Recursivity in source-based writing: a process analysis

Christian Tarchi[1] · Ruth Villalón[2] · Nina Vandermeulen[3] ·
Lidia Casado-Ledesma[1] · Anna Paola Fallaci[1]

Abstract

In university settings, writing argumentative essays from reading conflicting source texts is a common task for students. In performing this synthesis task, they must deal with conflicting claims about a controversial issue as they develop their own positions. Argumentative synthesis is characterized by writers' back-and-forth moves between reading source texts and writing their own texts—a self-regulatory process that can be termed recursivity. The present study investigated the recursive behavior of Italian university students as they wrote argumentative syntheses while reading conflicting sources. The 43 graduate students participating in the study read four source texts on a controversial topic, evaluation in academe, with the goal of writing an argumentative essay. Reading of the sources was studied through a think-aloud procedure, and recursivity in writing the syntheses was recorded through Inputlog software. Comparisons were made between 22 high-recursive and 22 low recursive writers for the quality of their argumentative essays and for the critical strategies that they had used in reading the sources. Descriptive and nonparametic analyses produced the following three findings: (1) The strategies most employed in prereading were all related to synthesis-related activities: voicing opinion, expressing agreement, and expressing doubts. (2) Recursivity occurred most often in the middle of the synthesis process, as writers developed their arguments, instead of at the beginning or end. (3) High-recursive writers surpassed low-recursive writers by producing argumentative essays of higher quality and obtained better recall scores. They also employed more critical processing relevant to synthesis when reading the sources. This study provides insight on how recursivity is involved in argumentative writing but still there is need for further research.

Keywords Recursivity · Think aloud · Argumentative synthesis writing · Process analysis

---

Extended author information available on the last page of the article

Composing texts based on the reading of multiple sources, commonly known as synthesis writing, has piqued the curiosity of the educational and scientific community in recent times. This heightened interest can be attributed to both its frequent requirement across various educational levels (Marttunen & Kiili, 2022) and the cognitive challenges it poses for students. Studies in the field of synthesis writing have highlighted the significant learning opportunities inherent in this type of task. The process of reading, rereading, integrating, organizing, and extending diverse source texts requires a profound transformation of knowledge (Solé et al., 2013; Spivey & King, 1989). However, it is not surprising that for students tackling these activities poses a significant challenge, given the high cognitive demands involved (Mateos et al., 2018; Solé et al., 2013). Students are often faced with multiple-text comprehension tasks with the goal of producing an argumentative essay. When this learning situation occurs, students have to generate a particular type of argumentative writing: an argumentative synthesis (Mateos et al., 2018). Argumentative synthesis is a hybrid task that implies the critical use of reading and writing. Indeed, when synthesizing texts, writers comprehend such sources and write an essay based on the information read, returning to source texts for further comprehension if necessary. Producing a synthesis requires students to read and reread the texts (Nelson, 2008; Vandermeulen et al., 2020c), both to identify relevant information and to elaborate and integrate it into the writing; i.e., synthesis writing is closely related to the recursion process.

Recursivity, which means returning to and repeating a procedure, has become a focus of research in synthesis writing since this is a central cognitive process in this type of activities (Nelson & King, 2023; Solé et al., 2013). The concept of recursion is widely known in the field of writing research. This recognition dates back to Emig's pioneering study in 1971 (Emig, 1971), which demonstrated that the writing process does not follow a strict, linear sequence comprising only the planning, writing, and revision phases. Rather, writers follow a recursive pattern, repeatedly returning to subprocesses such as planning or revision at different points in the composition process (Flower & Hayes, 1981; Perl, 1980). However, for research on writing from sources, the term recursion is used differently.

In the field of writing from sources, recursivity involves an iterative process of "back and forth" between the reading of sources and the writing itself (Vandermeulen et al., 2023). It is a self-regulatory cognitive process which makes it possible to monitor the writer's behavior, in order to introduce the relevant changes in the planning, textualization and evaluation phases (Mateos et al., 2018; Segev-Miller, 2007). Throughout the writing process, authors constantly revisit and reassess their ideas, arguments, and language choices, seeking coherence and effectiveness. This iterative process allows them to identify weaknesses, address inconsistencies, and refine their communication.

Despite the importance of recursivity in critical reading and writing, to date the studies focusing on this behavior are extremely scarce. In this study we aim to contribute to the literature on argumentative synthesis by investigating the relevance of recursivity and its interplay with critical reading processes.

Source-based writing

Writing activities in the academic context can take many forms. Students may be asked to write opinion essays on specific content, scientific reports, summaries of book chapters, etc. One task that stands out for its frequency and the difficulty it entails for students is source-based writing. Source-based writing requires the writer to read different sources and to synthesize information from them in response to an objective; for example, to develop a comprehensive view of a controversial topic (Braine, 1995; Weston-Sementelli et al., 2018). To adequately develop these writing tasks, students not only have to master different writing skills, but they also have to be proficient in reading and comprehending the different sources provided. Composing a high-quality text based on reading sources depends on both reading and writing skills and, therefore, there is an overlap between the processes of comprehension and language production (Spivey, 1990). This interdependence between the reading and writing processes (Graham et al., 2020) requires reading effectively in order to identify relevant information for the composition process and, in relation to the writing process, knowing how to incorporate this material into the text being created (Hirvela, 2004).

Argumentative synthesis writing

Synthesis writing is a type of source-based writing (Vandermeulen et al., 2023) and, therefore, it is a hybrid task (Spivey & King, 1989) that requires the combined use of reading and writing. Regarding reading processes, students need to evaluate the trustworthiness and relevance of the source-texts, identify the main perspective, identify and evaluate the strength of the main arguments (and counter-arguments), monitor their own comprehension and connect the new information with their prior knowledge and experiences. In other words, students need to read strategically. In addition, and because they are reading different sources, students need to perform the same actions across texts, to identify whether they hold compatible or opposing perspectives, and the extent to which they overlap in information provided and arguments discussed. Regarding writing processes, students need to plan, compose and revise (Hayes, 2012). In short, synthesis writing is an epistemic and a complex task (Segev-Miller, 2004) that requires the implementation of processes of selection, organization and connection of information related to different sources (Spivey, 1997), as well as intratextual (within one text) and intertextual integration (between two or more sources) processes in order to write a document with an original structure and content (Segev-Miller, 2007). To do so, a reader should consult the sources while writing his/her own text.

One aspect to take into account is that syntheses can be elaborated from sources that present complementary or conflicting information on a topic. Writing a synthesis from sources that present conflicting information can be understood as a particular type of argumentative writing, since it is necessary to consider the arguments and counterarguments related to the different perspectives (Mateos et al., 2018).

Addressing alternative perspectives on the controversial issue is critical to effective argumentation in argumentative synthesis writing; activities which are becoming increasingly important in the education of elementary and secondary students (e.g., De la Paz & Felton, 2010), as well as college students (e.g., Granado-Peinado et al., 2019; Luna et al., 2023; Mateos et al., 2018). In arguing personal opinion on a particular topic, different strategies can be implemented. A rebuttal strategy may be employed in case the arguments corresponding to the undefended position are considered erroneous or insufficiently justified. Another strategy may be to support one of the perspectives after assessing and weighing the arguments linked to the two positions. writers can point out the strengths and weaknesses of alternative perspectives and also refute positions and assertions with which they disagree (Reznitskaya et al., 2009; Toulmin, 1958). However, the emphasis can also be placed on intertextual integration processes when reading texts that address conflicting topics. In this regard, although rebuttal and weighing are well-recognized strategies in argumentation, Nussbaum and Schraw (2007) added another strategy in their theoretical framework concerning the integration of arguments and counterarguments: compromise/conciliation between alternative views. In this last strategy defined, the writer tries to propose a conciliatory solution that brings together the positive aspects of the two opposing positions. Importantly, even though all strategies described by Nussbaum and Schraw are employed in synthesizing, the authors use the term "synthesis" for one specific strategy: the development of a "conciliatory solution" to the problem being addressed. Moreover, Nussbaum and Schraw use the terms "argument" and "counterargument" for what many writing researchers would call "claim" and "counterclaim," while defining the term "argument" as a full argumentative text.

A rebuttal strategy may be employed in case the arguments corresponding to the undefended position are considered erroneous or insufficiently justified. Another strategy may be to support one of the perspectives after assessing and weighing the arguments linked to the two positions. The third and last strategy defined by these authors would be the strategy of synthesis, in which the writer tries to propose a conciliatory solution that brings together the positive aspects of the two opposing positions. Importantly, even though all strategies described by Nussbaum and Schraw are employed in synthesizing, the authors use the term "synthesis" for one specific strategy: the development of a "conciliatory solution" to the problem being addressed. Moreover, Nussbaum and Schraw use the terms "argument" and "counterargument" for what many writing researchers would call "claim" and "counterclaim," while defining the term "argument" as a full argumentative text. Furthermore, it is worth noting that while Nussbaum and Schraw refer to a synthesis strategy, it could also be called "compromise/conciliation between alternative views". However, in the field of research on argumentative synthesis writing from multiple sources, the term "synthesis" is commonly used to refer to this specific procedure.

Several studies have been conducted in the field of argumentative synthesis writing from sources with conflicting information (e.g., Casado-Ledesma et al., 2021; Granado-Peinado et al., 2023; Luna et al., 2023; Mateos et al., 2018). All these studies share a common feature, which is the design and implementation of intervention programs aimed at enhancing students' competence in writing argumentative syntheses. In doing so, they all draw upon the theoretical framework of Nussbaum and

Schraw regarding strategies for integrating arguments and counterarguments. With regard to our research, an argumentative synthesis writing task was implemented, that is, participants were asked to express an opinion on a topic and support it with the arguments and counter-arguments identified in the texts. Being that, our analytical approach also drew upon Nussbaum and Schraw's proposal regarding intertextual integration strategies. Besides, we adopted two process-tracing approaches: think-aloud procedures (Afflerbach & Cho, 2009) learn about reading strategies employed when writers read source texts after being informed that they will soon write argumentative texts from conflicting sources, and input logs (Leijten & Van Waes, 2013) to learn about recursivity during writing. We also used two product-oriented measures: text evaluation of the argumentative syntheses, with major attention to intertextual integration as in past studies in the field (Casado-Ledesma et al., 2021; Granado-Peinado et al., 2023; Luna et al., 2023; Mateos et al., 2018) and a delayed recall measure for addressing deep comprehension.

Recursivity in source-based writing

Recursivity when writing has received some attention from research. By recursivity we refer to the number of switches between sources and the writer's text document. Writers may go back to sources at different stages of the writing processes, namely when planning, composing or revising. Weak writers tend to follow a linear process, from reading to writing, which in turn produces low-quality texts (Fidalgo et al., 2014). Strong writers go back and forth from sources to their own text several times for, hypothetically, strategic reasons (Mateos & Solé, 2009; Solé et al., 2013).

The relevance of recursivity when writing is grounded in the levels-of-processing theoretical framework (Craik, 2002; Craik & Lockhart, 1972). According to this theory, people process information at different levels of depth, which are generally not processed linearly. Rather, people re-circulate information in their memory to further analyze it. Of course, this process depends on the quality of the working memory: the trace may get lost once people proceed to process different information. The repeated presentation of stimuli could support this process. Thus, recursivity exposes learners over and over again to the same information, which can be processed at different levels.

Past studies have investigated whether recursivity is associated with argumentative synthesis writing. Mateos and Solé (2009) analyzed the written products of students from different educational levels who had received a synthesis task from their teachers. They found that older students (university level) implemented more often a recursive rather than linear approach to the task than younger students. This finding was partially confirmed by Vandermeulen et al.'s study (2020d), showing that higher grade students switched more frequently between sources and their own text, at least in the beginning of the writing process. Moreover, the studies of Solé et al. (2013), with secondary students, and Du and List (2020), with undergraduate students, also support the idea that better quality products are related to more recursive patterns while reading multiple texts. Vandermeulen et al. (2020c) studied source use in upper-secondary students' argumentative and informative source-based writing.

Results showed that recursion was most frequent in the middle part of the writing process (as compared to the beginning and end phase), and that students switched to the sources more frequently when writing an argumentative text than when writing a narrative text. Additionally, these authors related source use to the quality of the text. A positive correlation between recursivity in the first phase of the process and text quality was found, while recursivity in the last phase of the process correlated negatively with text quality.

Process analysis in reading and writing

Most research on reading and writing has almost exclusively focused on the products of these activities (e.g., reading comprehension, recall, written text quality, coherence, and the like). At the same time, several scholars have turned their attention towards reading and writing processes, developing research methodologies able to provide us an insight into the students' metacognitive activity.

The think-aloud methodology has been used to address reading in writing from sources (Du & List, 2020; Mateos et al., 2018; Solé et al., 2013). This methodology helps researchers to identify cognitive and metacognitive processes implemented during a learning task (Ericsson & Simon, 1998; Pressley & Afflerbach, 1995). When performing a task, such as reading one or more texts, participants are asked to "think aloud", that is to voice any thought they have while reading, without filtering any thought. Thinking aloud while performing a task, rather than before (prospective think-aloud) or after (retrospective think-aloud) is considerate preferable as it addresses two limitations of these options, respectively people do not do what they say they do and people do not always recall accurately what they have done (Hu & Gao, 2017). Moreover, it provides direct access to reading processes, whereas other techniques, such as log-data or eye-tracking, indirectly infer metacognitive processes from behavior. Recent studies have demonstrated the substantial neutrality of think-aloud on target processes (Bannert & Mengelkamp, 2008; Tarchi, 2021).

One way to access cognitive and metacognitive processes such as recursivity during writing is through the use of keystroke logging tools such as Inputlog (Leijten & Van Waes, 2013). Inputlog makes it possible to observe the writing process unobtrusively as it runs in the background of a familiar word processor. Inputlog records (or logs) every keystroke, mouse movement, and window change. All the logged writing process activities are time stamped. The log files can be analyzed within Inputlog from different perspectives: fluency, pause, revision, and - of particular interest when studying recursivity - source use (Vandermeulen et al., 2020b). Studying the dynamics of the writing process using Inputlog allows us to understand the complexity of writing as a process; however, the conclusions that can be derived from the records are inferential and establishing a direct link between keystrokes and cognitive/metacognitive activities is often not evident (Galbraith & Baaijen, 2019). It is therefore advisable to complement this method with others that directly capture the cognitive/metacognitive activity of the subject when performing the task (Wengelin et al., 2019).

The present study

Recursivity seems deeply involved in source-based writing tasks, such as argumentative synthesis writing. It may help to connect reading and writing processes and to re-introduce relevant information in the students' working memory as they proceed in the writing task. However, it is still unclear whether recursivity is associated with strategic processes when going back to sources. Moreover, it is unclear to what extent recursivity is associated with argumentative synthesis performance. These aspects led us to propose the current research, through which we aimed to learn more about writers engaging in an argumentative synthesis task: (a) the strategies they employ in reading the source texts, (b) the recursivity that occurs in their writing of argumentative syntheses, and (c) the quality of the argumentative syntheses that they produce, especially intertextual integration. We were also interested in differences between high and low recursive writers in terms of their reading strategies, patterns of recursivity, and quality of their syntheses. In this study, university student writers read and wrote on the controversial topic of evaluation of education; specifically, about the advantages and disadvantages of standardized student assessment and the evaluation of teachers' professional practice. Thus, the objectives of this research were as follows:

1. To describe recursivity behavior (identified through keystroke logging) in university students while reading conflicting sources and while writing argumentative synthesis.
2. To compare high- versus low-recursive writers on the quality of argumentative essays and the recall of the sources.
3. To compare high- versus low-recursive writers on strategic behavior, assessed by a think-aloud protocol.

Based on past evidence, studies suggest that writing performances of students in synthesis tasks are still suboptimal, even at the higher education level, and that recursivity is may not found in the behavior of many subjects with less experienced (e.g., secondary school level, see Vandermeulen et al., 2020c; undergraduate students, Tarchi & Villalón, 2022). However, in our study the participants were postgraduate students and the task demanded the use of a significant number of sources, so we expected a moderately higher level of recursivity. Moreover, we hypothesized that recursivity is associated with higher quality in argumentative synthesis written essays. In particular, recursivity should be associated with a higher level of intertextual integration. Moreover, we hypothesized that recursivity would be associated with cognitive and metacognitive strategies while reading sources. In other words, we expected for high-recursive students to write more integrated essays and to be more strategic when reading then low-recursive students.

A recall measure was also included in the research design to investigate the impact of recursivity on retention and depth of processing. In this way, we could investigate whether recursivity influences the way sources are elaborated, besides the quality of students' written products. Recall allows to assess students' representation and long-term retention of the text content. Valid inferences, rather than literal comprehen-

sion, is a strong index for depth of comprehension, as it represents the links students did between text content and prior knowledge when reading (Diakidoy et al., 2015; Tarchi & Villalón, 2021).

The following variables were also assessed: perceived prior knowledge, prior beliefs, and need for cognition. These three variables have been found connected with argumentative synthesis writing (see Dai & Wang, 2007; Tarchi & Villalón, 2021) and may be associated with recursivity. Students with low perceived prior knowledge may struggle in strategically approaching the task and proceed more linearly. Students with skewed prior beliefs may find it unnecessary to process belief-inconsistent texts. Students with low levels of need for cognition may be not so engaged in a complex task such as argumentative synthesis.

As in much of the multi-text reading research (e.g., List & Alexander, 2020; Schoor et al., 2023), we divided the task into a reading phase and a written production phase. However, since synthesis writing is a hybrid task, we must acknowledge that much composing was, no doubt, occurring as students first encountered the sources during the reading phase of the study.

Method

Participants

Forty-three university students participated voluntarily in the study (13 males, 29 females, one preferred not to declare gender; age mean$=23.9\pm2.04$). All participants were enrolled in a Master's degree program in Educational Psychology. All participants were Italian and spoke Italian as their primary language. Data was collected anonymously (the participants included a personalized code in each task). The study was approved by the Ethics Committee of the University of Florence (Italy).

Different variables related to the participants were assessed; specifically perceived prior knowledge and prior beliefs about the topics addressed in the source texts, as well as need for cognition. Perceived prior knowledge was evaluated through an item ("What is your level of knowledge on the topic of evaluation in school"?) to be rated on a scale from 1 (minimum) to 6 (maximum). Prior beliefs were assessed through an 8-item questionnaire including four items reporting a pro-evaluation stance (*e.g., it is necessarily to evaluate teaching quality*) and four items reporting an against-evaluation stance (*e.g., There is no sufficiently well-founded consensus on what constitutes good teaching practices to create an evaluation system*). The four against-evaluation items were reverse coded. The composite score was obtained by adding up all the ratings: the higher the score, the more pro-evaluation the beliefs were. The reliability of the scale was adequate ($\alpha=0.71$). Need for cognition (Cacioppo & Petty, 1982) was assessed using an 18-item questionnaire (*e.g., I like tasks that require little reflection once they have been learnt*). Participants scored each item on a 5-point Likert scale (from 1=completely false to 5=completely true). The reliability of the scale was adequate ($\alpha=0.87$).

Materials

Source texts

We used four texts previously employed in studies about argumentative synthesis writing (e.g., Granado-Peinado et al., 2019; Mateos et al., 2018). The texts discussed the topic of how to enhance the quality of teaching and learning in the school system.

Two texts addressed the topic of teachers' evaluation; namely, the advantages and disadvantages of conducting an evaluation of teachers' professional practice, in order to improve the quality of instructional processes (one of the texts addressed the advantages, and the other, the disadvantages). The text in favor of teachers' evaluation received the name of "Improving the quality of teaching" (599 words) and presented arguments supporting the use of teachers' evaluation to improve teaching quality. The text against teachers' evaluation was titled "Good intentions, bad outcomes" (594 words), and included the problems regarding the implementation of instructors' evaluation.

The other two texts dealt with the topic of student assessment, through standardized and external performance tests, one taking a positive position and the other taking a negative position. The text related to the advantages of students' evaluation received the title of "Students 'assessment and education quality" (502 words) and included arguments supporting the use of students' performance evaluation as a way to improve the quality of educational processes at school. The text related to the disadvantages of students' evaluation was named "The performance evaluation trap" (612 words), and it included arguments related to the difficulty of deriving improvements in education from these standardized and external evaluations.

The original texts were written in Spanish, adapted by the second author based on texts used in previous studies (Authors, XXXX), so prior to the implementation of the study they were translated into Italian. Cultural adaptability to the Italian educational context was ensured by the first author. Texts had similar readability scores (calculated through the Gulpease, a legibility index for Italian, range 0-100): "Assessment and quality of teaching" (Gulpease index=45), "The performance evaluation trap" (Gulpease index=47), "Improving the quality of teaching" (Gulpease index=43), "Good intentions, bad outcomes" (Gulpease index=48). Overall, texts were balanced by length, difficulty and number of supporting arguments (seven each text). Excerpts from texts are included in the Supplementary Material A.

Procedure

To aim our objectives, the following procedure was followed. Firstly, the participants were asked to fill in a questionnaire including an assessment of individual variables and demographic information. Secondly, participants were asked to perform a source-based writing task. They were asked to read four texts on a controversial topic. While reading, the participants were asked to think-aloud. Then, participants were asked to write on a personal computer an argumentative essay based on the sources that they had just read. They were asked to write the essay (with access to

sources) while keystroke logging software Inputlog was working in the background. Finally, a week later, they completed a free recall task.

The reading-writing task was conducted online with the direct supervision of an experienced researcher. Prior to the experimental session, students were: (1) instructed how to think-aloud, (2) asked to practice thinking-aloud with two texts provided by the researcher, (3) asked to send a sample of the think-aloud to the researcher. Finally, they received feedback on their think-aloud practice. Then, students were: (1) instructed how to install Inputlog on their device, (2) asked to practice starting and ending the writing sessions with Inputlog, (3) and asked to send a sample of the output to the researcher. Finally, they received feedback on their think-aloud practice. Think-aloud and Inputlog practice sessions were all well performed by the participants on their first attempt. In the experimental session, students were asked to work in a quiet environment and perform the task without interruptions and in the same session. The researcher was available for an online meeting throughout their session for any issue. First, students received the four texts and were asked to read while thinking-aloud. Participants recorded their think-alouds and sent them to the researcher. Immediately after the task, students activated Inputlog and performed the writing task. As soon as they had finished, they were asked to submit the Inputlog output to the researcher. The exchange of materials between students and the researcher was performed through a learning management system. All participants completed the task with no issues. Think-alouds and Inputlogs files were carefully reviewed by the researchers to identify any invalid performance.

Reading task

Students were given four digital texts on the debated topic (see paragraph on texts within the material section for details). They were given the following instructions: "You will now read four texts that argue positions on a controversial topic in education. You can read them as many times as you like and return to them as many times as you like. When you have finished reading the passages, move on to writing. You will be asked to write an essay that discusses the positions expressed in each text and includes a conclusion that integrates the strengths of the positions expressed." This instruction was given so that participants knew that they had to read texts with the purpose of writing an argumentative synthesis essay.

While the participants were reading, they were asked to think-aloud, that is: "say out loud everything that is on your mind, whether inherent in the text you read or not. You should verbalize as much as you can, in any case at least every two minutes (a timer will help you keep time)." Before the reading task, participants practiced think-aloud with a practice text and received feedback from the researcher. The whole reading task was recorded through a screencast software to capture both the reading activity and the thoughts voiced aloud.

Writing task

The participants were given the following instructions: "After reading the texts, you will have to write an essay that, based on the texts you have read, discusses the posi-

tions expressed in each text and includes a conclusion that integrates the strengths of the positions expressed. This is a time and effort-consuming task, as it involves consulting the texts, extracting and connecting the key ideas from the four texts, and writing an essay that draws your own conclusion and explains in a well-argued manner why you came to that conclusion. You can go back and read the texts as many times as you like. There is no time limit for this exercise, but it is very important to perform the reading and writing task in one work session, without interruptions." This instruction was given to help students understand what an argumentative synthesis task is. This type of task is uncommon in the Italian educational system, and students needed some explanation of what it was expected from them.

While performing the writing task, Inputlog was running in the background and logging the writing process. Students were instructed not to take notes on paper. In this way, Inputlog could register every instant the students switched between their own text document and the digital sources, in this way, students' recursive behavior was logged.

Free recall

After one week, the participants were asked to recall as much content as they could from the texts that they had read (without accessing them). This measure provides an indication of long-term comprehension of the texts.

Measures

Strategic reading from think-aloud protocols

Strategic reading was assessed through a think-aloud protocol, which was transcribed and coded following a category system elaborated following a deductive-inductive process. First, we analyzed the scientific literature, identified the studies that investigated strategic reading through think-aloud and created a list of reading strategies (e.g., Bereiter & Bird, 1985; Bråten & Strømsø, 2003). Then we examined 10% of the protocols to identify reading strategies that were not included in the list. This was the final list of reading strategies: Summarizing, Linking to prior knowledge, Digressing from topic, Expressing agreement with text, Linking to prior experiences, Identifying new information, Making proposals, Expressing disagreement with text, Voicing opinion, Identifying new perspectives, Expressing doubts, Assessing source, Comparing texts.

The protocols were coded through Qcamap (Fenzl & Mayring, 2017) by two independent coders, with a good inter-rater agreement ($k=0.85$). Then, we proceeded to calculate a composite score by adding the frequencies of all the functional reading strategies implemented (prior knowledge+agreement with text+prior experiences+new information+proposals+disagreement with text+personal opinion+perspective on topic+doubts+source relevance texts comparison). Verbosity was also assessed (total number of words expressed).

Recursivity in writing from Inputlog

Recursivity was assessed through Inputlog while students were writing, capturing the degree of recursivity between the essay and the sources, among several other indices of the writing process. We counted the number of transitions between the essay and the source texts, which were available when students were writing (absolute recursivity). The total number of transitions was then divided by the total time on task, resulting in a recursivity indicator: the total number of transitions between the sources and the essay per minute. Since the time participants spent on the task differed, it is also recommended to work with relative measures, so that recursivity can be compared between participants (relative recursivity).

Quality of syntheses from text analyses

Students were asked to write an argumentative essay on the topic discussed in the texts. The quality of the essays was assessed considering three different dimensions:

1) The level of *argument-counterargument integration.* As mentioned in the introduction, in this study we have adopted an analytical approach consistent with the proposal of Nussbaum and Schraw (2007), based on the intertextual integration of arguments and counterarguments (elements defined from other theoretical perspectives as claim and counter claim or position and counter position). Regarding this criterion, we employed the following coding tool developed by Mateos et al. (2018); authors who also rely on the framework of integrating arguments and counterarguments. See Table 1 (see supplementary materials B for an extended version):

As seen in the coding system, refutation strategies are considered to be of lower level than weighing and synthesis strategies. This is due to the association of refutation with processes still linked to the bias of one-sided reasoning (Mateos et al., 2018; Nussbaum, 2008).

2) *Intertextual theme*: whether students are able to identify the storyline connecting the texts to each other and whether they explicitly state it in their essays. We assigned the following scores: 0 (students do not identify the common theme); 1 (students only mention the common sub-topic of two texts); 2 (students identify the two sub-topics discussed in the four texts and explicitly state it in the essay).

3) *Supraintegration*: if the students are able to propose solutions that respond to the controversies addressed in four texts, i.e., not only based on one of the sub-topics.

Table 1 Summary of the coding system of argumentative synthesis essays

Score	Argumentative strategy
0	Personal opinion not based on source texts
1	Does not take a clear position.
2	Argues in support of one perspective only
3	Integration via refutation. Essay with or without a conclusion.
4	Minimum integration via weighing or synthesizing. No or partial conclusion.
5	Average integration via weighing or synthesizing. Partially integrated conclusion.
6	Maximum integration via weighing or synthesizing. Conclusion with overall integration.

We assigned the following scores: 0 (the student focuses on one of the two sub-topics - either external evaluation tests or teacher evaluation - without proposing solutions that address both aspects); 1 (the student is able to mention arguments linked to the two issues, but not to propose solutions for both aspects); 2 (minimal supraintegration: the student proposes at most two solutions to give a combined answer to the problems of the two sub-topics); 3 (maximum supraintegration: the student proposes more than two solutions to give a combined answer to the problems of the two sub-topics).

Two independent judges (authors 2 and 4 of the paper) coded 38% of the argumentative essays to calculate the inter-rater reliability. Reliability indexes were appropriate for the three dimensions (ICC Integration: 0.85; ICC Intertextual theme: 0.81; ICC Supraintegration: 0.67). The cases in which there was no agreement were resolved by consensus, and the remaining 62% of the essays were evaluated by one of these researchers using the established criteria. Essay length was also assessed.

Delay recalls

A week after reading the texts, students were asked to recall what they had read. The outcome variable was the number of valid inferential clauses, as a measure of depth of comprehension. Valid inferences are logical connection across content discussed in different parts of a text (local inferences) or in different texts (intertextual inferences). Moreover, we also considered valid inferences logical connection between new information from the texts and students' prior knowledge (global inferences) (Diakidoy et al., 2015; Tarchi & Villalón, 2021). Two raters coded independently the protocols, with a good inter-rater agreement ($k = 0.90$).

Data analysis

Research objectives were investigated through descriptive statistics and non-parametric statistical analyses, given the low sample-size and the non-normal distribution of data. To address the first objective (description of recursivity behavior), we analyzed the descriptive statistics and calculated through a series of non-parametric comparisons for paired samples (Wilcoxon test) to determine in which interval (relative) recursivity was higher. Rank biserial correlations were used as a measure of effect size.

To address the second objective (comparison between high- versus low-recursive writers in argumentative quality), we analyzed the interaction between recursivity and outcome variables through a series of non-parametric comparisons for independent samples (Mann-Whintey test), with rank biserial correlations as a measure of effect size. To this end, high- (n = 22) versus low-recursive writers (n = 21) were identified through a median split of the relative recursivity score. While this approach is less than ideal from a statistical perspective, it helps to provide some initial data on reading and writing processes. Preliminarily, we investigated if there were pre-existing difference between groups in prior knowledge, beliefs, or need for cognition.

To address the third objective (comparison between high- versus low-recursive writers in strategic reading), we conducted a series of Mann–Whitney U tests on each

reading strategy, with rank biserial correlation as a measure of effect size. The same two groups of high- and low-recursive participants were used in this analysis.

Results

Descriptive statistics for individual variables related to the participants (i.e., perceived prior knowledge, prior beliefs, need for cognition and time on task), process variables (recursivity, strategic reading) and outcome variables (from the essay and free recall tasks) are reported in Supplementary Materials C. Descriptive analyses revealed that the strategies most employed in prereading were all related to synthesis activities; specifically voicing opinion, expressing agreement, and expressing doubts.

Description of recursive behavior

Overall, students spent 82.5 min completing the task (with a median of 76.50). In terms of absolute recursivity values, students went back and forth between the text they were writing and the sources they were reading 55.05 times (with a median of 40). In terms of relative recursivity values, students switched on average 0.58 times per minute (with a median of 0.51). To address our description objective, relative recursivity was used as an independent variable. Students' performance measured with Inputlog was split into three time intervals: beginning, middle and end This was done by dividing each writer's total time on task into three equal parts. Because of the complexity of the research design, it was only possible to collect data on a small number of subjects. Due to the sample size of the study and the non-normal distribution of some of the variables, nonparametric tests were performed.

According to Wilcoxon's test, recursivity in the middle (Median=0.64) was higher than recursivity in the beginning (Median=0.43) and in the end (Median=0.37), see Table 2.

The following two cases (see Fig. 1) serve as an example to illustrate the recursive behavioral pattern over the three phases (i.,e., time interval) of the writing process as measured with keystroke logging. As there is quite some variance in recursivity among the students, we present a case of a high-recursive writer (Fig. 1, case on the left side) and a case of a low-recursive writer (Fig. 1, case on the right side). Recursivity is visually represented at the bottom of these graphs by the orange line. When the orange line runs at the top, the focus is on the sources. Every red dot represents a source text. When the orange line runs at the bottom, the focus was on the student's synthesis text. The blue and green lines show the text production (y-axis: number of characters) at a certain point in time (x-axis). The blue line shows the production during the process, while the green line represents the production in the document. We

Table 2 Within-subject comparisons of relative recursivity across time intervals

Comparisons	Wilcoxon's W	p	Rank biserial correlation
Beginning vs. Middle	264	<0.001	−0.55
Beginning vs. End	433	0.17	−0.23
Middle vs. End	857	0.015	0.40
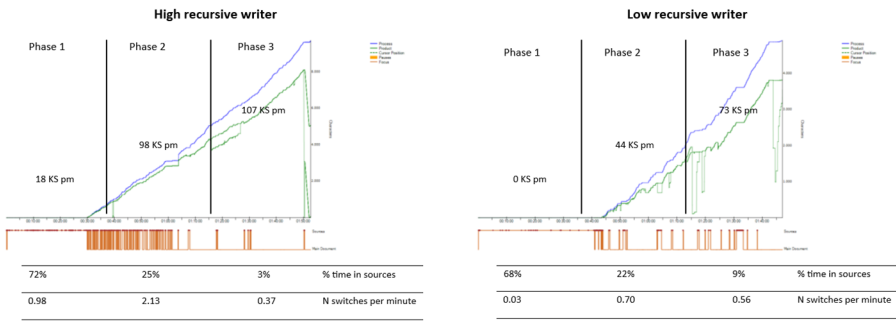
Fig. 1 Illustrative cases: Process graphs generated by Inputlog of the writing process of a high-recursive and a low-recursive writer

refer to Vandermeulen et al. (2020b) for a more complete description of the process graph.

As can be observed in the process graphs, both the high- and the low-recursive writer start the process with a focus on the sources. The second phase of the writing process is marked by text production and a certain degree of recursivity. Also in the third and final phase, text production is dominant. These patterns are in line with findings from previous studies on writing processes of source-based tasks. Synthesis writing processes are generally marked by an initial reading phase (Chau et al., 2022; Vandermeulen et al., 2020d) followed by text production in the middle part of the process. Additionally, recursivity is important for the integration of information or arguments (Vandermeulen et al., 2020c).

In the beginning of the writing process, both the students read the sources without going to their own text document (the orange line runs at the top), so (almost) no text production is taking place. The second process phase is marked by text production. After reading the sources, the students start writing their own text. Both production lines are increasing. An analysis of the keystroke logging data of these two cases shows that the high-recursive writer produces 98 characters per minute in the middle part of the process, thus text is produced rather fluently. At the same time, this student displays a rather high recursivity in the middle phase; this is reflected in the switches between the synthesis text and the sources (2.13 switches per minute). The time spent in the sources is considerably lower than in the first process phase (25% in the second part versus 72% in the first part) as it concerns quick switches between the text document and the sources. Based on these observations, we can argue that it is plausible that the high-recursive writer regularly goes back to the sources to look for information to incorporate in their text. It can be assumed that it is a goal-oriented activity as the checking of the sources is combined with fluent text production.

Although the low-recursive writer switches considerably less frequently between the synthesis text and the sources than the high-recursive writer, recursivity is the highest in the middle part of the process (0.70 switches in phase 2). Although the writer starts producing text in the middle phase of the process, text production is not fluent as this writer types 44 characters per minute. This is not surprising given that it is rather hard to produce text fluently when one relies on their memory to retrieve information from the sources that were read in the first phase of the process.

Table 3 Significant results from Mann-Whitney to compare high- versus low-recursive writers in argumentative quality in strategic reading categories

Variables	U	p	Rank biserial correlation
Total score	137	0.02	0.37
Voicing opinion	127	0.01	0.41
Expressing doubts	157	0.05	0.27
Comparing texts	152	0.04	0.30

Table 4 Significant results from Mann-Whitney to compare high- versus low-recursive writers in outcome variables

Variables	U	p	Rank biserial correlation
Intertextual theme	175	0.02	0.34
Supraintegration	175	0.02	0.33
Recall	101	0.004	0.50

Differences between high-recursive and low-recursive writers

Differences in strategic reading (process variables)

For this analysis, we referred to absolute recursivity as relative recursivity was not associated with strategic reading. Overall, high-recursive students had more strategic reading than low-recursive students did (U=137, p<.05). As a post-hoc analysis, we repeated the Mann-Whitney test on each category. It must be noticed however, that since we are implementing a multiple testing procedure, results should be interpreted with caution. High-recursive writers voiced more their opinions about text content, expressed more doubts and compared the texts more frequently (see Table 3).

Differences in argumentative synthesis writing and delayed recall

To address the second objective, we analyzed the interaction between recursivity and outcome variables through a series of non-parametric comparisons (with rank biserial correlations as a measure of effect size). We also identified high- (n=22) versus low-recursive writers (n=21) through a median split of the relative recursivity score. While this approach is less than ideal from a statistical perspective, it helps to provide some initial data on reading and writing processes. Students with different recursivity levels (high- versus low-recursive students) did not differ in any individual variables, namely perceived prior knowledge (U=156, p>.05), prior beliefs (U=158, p>.05) or need for cognition (U=112, p>.05).

According to the results from the Mann–Whitney U test (employed because of the non-normal distribution of the data), intertextual activity and recall of valid inferences differed across recursivity levels. In both cases, high-recursive writers outperformed low-recursive writers. To better understand at what step in the intertextual integration process recursivity may have an impact, we repeated the Mann–Whitney U test on each level of intertextual integration (see Table 4). High-recursive writers outperformed low-recursive writers in intertextual theme identification and supraintegration, but not in intertextual integration.

Discussion

Source-based writing and argumentative reasoning are two fundamental skills in today's world. We are exposed to complex and controversial topics such as climate change, geopolitical conflicts, pandemics, which require the ability to develop an informed opinion which takes into consideration multiple perspectives and supporting arguments. For these reasons, students should be engaged in argumentative synthesis writing, a type of task in which learners are asked to synthesize multiple perspectives based on sources. Unfortunately, research has demonstrated that students' competence in writing argumentative synthesis essays are suboptimal, even in higher education (Hyytinen et al., 2021; Marttunen & Kiili, 2022; Nelson & King, 2023; Tarchi & Villalón, 2021). To contribute to the scaffolding of students' competences in argumentative synthesis writing tasks, we focused our attention on recursivity, that is, going back and forth between the text we are writing and the sources we are reading (Du & List, 2020; Mateos & Solé, 2009; Tarchi & Villalón, 2022), to provide evidence of the writing process by keystroke logging. Moreover, it is still unclear to what extent recursivity is a strategic process. The present study aimed at addressing these two issues and also to provide more information on the recursivity variable itself.

In the present study, participants displayed an overall minimal level of integration across texts in their essays. Most of the essays were rated as "Minimum integration via weighing or synthesizing with no or partial conclusion." (Mode=4). Regarding our first objective, describing the participants' recursivity behavior, if we look at absolute scores, the level of recursivity among university students involved in an argumentative synthesis writing task seems reasonably high (half of the participants with at least 40 switches between written text and sources), although with a high dispersion of data points, illustrating a consistent variance of recursivity within our sample. Although the absolute number of switches seems high, when we take into account how long they worked on the task, we notice that participants did not switch that often. In respect to the relative scores, our results are coherent with past studies that have indicated that recursivity is most frequently carried out in the middle part of the writing process (Vandermeulen et al., 2020c). Moreover, overall, the relative level (number of switches per minute) was relatively low, compared to performances reported in previous studies. For instance, inspection of data gathered as part of national baseline study in the Netherlands (Vandermeulen et al., 2020a) shows that Dutch students in their last year of upper-secondary school, switched on average 3.02 times per minute between the sources and their text when writing an argumentative text based on conflicting sources. Conversely, in our study we found an average of 0.58 of switches per minute. There are several reasons that may explain this result. Firstly, in the previously referenced national baseline study (Vandermeulen et al., 2020b), students wrote for a maximum of 45 min, whereas in our study the task was open and students took an average of 82.5 min. This could depend on a higher complexity of the task (depending on the topic or the texts) or a higher engagement. Secondly, university students may have a more strategic approach or a higher expertise when reading sources, thus requiring to switch from sources to text less frequently. On the contrary, our sample was quite homogeneous for other control variables. This

might be also the reason because we found no effect of the control variables we explored.

The hypothesis we had for the second objective was substantially supported by our data analysis and coherent with previous studies (Du & List, 2020; Solé et al., 2013). High-recursive students had a better performance in identifying the complexity of the issue explored (intertextual theme identification and supraintegration). However, intertextual integration performances in argumentative essays did not different across recursivity levels. This last result contradicts our research hypothesis, and it may depend that on the fact that the intertextual integration we used (Mateos et al., 2018) was originally designed and employed for intervention studies in which students were being taught the three strategies described by Nussbaum and Schraw (2007) and were expected to use them. Participants in those studies also had less complex pro-con tasks, with only a single major issue and only one pro-text and one con-text.

Moreover, the recall of valid inferences was also associated with a higher recursivity, indicating that a more effortful and nonlinear processing of the sources during writing fosters reading comprehension. These results are, to the best of our knowledge, the first direct evidence supporting the relevance of recursivity for intertextual integration and depth of comprehension in source-based writing. However, recursivity is not frequently found in the common behavior of secondary or even undergraduate students (Fidalgo et al., 2014; Mateos et al., 2018: Solé et al., 2013) For that reason, it is essential that they receive instruction that includes this element, although it seems it is not easily incorporated. Tarchi and Villalón (2022) tested whether it is possible to scaffold university students' recursivity through critical questions. The intervention was effective in improving text quality and induced, at least in some participants, a higher recursivity level as compared to the control group.

In this line, the hypothesis we had for the third objective was also supported by our data analysis. Recursivity was associated with strategic processing during reading, as assessed through the think-aloud methodology. This is in line with previous research (Du & List, 2020; Solé et al., 2013), pointing out that recursivity is linked to self-regulated writers. Past research on thinking aloud when reading multiple texts has emphasized the importance of organization and comprehension confirmation strategies in high-grade students, whereas most of the sample engaged in more shallow processing of texts and implemented memorization and elaboration strategies (Bråten & Strømsø, 2003). In this study, expressing opinions and doubts, and comparing the texts were associated with recursivity, suggesting that students may have looked back at the sources while writing their own text to integrate content across texts or text information with prior beliefs.

Limitations and directions for future research

When interpreting the findings of the current study, some limitations should be taken into account. Firstly, the sample size was quite low, although larger than in previous studies with similar methodologies (Du & List, 2020; Solé et al., 2013). For that reason, it was not allowed to run more complex analysis. Nevertheless, the sample size was appropriate for the statistical analysis performed in this study. As we pro-

vided evidence supporting the relevance of recursivity, future research should further investigate it.

Secondly, recursivity was associated with strategic reading but not with strategies implemented while writing. This was done as think-aloud is a methodology validated for reading but not for writing. The use of retrospective think-aloud protocols may address this issue (although participants do not always recall correctly what they were thinking). Moreover, we used Inputlog only when writing and not when reading not to overload participants, but in future the reading and writing activities should be studied more in terms of a flow of interweaved processes and activate Inputlog and or think-aloud from when they start reading to when they finish writing.

Thirdly, working memory, along with several other individual differences, may have influenced learners' performances (e.g., the free recall measure or the actual need for recursivity). Given that the present research design does not allow us to assess working memory, future studies should investigate the influence of working memory on recursivity.

Conclusions

Recursivity is a behavior that can be tracked with softwares such as Inputlog. Thus, it represents a good candidate for being a learning analytics associated with quality of writing. As the reliance on online platforms to support learning processes is increasing, there is a high demand for automated assessments of writing products and processes (Strobl et al., 2019). Recursivity may be tracked to provide feedback to students as they progress in their writing. For instance, students displaying a low level of recursivity may receive a warning to go back to sources while writing, to support either planning, composing, or revising.

Importantly, the qualitative analysis of two writers suggests that high- versus low-recursive writers seem to address the task with different approaches. Good writers refer more often to sources at the beginning of the process, whereas in both cases they go back to sources in the middle part of the process. Our study suggests that more research is needed to investigate what good writers do in the initial stages of writing.

In the current society, citizens need to deal with information from different sources on a controversial topic and they should be able to express their own view in writing. Given that recursivity is a central element when composing a source-based text, students need evidence-based instruction which marks the role of it (Castells et al., 2022; van Ockenburg et al., 2019). In order to develop such instruction, it is of utmost importance to gain a better understanding of recursion processes. Past studies have shown that instruction may improve recursivity (Tarchi & Villalón, 2022). However, insights obtained from this study could provide valuable input to develop interventions aimed at supporting students' source-based writing and, more in particular, the recursive process. More research on how recursivity is developed and promoted should be carried out, but this study is a first step.

Declarations

Competing interests Authors have no competing interest to declare.

References

Afflerbach, P., & Cho, B. Y. (2009). Determining and describing reading strategies: Internet and traditional forms of reading. In S. H. Waters, & W. Shneider (Eds.), *Metacognition, Strategy Use, and instruction* (pp. 201–225). Guilford Press.

Bannert, M., & Mengelkamp, C. (2008). Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted. *Does the Verbalisation Method Affect Learning? Metacognition and Learning*, *3*, 39–58. https://doi.org/10.1007/s11409-007-9009-6

Bereiter, C., & Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. *Cognition and Instruction*, *2*, 131–156. https://doi.org/10.1207/s1532690xci0202_2

Braine, G. (1995). Writing in the natural sciences and engineering. In D. Belcher, & G. S. Braine (Eds.), *Academic writing in a second language: Essays on research and pedagogy* (pp. 113–134). Albex.

Bråten, I., & Strømsø, H. I. (2003). A longitudinal think-aloud study of spontaneous strategic processing during the reading of multiple expository texts. *Reading and Writing*, *16*, 195–218. https://doi.org/10.1023/A:1022895207490

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

Casado-Ledesma, L., Cuevas, I., Van den Bergh, H., Rijlaarsdam, G., Mateos, M., Granado-Peinado, M., & Martín, E. (2021). Teaching argumentative synthesis writing through deliberative dialogues: Instructional practices in secondary education. *Instructional Science*, *49*, 515–559. https://doi.org/10.1007/s11251-021-09548-3

Castells, N., Minguela, M., & Nadal, E. (2022). Writing a synthesis versus reading: Strategies involved and impact on comprehension. *Reading and Writing*, *36*, 849–880. https://doi.org/10.1007/s11145-022-10341-y

Chau, L., Leijten, M., Bernolet, S., & Vangehuchten, L. (2022). Envisioning multilingualism in source-based writing in L1, L2, and L3: The relation between source use and text quality. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2022.914125

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

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

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*, 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*, 174–192. https://doi.org/10.1016/j.cedpsych.2010.03.001

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*, 300–315. https://doi.org/10.1111/bjep.12074

Du, H., & List, A. (2020). Researching and writing based on multiple texts. *Learning and Instruction*, *66*, 101297. https://doi.org/10.1016/j.learninstruc.2019.101297

Emig, J. (1971). *The composing processes of twelfth graders*. National Council of Teachers of English.

Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind Culture and Activity*, *5*, 178–186. https://doi.org/10.1207/s15327884mca0503_3

Fenzl, T., & Mayring, P. (2017). QCAmap: Eine interaktive webapplikation für qualitative inhaltsanalyse. *Zeitschrift für Soziologie der Erziehung und Sozialisation ZSE*, *37*, 333–340.

Fidalgo, R., Torrance, M., Arias-Gundín, O., & Martínez-Cocó, B. (2014). Comparison of reading-writing patterns and performance of students with and without reading difficulties. *Psicothema*, *26*, 442–448. https://doi.org/10.7334/psicothema2014.23

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*, 365–387.

Galbraith, D., & Baaijen, V. M. (2019). Aligning keystrokes with cognitive processes in writing. In E. Lingren & K. Sullivan (Eds.), Observing Writing. Insights from keystroke logging and handwriting (pp. 306–325). Brill. https://doi.org/10.1163/9789004392526_015

Graham, S., Graham, S., Fulton, M., & Lou (2020). Reading and writing connections: A commentary. In R. A. Alves, T. Limpo, & R. M. Joshi (Eds.), *Reading-writing connections. Literacy studies* (19 vol., pp. 313–317). Springer. https://doi.org/10.1007/978-3-030-38811-9_19

Granado-Peinado, M., Cuevas, I., Olmos, R., Martín, E., Casado-Ledesma, L., & Mateos, M. (2023). Collaborative writing of argumentative syntheses by low-performing undergraduate writers: Explicit instruction and practice. *Reading and Writing*, *36*, 909–936. https://doi.org/10.1007/s11145-022-10318-x

Granado-Peinado, M., Mateos, M., Martín, E., & Cuevas, I. (2019). Teaching to write collaborative argumentative syntheses in higher education. *Reading and Writing*, *32*, 2037–2058. https://doi.org/10.1007/s11145-019-09939-6

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, *29*, 369–388. https://doi.org/10.1177/0741088312451260

Hirvela, A. (2004). *Connecting Reading & writing in Second Language writing instruction*. The University of Michigan Press.

Hu, J., & Gao, X. (2017). Using think-aloud protocol in self-regulated reading research. *Educational Research Review*, *22*, 181–193. https://doi.org/10.1016/j.edurev.2017.09.004

Hyytinen, H., Siven, M., Salminen, O., & Katajavuori, N. (2021). Argumentation and Processing Knowledge in Open-Ended assignment tasks: Challenges and Accomplishments among Pharmacy Students. *Journal of University Teaching and Learning Practice*, *18*, 37–51.

Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing Research: Using Inputlog to analyze writing processes. *Written Communication*, *30*, 358–392. https://doi.org/10.1177/0741088313491692

List, A., & Alexander, P. A. (2020). Strategy use in learning from multiple texts: An investigation of the integrative framework of learning from multiple texts. *Frontiers in Education*, *5*, 578062. https://doi.org/10.3389/feduc.2020.578062

Luna, M., Villalón, R., Martínez-Álvarez, I., & Mateos, M. (2023). Online interventions to help college students to improve the degree of integration of their argumentative synthesis. *Reading and Writing*, *36*, 937–963. https://doi.org/10.1007/s11145-021-10248-0

Marttunen, M., & Kiili, C. (2022). Supporting university students' argumentative source-based writing. *Written Language & Literacy*, *25*, 228–252. https://doi.org/10.1075/wll.00068.mar

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*, 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*, 435–451. https://doi.org/10.1007/BF03178760

Nelson, N. (2008). The reading–writing nexus in discourse research. In C. BaZerman (Ed.), *Handbook of research on writing: History, society, school, individual, text* (pp. 435–450). Lawrence Erlbaum.

Nelson, N., & King, J. R. (2023). Discourse synthesis: Textual transformations in writing from sources. *Reading and Writing*, *36*, 769–808. https://doi.org/10.1007/s11145-021-10243-5

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

Nussbaum, M. E., & Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *Journal of Experimental Education*, *76*, 59–92. https://doi.org/10.3200/JEXE.76.1.59-92

Perl, S. (1980). Understanding composing. *College Composition and Communication*, *31*, 363–369. https://doi.org/10.2307/356586

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Erlbaum.

Reznitskaya, A., Kuo, L., Clark, A., Miller, B., Jadallah, M., Anderson, R. C., & Nguyen-Jahiel, K. (2009). Collaborative reasoning: A dialogic approach to group discussions. *Cambridge Journal of Education*, *39*, 29–48. https://doi.org/10.1080/03057640802701952

Schoor, C., Rouet, J. F., & Britt, M. A. (2023). Effects of context and discrepancy when reading multiple documents. *Reading and Writing*, *36*, 1111–1143. https://doi.org/10.1007/s11145-022-10321-2

Segev-Miller, R. (2004). Writing from sources: The effect of explicit instruction on college students' processes and products. *L1-Educational Studies in Language and Literature*, *4*, 5–33. https://doi.org/10.1023/B:ESLL.0000033847.00732.af

Segev-Miller, R. (2007). Cognitive processes in discourse synthesis: The case of intertextual processing strategies. In M. Torrance, L. van Waes, & D. Galbraith (Eds.), *Writing and cognition: Research and applications* (pp. 231–250). https://doi.org/10.1108/S1572-6304(2007)0000020016

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*, 63–90. https://doi.org/10.1177/0741088312466532

Spivey, N. N. (1990). Transforming texts: Constructive processes in reading and writing. *Written Communication*, *7*, 256–287. https://doi.org/10.1177/074108839000700

Spivey, N. N. (1997). *The constructivist metaphor: Reading, writing, and the making of meaning*. Academic Press. https://doi.org/10.2307/358470

Spivey, N. N., & King, J. R. (1989). Readers as writers composing from sources. *Reading Research Quarterly*, *24*, 7–26.

Strobl, C., Ailhaud, A., Benetos, K., Devitt, A., Kruse, O., & Rapp, C. (2019). Digital support for academic writing: A review of technologies and pedagogies. *Computers & Education*, *131*, 33–48. https://doi.org/10.1016/j.compedu.2018.12.005

Tarchi, C. (2021). Effects of think-aloud on students' multiple-documents comprehension. *Applied Cognitive Psychology*, *35*, 526–537. https://doi.org/10.1002/acp.3782

Tarchi, C., & Villalón, R. (2021). The influence of thinking dispositions on integration and recall of multiple texts. *British Journal of Educational Psychology*, *91*, 1498–1516. https://doi.org/10.1111/bjep.12432

Tarchi, C., & Villalón, R. (2022). Fostering university students' written argumentation via recursive reading: A randomized controlled trial. *Journal of College Reading and Learning*, *52*, 42–63. https://doi.org/10.1080/10790195.2022.2021771

Toulmin, S. E. (1958). *The uses of argument*. Cambridge University Press.

Vandermeulen, N., De Maeyer, S., Van Steendam, E., Lesterhuis, M., van den Bergh, H., & Rijlaarsdam, G. (2020a). Mapping synthesis writing in various levels of dutch upper-secondary education. A national baseline study on text quality, writing process and students' perspectives on writing. *Paedagogische Studiën*, *97*, 187–236.

Vandermeulen, N., Leijten, M., & Van Waes, L. (2020b). Reporting writing process feedback in the classroom: Using keystroke logging data to reflect on writing processes. *Journal of Writing Research*, *12*, 109–140. https://doi.org/10.17239/jowr-2020.12.01.05

Vandermeulen, N., van den Broek, B., Van Steendam, E., & Rijlaarsdam, G. (2020c). In search of an effective source use pattern for writing argumentative and informative synthesis texts. *Reading and Writing*, *33*, 239–266. https://doi.org/10.1007/s11145-019-09958-3

Vandermeulen, N., Van Steendam, E., & Rijlaarsdam, G. (2020d). Dataset - Baseline data LIFT Synthesis Writing project [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3893538

Vandermeulen, N., Van Steendam, E., & Rijlaarsdam, G. (2023). Introduction to the special issue on synthesis tasks: Where reading and writing meet. *Reading and Writing*, *36*, 747–768. https://doi.org/10.1007/s11145-022-10394-z

van Ockenburg, L., van Weijen, D., & Rijlaarsdam, G. (2019). Learning to write synthesis texts: A review of intervention studies. *Journal of Writing Research*, *10*, 401–428. https://doi.org/10.17239/jowr-2019.10.03.01

Wengelin, Å., Frid, J., Johansson, R., & Johansson, V. (2019). Combining keystroke logging with other methods: Towards an experimental environment for writing process research. In Observing Writing (pp. 30–49). Brill.

Weston-Sementelli, J. L., Allen, L. K., & McNamara, D. S. (2018). Comprehension and writing strategy training improves performance on content-specific source-based writing tasks. *International Journal of Artificial Intelligence in Education*, *28*, 106–137. https://doi.org/10.1007/s40593-016-0127-7

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

**Christian Tarchi[1] · Ruth Villalón[2] · Nina Vandermeulen[3] ·
Lidia Casado-Ledesma[1] · Anna Paola Fallaci[1]**

✉ Christian Tarchi
christian.tarchi@unifi.it

[1] Dept. of Education, Languages, Interculture, Literatures, and Psychology, University of Florence, via San Salvi, 12, Florence 50135, Italy

[2] Universidad de Cantabria, Santander, Spain

[3] Umeå University, Umeå, Sweden