psychologist*, *52*(3), 143–147. https://doi.org/10.1080/00461520.2017.1328309

Lucisano, P., & Piemontese, M. E. (1988). GULPEASE: una formula per la predizione della difficoltà dei testi in lingua italiana (en. tr. GULPEASE: a formula to predict the difficulty of texts in Italian). *Scuola e Città, 3*, 110–124.

Luna, M., Villalón, R., Martínez-Álvarez, I., & Mateos, M. (2022). Online interventions to help college students to improve the degree of integration of their argumentative synthesis. Special Issue on Synthesis Writing. *Reading and Writing*. https://doi.org/10.1007/s11145-021-10248-0

MacArthur, C. A., Jennings, A., & Philippakos, Z. A. (2019). Which linguistic features predict quality of argumentative writing for college basic writers, and how do those features change with instruction? *Reading and Writing*, *32*(6), 1553–1574. https://doi.org/10.1007/s11145-018-9853-6

Macpherson, R., & Stanovich, K. E. (2007). Cognitive ability, thinking dispositions, and instructional set as predictors of critical thinking. *Learning and Individual Differences*, *17*(2), 115–127. https://doi.org/10.1016/j.lindif.2007.05.003

Mason, L., & Scirica, F. (2006). Prediction of students' argumentation skills about controversial topics by epistemological understanding. *Learning and Instruction*, *16*(5), 492–509. https://doi.org/10.1016/j.learninstruc.2006.09.007

Mateos, M., Cuevas, I., Martín, E., Martín, A., Echeita, G., & Luna, M. (2011). Reading to write an argumentation: The role of epistemological, reading and writing beliefs. *Journal of Research in Reading*, *34*(3), 281–297. https://doi.org/10.1111/j.1467-9817.2010.01437.x

Mateos, M., Martín, E., Cuevas, I., Villalón, R., Martínez, I., & González-Lamas, J. (2018). Improving written argumentative synthesis by teaching the integration of conflicting information from multiple sources. *Cognition and Instruction*, *36*(2), 119–138. https://doi.org/10.1080/07370008.2018.1425300

Mateos, M., & Solé, I. (2009). Synthesising information from various texts: A study of procedures and products at different educational levels. *European Journal of Psychology of Education*, *24*(4), 435–451. https://doi.org/10.1007/BF03178760

Mateos, M., Villalón, R., De Dios, M. J., & Martin, E. (2007). Reading-and-writing tasks on different university degree courses: What do the students say they do? *Studies in Higher Education*, *32*(4), 489–510. https://doi.org/10.1080/03075070701476183

Nussbaum, M. E. (2005). The effect of goal instructions and need for cognition on interactive argumentation. *Contemporary Educational Psychology*, *30*(3), 286–313. https://doi.org/10.1016/J.CEDPSYCH.2004.11.002

Nussbaum, M. E. (2008). Using argumentation vee diagrams (AVDs) for promoting argument-counterargument integration in reflective writing. *Journal of Educational Psychology*, *100*(3), 549–565. https://doi.org/10.1037/0022-0663.100.3.549

Nussbaum, M. E., & Edwards, O. V. (2011). Critical questions and argument stratagems: A framework for enhancing and analyzing students' reasoning practices. *Journal of the Learning Sciences*, *20*(3), 443–488. https://doi.org/10.1080/10508406.2011.564567

Perfetti, C. A., Rouet, J.-F., & Britt, M. A. (1999). Toward a theory of documents representation. In H. Van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representation during reading* (pp. 99–122). Erlbaum.

Perkins, D. N., Faraday, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 83–105). Lawrence Erlbaum Associates.

Primor, L., & Katzir, T. (2018). Measuring multiple text integration: A review. *Frontiers in Psychology*, *9*, 2294. https://doi.org/10.3389/fpsyg.2018.02294

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist*, *52*(3), 1–19. https://doi.org/10.1080/00461520.2017.1322968

Rijlaarsdam, G., Janssen, T., Rietdijk, S., & Van Weijen, D. (2017). Reporting design principles for effective instruction of writing: Interventions as constructs. In R. Fidalgo, K. Harris, & M. Braaksma (Eds.), *Design principles for teaching effective writing: Theoretical and empirical grounded principles* (pp. 280–313). Brill Publishers. https://doi.org/10.1163/9789004270480_013

Schumm, W. R., Pratt, K. K., Hartenstein, J. L., Jenkins, B. A., & Johnson, G. A. (2013). Determining statistical significance (alpha) and reporting statistical trends: Controversies, issues, and facts. *Comprehensive Psychology, 2*: 10. https://doi.org/10.2466/03.CP.2.10

Solé, I., Miras, M., Castells, N., Espino, S., & Minguela, M. (2013). Integrating information: An analysis of the processes involved and the products generated in a written synthesis task. *Written Communication*, *30*(1), 63–90. https://doi.org/10.1177/0741088312466532

Stanovich, K. E., & West, R. F. (2008). On the failure of cognitive ability to predict myside and one-sided thinking biases. *Thinking & Reasoning*, *14*(2), 129–167. https://doi.org/10.1080/13546780701679764

Strømsø, H. I., Bråten, I., & Britt, M. A. (2010). Reading multiple texts about climate change: The relationship between memory for sources and text comprehension. *Learning and Instruction*, *20*, 192–204. https://doi.org/10.1016/j.learninstruc.2009.02.001

Tarchi, C., & Mason, L. (2020). Effects of critical thinking on multiple-document comprehension. *European Journal of Psychology of Education*, *35*(2), 289–313. https://doi.org/10.1007/s10212-019-00426-8

Tarchi, C., & Villalón, R. (2021). The influence of thinking dispositions on integration and recall of multiple texts. *British Journal of Educational Psychology*, e12432. https://doi.org/10.1111/bjep.12432
Toulmin, S. E. (1958). *The uses of argument*. Cambridge University Press.
Voss, J. F., & Van Dyke, J. A. (2001). Argumentation in psychology: Background comments. *Discourse Processes*, *32*(2–3), 89–111. https://doi.org/10.1080/0163853X.2001.9651593
World Medical Association. (2013). Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA, 310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

## Appendix A

**Instructions provided in the Recursivity-Induced condition**

You will read two texts presenting two contrasting stances on a controversial topic in the field of education. Your task is to write about your perspective, taking into consideration what you read and your ideas.

The strategy that is most frequently applied when we compare contrasting perspectives is to assume a position from the beginning and try to defend it, finding supporting arguments and refuting the argumentation supporting the opposite position. In other words, we try to persuade the other ones about the validity of one of the positions.

However, this strategy is not the best choice. It is better to adopt a strategy that integrates, that is that allows to explore more in depth the different positions before getting to a conclusion. Given that each position has pros and cons, what you should do is to compare them to find a conclusion that integrates them. In Psychology, we know that our knowledge, in this case the stance offered by each text, is just a perspective, chosen by the author to represent the issue. To learn more, and achieve an increasingly complete level of knowledge, we should understand that there are different perspective and positions on several topics, without one being necessarily truer than the other ones. Thus, we need to deeply comprehend each position, give value to their supporting arguments, and think about new positions that could overcome the limitations of the prior ones. To this aim, we ask you to take into consideration four elements, to be included in a recursive process (rather than as linear steps):

(a) Identifying and exploring different position on a topic (what are the supporting reasons in each text?)
(b) Compare and contrast the positions (Do some of the argumentations supporting one position contrast with the argumentations supporting the opposite position?)
(c) Draw conclusions (are there argumentations that weigh more than other ones? Is there an intermediate solution that would allow to integrate pros and/or overcome cons of the two positions?)
(d) Write the argumentative essay.

We hope that taking into consideration these elements could help you to successfully perform the integration task.

[PARTICIPANTS READ TEXT #1, AT THE BOTTOM OF WHICH THE FOLLOWING QUESTIONS ARE LISTED]

**Please, answer to the following questions**

(1) **What is your position on the topic discussed in the text?**
(2) **What are the main argumentations of the text?**
(3) **Is this text in line with your prior beliefs on the topic? Please, explain.**
(4) **Do the ideas expressed in the text changed in some way your prior beliefs on the topic? Please, explain.**

(5) **What is your opinion on the debated topic now?**

[PARTICIPANTS ANSWER QUESTIONS AND THEN MOVE TO TEXT #2 AND READ IT. AT THE BOTTOM OF TEXT #2 THE FOLLOWING QUESTIONS ARE LISTED]

**Please, answer to the following questions**

(1) **What is your position on the topic discussed in the text?**
(2) **What are the main argumentations of the text?**
(3) **Is this text in line with the other text? Please, explain.**
(4) **Is this text in line with your prior beliefs on the topic? Please, explain.**
(5) **Do the ideas expressed in the text changed in some way your prior beliefs on the topic? Please, explain.**
(6) **What is your opinion on the debated topic now?**
(7) **Did you need to read again the first text?**

**Instructions provided in the Active Control condition**

You will read two texts presenting two contrasting stances on a controversial topic in the field of education. Your task is to write about your perspective, taking into consideration what you read and your ideas.

The strategy that is most frequently applied when we compare contrasting perspectives is to assume a position from the beginning and try to defend it, finding supporting arguments and refuting the argumentation supporting the opposite position. In other words, we try to persuade the other ones about the validity of one of the positions.

[PARTICIPANTS READ TEXT #1, AT THE BOTTOM OF WHICH THE FOLLOWING INSTRUCTION WAS PROVIDED]

**Write a summary of the text, including all the main points.**

[PARTICIPANTS ANSWER TO QUESTIONS AND THEN MOVE TO TEXT #2 AND READ IT. AT THE BOTTOM OF TEXT #2 THE FOLLOWING QUESTIONS ARE LISTED]

**Write a summary of the text, including all the main points.**

## Appendix B

**Prior topic beliefs questionnaire**

Please, refer to the following scale to answer to the questions.

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Completely disagree | Quite disagree | Slightly disagree | Slightly agree | Quite agree | Completely agree |

| | |
|---|---|
| (1) It is necessary to evaluate teachers' quality of teaching. | 1 2 3 4 5 6 |
| (2) We should trust each teacher's initiative and goodwill for the assessment of the quality of teaching. | 1 2 3 4 5 6 |
| (3) It is not positive for all teachers to have access to the same benefits, regardless of the quality of their work. | 1 2 3 4 5 6 |
| (4) It is important to take students' learning outcomes as an indicator of teaching quality. | 1 2 3 4 5 6 |
| (5) It is important to promote the theme of the evaluation of the teaching quality in school policies. | 1 2 3 4 5 6 |
| (6) You cannot attribute students' outcomes to the influence of a single professor. | 1 2 3 4 5 6 |
| (7) There is no enough consensus about what the best teaching practices are in order to create an evaluation system. | 1 2 3 4 5 6 |
| (8) We do not have enough resources to sustain an appropriate evaluation system of the teachers. | 1 2 3 4 5 6 |