Studies in Second Language Acquisition **41** (2019), 555–582
doi:10.1017/S0272263119000251

Research Article

MAPPING MASTER'S STUDENTS' USE OF EXTERNAL SOURCES IN SOURCE-BASED WRITING IN L1 AND L2

Mariëlle Leijten*

University of Antwerp

Luuk Van Waes

University of Antwerp

Iris Schrijver

University of Antwerp

Sarah Bernolet

University of Antwerp

Lieve Vangehuchten

University of Antwerp

Abstract

This article examines how master's students consult and process sources in source-based writing tasks in L1 and L2. Two hundred eighty master's students wrote a text in their L1 (Dutch) and L2 (English) at the beginning and end of the academic year. They wrote these texts based on three sources: a report, a web text, and a newspaper article. Their writing processes were registered using the Inputlog keylogging program. This allowed us to determine how much time the students spent reading the sources, when they did so, which sources they consulted most frequently, and how often they switched between the various (types of) sources. The quality of the texts was assessed

First of all, we would like to thank the master's students in multilingual professional communication at the University of Antwerp for their contribution to the research. Next, we would like to extend warm thanks to the members of the team of teachers who evaluated the texts using D-pac (Mike Beyers, Julie Declerck, Caroline Dothee, Suzy Stals, Jimmy Ureel, Katrien Verreycken). A word of thanks to Sven De Maeyer, who patiently explained the statistical techniques in R. Finally, we owe thanks to Eric Van Horenbeeck, who programmed the source analysis of Inputlog for interval analyses.

*Correspondence concerning this article should be addressed to Mariëlle Leijten, University of Antwerp, Prinsstraat 13, SZ 458, BE 2000 Antwerp. E-mail: marielle.leijten@uantwerpen.be

Copyright © Cambridge University Press 2019

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

holistically using pairwise comparisons (D-pac). Confirmative factor analysis showed three components to be relevant to describe source use in L1 and L2 writing: (a) initial reading time, (b) interaction with sources, and (c) the degree of variance in source use throughout the writing process. Individual text quality remained stable in L1 and L2 throughout the academic year. Structural equation modeling showed that the approach in source use, especially source interaction, is correlated with text quality, but in L1 only.

INTRODUCTION

Today's society demands a great deal from the reading and writing skills of master's students. Not only is the flow of online written information immense, but information is also offered in different ways, and in many languages. Moreover, information can highlight multiple aspects of a topic and may or may not be relevant. It is a challenge for educators to teach students how to manage the flow of information and to convert it into new texts (Martínez, Mateos, Martín, & Rijlaarsdam, 2015). These new texts need to be correct, clear, and coherent and can take many forms, although they must be a faithful representation of the information used as source material. Consequently, because few texts are still written from scratch (Leijten, Van Waes, Schriver, & Hayes, 2014), writers almost always take on alternately the role of reader and writer (Mateos, Solé, Martín, Cuevas, Miras, & Castells, 2014) to arrive at the new text. This reading-to-write process is defined in the literature as a meaning-making genre-based task, in which writers simultaneously have to take into account the context, the readership, the purpose of the text, and culturally and socially determined reasoning techniques (Byrnes & Manchón, 2014; Ellis, 2009; Kellogg, 2008; Lindgren, Leijten, & Van Waes, 2011; Schriver, 2012).

Reading-to-write tasks are cognitively complex because they rely not only on the linguistic skills of writers but also on their reasoning skills and problem-solving behavior (Plakans, 2008; Raedts, 2008; Spivey & Nelson, 1997). Writers must not only understand sources correctly and select the information relevant to the new text but they must also decide how detailed the information should be in that new text. In addition, writers must ensure that the information in the new text is presented in a logical order and forms a coherent whole. The writing task becomes even more complex if the new text has to be written for a different kind of audience than the readers at whom the source texts were directed. Writers will then have to adapt the register and style to match the expectations, needs, and preferences of the new readers (Kellogg, 2008; Schriver, 2012). Writing this kind of text in a nonnative language (L2) poses an even heavier burden on writers, as their productive knowledge of grammar and lexis is more limited than in their native language (L1) (Hinkel, 2003). Consequently, problems with reading the source materials and selecting the relevant information might add to lexical and formulation problems that occur while writing in the L2.

In a recent study (Leijten, Van Waes, Schrijver, Bernolet, & Vangehuchten, 2017), we investigated source use during source-based writing in a group of masters' students (N = 60). The students wrote a text on the basis of source texts in their L1 (Dutch). Using the keystroke logging tool Inputlog (Leijten & Van Waes, 2013), we were able to measure

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

(a) how much time students spent on reading and consulting the source texts, (b) when they did so in the production process, (c) which sources they consulted most frequently, and (d) how often they switched between the different sources. The results of this study indicated that the quality of the texts was related to source use: relatively long and attentive reading before writing, combined with frequent switches between sources during writing lead to high-quality texts. Another interesting finding was that neither the students' source-use strategies nor their text scores in L1 changed over the course of an academic year. Despite these insights, it remains unknown whether the same trends are true for L2 writing because writers may encounter more problems while reading the sources and producing the text in their L2 than in their L1 (as stated previously). This question inspired the current study. Our approach is innovative in the sense that it is the first study in L2-writing research that explores the key variables necessary to describe the use of external sources. Be aware that we deliberately focus exclusively on source use from a process perspective, but that research in the broader field of L2-writing research can benefit from our approach to pinpoint to important constructs in writing processes.

LITERATURE REVIEW

SOURCE-BASED WRITING

As Cumming, Lai, and Cho (2016) convincingly state in their review article, there seems to be a consensus in the literature on writing that learning to write effectively from sources is a fundamental academic outcome, both for L1 and L2 learners. However, how to teach this specific writing competence is still open to debate, according to these same authors. Recent intervention studies fail to come to a one-size-fits-all solution, and this is mainly due to the complex cognitive load of a reading-to-write task (Spivey & Nelson, 1997). After all, writers must compare information from different sources, contrast this information and then organize it into a logical text, by relating the information to their own knowledge and ideas (Martínez et al., 2015; Mateos et al., 2014; Segev-Miller, 2004). Doolan and Fitzsimmons-Doolan (2016) mention several variables influencing this process, such as basic reading competence and vocabulary knowledge but also knowledge of genre-dependent text structure and other discursive features necessary for a correct understanding and interpretation of the reading materials. Indeed, source texts may contain complementary or contradictory information, which makes the interpretation of multiple sources a complex task. Obvious differences may also appear between L1 and L2 performances because of the variation in language proficiency. However, the existing empirical research seems to indicate that currently an absolute distinction between L1 and L2 cannot be made, which may be due to the nature of the (limited) research: "The universal nature and challenges of learning to write effectively from sources make it difficult to draw absolute distinctions between the writing of L1 and L2 students, particularly given the small numbers of student populations and writing tasks examined to date" (Cumming et al., 2016, p. 52).

Consequently, it is not surprising that the only common predictive factors for successful source-based writing in all studies are education and experience (Marzec-Stawiarska, 2016). Clearly, reading-to-write is a competence that needs progressive and iterative instruction and training (Cuevas et al., 2016). However, research into the

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

instructional domain shows that students receive relatively little education and training on multiple source-based writing (see Solé, Miras, Castells, Espino, & Minguela, 2013 for an overview). Without specific instruction, the strategies are often limited to a single reading of the sources and the transfer of this information in the writing in a fairly straightforward manner, or even through copy pasting (Cerdán & Vidal-Abarca, 2008; Lenski & Johns, 1997; McGinley, 1992; Plakans & Gebril, 2012; Solé et al., 2013). Such copy-pasting behavior immediately raises the question of plagiarism, which explains why much research on source-based writing specifically focuses how to deal with source integration in an appropriate and ethical way (for a review we refer to Liu, Lin, Kou, & Wang, 2016). Of course, "textual borrowing" is more tempting in L2 writing than in L1 because of the already mentioned vocabulary knowledge factor (Neumann, Leu, & McDonough, 2019; Nguyen & Buckingham, 2019; Plakans & Gebril, 2013). A study by Gebril and Plakans (2016) confirms that lexical diversity in L2 reading-to-write tasks is indeed best predicted by the source material rather than by the learners' lexical competence.

In the research presented in this article, we will focus on the within-subject comparison of source use in L1 and L2 writing, as suggested in the review article by Cumming et al. (2016, p. 53) as one of the necessary future research directions for source-based writing studies. This way, we want to get a better grasp of the underlying processes that advanced writers go through to incorporate source information into their texts. Unfortunately, these underlying writing processes have been largely ignored in the literature, although there are a few case studies that describe the interaction with sources. For example, McGinley (1992) and Lenski and Johns (1997) advance two models of source-use strategies: a linear or sequential approach, on the one hand, and a recursive approach, on the other. In linear or sequential strategies, the information from the source texts fails to be integrated into a whole. These linear writing processes can be situated at the level of pure knowledge telling or retrieval (Bereiter & Scardamalia, 1987). In recursive processes, writers transcend this level and succeed in successfully integrating the information from the sources into a new and coherent text. We then speak of knowledge transformation (Bereiter & Scardamalia, 1987) or even knowledge crafting if the perspective of the reader is also taken into account in the message of the text (Kellogg, 2008).

METHODOLOGICAL ISSUES IN SOURCE-BASED WRITING RESEARCH

The previously mentioned models on the use of strategies in source-based writing have been formulated on the basis of think-aloud protocols. A distinct advantage of thinking-aloud protocols is that they provide information about what writers think during the task, which gives insight into why they are performing the observed writing activities. However, this is also the main disadvantage: the reading and writing processes are interrupted and to some extent also disrupted, which explains why thinking aloud is considered a reactive research method (Bowles, 2010). With the current technology of keystroke logging, such as Inputlog and Scriptlog, it is possible to observe (digital) writing processes in a way that does not disturb the writer (Chan, 2017; Johansson, Wengelin, Johansson, & Holmqvist, 2010; Leijten & Van Waes, 2013; Lindgren & Sullivan, 2019). Leijten et al. (2014) used this technology, for example, to examine the source use in writing processes by a professional writer in an elaborate case study. The

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

results of this study allowed them to refine Hayes's latest writing model (2012) by adding a Searcher, among others. The Searcher refers to the strategies that are used by professionals in their writing process to search for and in sources according to their writing goals. Following the example of the Leijten et al. (2014) case study, Chan (2017) used a similar research method to compare how a PhD student and a postgraduate master's student wrote an essay based on two sources. The writing process was recorded with Inputlog and a retrospective interview was conducted to gain insight into the choices the writers made during the process. The laborious triangulation of the qualitative data with the key-registration data showed that one writer consulted sources less often to formulate his own text immediately after reading, while the other worked much more fragmentedly by frequently switching between the text and the sources. The former managed to translate the ideas from the sources much better, whereas the latter had great difficulty producing his own text.

Keylogging, and Inputlog in particular, makes it possible to gather—in a very detailed way—information about the progress and dynamics of both source use and writing processes. However, as the two case studies described show, it is quite challenging to get a grip on the enormous number of variables, approximately 900 variables, that are generated by various fine-grained analyses of Inputlog. Thus, the central question is what variables are related to one another and are relevant to serve as indicators for the description of source use during writing. Leijten et al. (2017) presented a first attempt to uncover these variables. The current study aimed to validate the model proposed in Leijten et al. (2017) and to verify whether the same variables are relevant in the production of L1 and L2 texts.

SOURCE USE AND COGNITION

When writers use sources during their writing process, they interrupt their text production. In the literature, the continuous production of text between interruptions is called a burst. Bursts can be defined from different perspectives (Alves, Castro, & Olive, 2008; Chenoweth & Hayes, 2001, 2003; Limpo & Alves, 2017). For example, a P-burst is an uninterrupted writing unit delineated by pauses of a certain threshold value. An R-burst is an uninterrupted writing unit that is bound by revisions (Hayes & Chenoweth, 2006).

Research into cognitive processes during writing tasks pays special attention to these moments (Barkaoui, 2019; Van Waes & Leijten, 2015). The reason why bursts have aroused the interest of writing researchers is that earlier empirical work had shown that the burst length gives insight into the skills and fluency of writers. The theory assumes that the ability to write in longer bursts means that a writer has more mental space available to activate different subprocesses of writing simultaneously (Alves & Limpo, 2015, p. 3; Limpo & Alves, 2017). The way in which writing research looks at bursts has so far been limited to revision or planning: either the production of text is interrupted to remove, change, or insert a piece of text (R-burst, aimed at revision) or the production of text is interrupted by a pause after which the production of text continues (P-burst, aimed at planning, or formulation [Révész, Kourtali, & Mazgutova, 2017]). Because source use—of whatever nature—has become essential in today's digital text production, Leijten et al. (2014) argue that so-called "S-bursts'" (due to pauses targeted on sources)

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

also play an important role in fluency: the production is interrupted by a switch or transition to an external source. How writers consult digital sources influences, among other things, the number and length of S-bursts. In other words, whether writers reread (fragments from) sources regularly or in very concentrated phases, and whether they choose to copy text from the source, are variables that influence fluency, or the fragmentation of the writing process.

This new, extended view on bursts was demonstrated in the explorative case studies by Leijten et al. (2014) and Chan (2017), exemplifying the added value of keystroke logging for this new approach. More recently, Leijten et al. (2017) deepened the reflection on S-bursts in writing processes. We first distinguished between the time students spent on reading sources and the time spent on writing. Subsequently, we focused on when the students read the sources (i.e., before starting to write the text or during writing) and when in the process the students paused and produced text (see Figure 1).

Subsequently, we used a principal component analysis (PCA) to find within the 900 variables those variables that were most fit to describe and explain the use of sources during source-based writing in L1 (see "Methodological Issues in Source-Based Writing Research"). This analysis enabled us to find clusters of variables that are related, suggesting that they measure aspects of the same underlying dimension (i.e., components). It also allowed us to establish how an individual variable contributed to that component. In the component extraction, we used a factor loading of .60 or higher and exclusively selected those variables that concerned ratios and average values, which implied that they were generalizable for different writing processes (e.g., variables such as absolute total process time was excluded). The PCA was carried out on Inputlog data collected among a group of 60 masters' students at the beginning and the end of the academic year (Leijten et al., 2017). The results of the PCA indicated that source use during source-based writing can be described by a combination of three different components consisting of eight underlying variables (see Table 1).

As mentioned in the preceding text, these three components were derived from the data. The underlying variables for the first component *initial reading time* are the proportion of initial reading time for the sources on the total time spent on sources and the number of switches between sources in the first interval. The second component, *source interaction*, refers to the number of sources that is consulted per minute, both in general and per source content category, as well as the number of switches between sources. *Variance in source use* indicates the extent to which writers vary in their

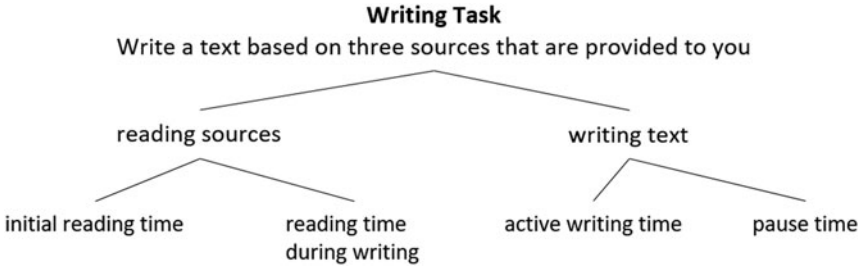


FIGURE 1. Distribution of activities during source-based writing processes.

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

TABLE 1. Three components and eight variables that describe source use in source-based writing

Components	Variables
Initial reading time	Proportion initial reading time vs. total reading time of sources
	Number of source switches during initial reading time (per min)
Source interaction	Number of sources (per min)
	Number of sources per content categories (per min)
	Number of switches between sources (per min)
Variance of source use	Relative time spent on text during process
	Relative time spent on text during first interval (out of 5)
	Variance of time spent on text during process

interaction with sources during the writing process, both in relative time spent on source consultation, time spent on writing in the first interval, and relative time spent on writing during the task. We consider this last component as particularly relevant because previous studies have explained the importance of within-subject variance (or fragmentation) in the writing process: for example, process variation as a component of describing fluency in L1 and L2 (Van Waes & Leijten, 2015), but also pausing variation of skilled and less skilled L2-writers (Chukharev-Hudilainen, Saricoaglu, Torrance & Feng, 2019; Xu & Qi, 2017).

Together, the three components explained 74.6% of the variance in the process data. When we compared the results between the two test moments, we found no differences in the component loads, suggesting that the students' writing and source-use strategies did not differ at the beginning and at the end of the academic year. Interestingly, the quality of the texts also remained stable across both test moments. There was no overall relation between text quality and one of the three components described previously, but, when we compared texts with high scores (highest 20%) and low scores (lowest 20%), there were significant differences to be observed. As mentioned in the Introduction, the initial reading time and the interaction with sources was significantly different: students who wrote higher-quality texts read relatively long and attentive before writing, and switched more often between sources during writing than students who wrote texts of poorer quality.

The current study is a follow-up study of Leijten et al. (2017), but with a wider scope. It presents a within-participant comparison between source-based writing in L1 (Dutch) and L2 (English).[1] Based on the literature, we know that some aspects of the writing process will differ between L1 and L2: L1 and L2 written texts vary in their lexical choices (Crossley & McNamara, 2009), for example, and writers pause significantly more in L2 than in L1 (Van Waes & Leijten, 2015). On a higher, metalinguistic level, however, L1 and L2 writing are not always very different. Knowledge about common text features (e.g., discourse markers and counterarguments), for example, appears to be merged between L1 and L2 as writers' writing knowledge increases, forming a single system (Kobayashi & Rinnert, 2012). Another metalinguistic skill in L1 and L2 writing, which could be similar across languages, is the use of sources during writing. Hence, the

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

first aim of the study was to verify whether the same variables can be used to describe and predict source use during source-based writing in L1 and L2. The first step to do this was to conduct a confirmative factor analysis on the L1 dataset, which comprises the exact data ($N = 101$) of Leijten et al. (2017) as well as newly collected data ($N = 280$), to confirm the exploratory findings from Leijten et al. (2017). The second aim of the present study was similar to the objective of Leijten et al. (2017): we investigated which variables linked to source use are correlated with text quality. On the basis of text quality and process data we investigated whether, like in L1, the students' approach to source use and text quality remains stable across both test moments.

The preceding principles led to the following research questions:

1. How do master's students deal with sources when writing source-based texts in L1 and L2?
2. How stable is the individual text quality in L1 and L2 during an academic year?
3. To what extent is there a relationship between the approach to the use of sources in L1 and L2 and the quality of the text?

METHOD

PARTICIPANTS

There were 280 students of the master in multilingual professional communication from the University of Antwerp who took part in this study (16% male, 84% female, which resembles the gender proportion in the master's program). The data were collected in three consecutive academic years; the first-year data ($N = 60$) have been previously used in an exploratory study (Leijten et al., 2017). All students (aged between 20 and 34 years old; average: 22 years and 10 months), native speakers of Dutch, spoke and wrote Dutch at C2 level (cf. Common European Framework of Reference).[2] The students followed the master's program of communication courses both in Dutch and in one—or more—second/foreign languages. (English, $N = 138$, C1 level; the other students follow either French, German, or Spanish; they were tested or asked for B2 accreditation as a necessary requirement to start the course at B2 level when starting the course, aiming at C1 by the end). In this study we will focus exclusively on L1 Dutch and L2 English. The selection of this particular population of master's students is motivated by our interest in source use in source-based writing by writers with advanced writing skills. The master's students under study had previous experience in academic writing, a particular form of source-based writing that requires including key ideas from sources, using multiple sources, and integrating the source material appropriately (Plakans & Gebril, 2013). Moreover, they had a near-native proficiency in their L2 (C2 in Dutch; C1 in English). By choosing this particular student population for this study, we wanted to avoid that participants' source use would be determined by lower-level writing problems (e.g., selecting important information from sources and appropriately synthesizing source materials) rather than higher-order synthesis skills, especially in the L2 (see Plakans & Gebril, 2013), to keep the comparison between L1 and L2 source use as untainted as possible. Furthermore, we wanted to examine how source use is characterized at a high level of writing competence to infer what should be focused on in previous years of first and second language education to reach that level.

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

MATERIALS AND TASKS

The participants were given two similar writing tasks at the beginning and end of their master's program in L1 (Dutch) and L2 (English), with an interval of 6 months. So, each student wrote four texts for this experiment during the academic year. The students wrote a text of 200 to 250 words based on three digital sources on one of the following topics: humanitarian aid, renewable energy, animal rights, and climate change. The full prompt can be found in Appendix A.

For each of the four topics, we developed similar source material in three different genres: a report, a web text, and a newspaper article. The source texts for all topics were drawn up in such a way that they were very similar in terms of difficulty, for example by keeping the text length, the average number of words per sentence, and the average word frequency as comparable as possible. Table 2 shows an overview of the material in Dutch and English.

The participants were asked to write for a specific target group: last-year students of secondary education. We opted for a target audience to which the participants could relate. The diversity of content and language between the three sources, as well as the difference in target audience of the sources and the new text, were deliberately chosen to increase task complexity. The master's students who took part in this research were supposed to have advanced reading and (academic) writing skills. By using a writing task that resembled academic writing but that focused instead on the use of sources directed at different target audiences than the text to be produced, we were able to test whether these students were able to write a coherent new text for a younger target audience.

The participants were given 40 minutes per writing task and they were free to consult the Internet for more content and linguistic information and use online tools such as dictionaries. The topic, language, and order of the writing tasks were counterbalanced.

DATA COLLECTION

The writing processes were registered with Inputlog 7. Inputlog is a key registration program that records all keystrokes, mouse movements, and Microsoft Windows activities (Leijten & Van Waes, 2013; Leijten et al., 2014; Inputlog is freely available for

TABLE 2. Overview of the text length of the provided sources per theme and per language (mean number of words per sentence)

	Humanitarian aid	Energy	Animal welfare	Climate change
Dutch				
Report	316 (18.6)	321 (18.9)	321 (18.9)	323 (19.0)
Web text	237 (8.4)	229 (8.2)	227 (8.1)	227 (8.1)
Newspaper article	244 (15.3)	239 (15.0)	246 (15.4)	246 (15.4)
English				
Report	325 (19.1)	313 (18.4)	318 (18.7)	304 (19.0)
Web text	246 (8.8)	230 (8.9)	239 (8.9)	242 (8.9)
Newspaper article	251 (15.6)	242(15.1)	250 (15.6)	244 (15.3)

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

researchers and can be downloaded from www.inputlog.net). All activities of a writer to arrive at a final text are recorded with the corresponding time indication. This detailed log file formed the basis for further analyses (see "Data Preparation and Data Analyses").

PROCEDURE

The experiment, in which all students took part simultaneously, took about two hours and consisted of three parts: a typing task and two writing tasks. The students started their session with a typing task (Van Waes, Leijten, Mariën, & Engelborghs, 2017), which allowed us to determine the general typing speed of the students, but at the same time permitted the students to get used to the computer room and the experimental setting and to "warm up."[3] Subsequently, they were allowed to start the first writing task.

For each writing task (both at the beginning and at the end of the academic year), students were given a maximum of 40 minutes. Students who needed less than 40 minutes for the writing task performed an in-between task in the remaining minutes, in which they were asked to describe the strengths and weaknesses of their master's program in the L1. By letting everyone type for 40 minutes we tried to avoid that students would be distracted and/or influenced by fellow students who took less time for the writing task. After a short break, the students performed the second writing task, for which they were again given 40 minutes. At the end of the session, all students completed a form in which they gave us permission to process the experimental data and in which they were informed about their right to opt out of the experiment at any time.

TEXT ASSESSMENT

The overall quality of the texts was assessed holistically using pairwise comparisons. For this purpose, we used the Digital Platform for Assessment of Competences (D-pac, www.d-pac.be; Van Daal, Lesterhuis, Coertjens, van de Kamp, Donche, & De Maeyer, 2017). This platform shows two texts (A vs. B) and offers a choice between them: "Which text is the best elaboration of the task?" We opted for this comparative assessment procedure because research has shown that evaluators estimate the quality of texts in a more reliable way when they can compare two texts than when they have to score an isolated text (Pollitt, 2012). Furthermore, through pairwise comparisons each text is judged by multiple experts and therefore is a result of the judges' collective expertise (Pollitt, 2012).[4]

For the assessment in D-pac, we selected the texts of those students that performed the writing task on both measurement moments. For L1 Dutch, 68 students (N *texts* = 136) produced texts at the beginning and end of the academic year. Ten experienced assessors, all of whom were involved in the master in multilingual professional communication, assessed the texts independently of one another. The texts were anonymized. The order in which all texts were assessed was also completely randomized for the 10 evaluators. The average Intra Class Correlation (ICC) based on standardized scores was .79 (high reliability). For L2 English the texts of 35 students were assessed (N *texts* = 70) by four experienced assessors, who were all affiliated with the Faculty of Arts. The same procedure as for Dutch was followed. The average ICC based on standardized scores was .78 (high reliability). We strived to reach high reliability and therefore each text in Dutch

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

was assessed 20.1 times on average and each text in English was assessed 19.8 times (van Daal et al., 2017).

DATA PREPARATION AND DATA ANALYSES

PREPARATION AND ANALYSIS OF INPUTLOG DATA

The keylogging data ($N = 280$) were iteratively prepared in Inputlog version 7.1 for further data analysis. In doing so, we used the following analyses that are integrated in Inputlog (see Inputlog User's Manual; www.inputlog.net):

- The summary analysis: focusing on the process duration, text length, produced characters, and distribution of pause and writing time
- Pause analysis: focused on the pause behavior at different text levels (within words, between words, phrases, sentences, paragraphs) and time intervals.
- The source analysis: focusing on the use of digital sources during the writing process (number of sources, type of sources, time in the sources, distribution of source use over time intervals).

First of all, we carried out a general data analysis on the complete data files of the whole writing task (total time on task: initial reading and activities both within Microsoft Word and in other sources) using the summary analysis (with pause threshold 2000ms) as well as the pause analysis (with pause threshold 200ms). On the one hand, we opted for a general pause threshold of 2 seconds to calculate the number of P-bursts (Van Waes, Leijten, Lindgren, & Wengelin, 2015). These bursts indicate a pronounced cognitive load during the process interruption (linked to planning or revision). On the other hand, we opted for a pause threshold of 200ms to take into account the pauses related to a lower cognitive load (Van Waes et al., 2015). This enabled us to only consider those process delays that are most likely not linked to the time it takes writers to move their fingers from one key to another. In a second step, we performed both the summary analysis and the pause analysis again, but now only for the writing processes that took place within the text document.

Finally, we conducted a source analysis. Inputlog identified a total of 3,104 different focus events. A focus event is a technical approach to the source interaction, in which each computer window opened by the writer is registered and identified in the log file. We have used this technical approach as the basis for a substantial recoding of the data: main text, report, web text, newspaper article, and other sources. We divided the other sources that the students consulted into six categories and recoded them. Table 3 provides an overview and description of these additional external resources and tools that the participants have consulted. A detailed description of this categorization procedure can be found in the online supplementary information.

CONFIRMATIVE FACTOR ANALYSIS AND STRUCTURAL EQUATION MODELING

One of the aims of the present study was to provide a diversified insight into the use of sources and the interaction with the texts. Key was to explore whether and what underlying structure can be found in the large number of variables that can be used as indicators for source use during writing. After all, it is important to reduce the large

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

TABLE 3. Six categories of the other consulted sources

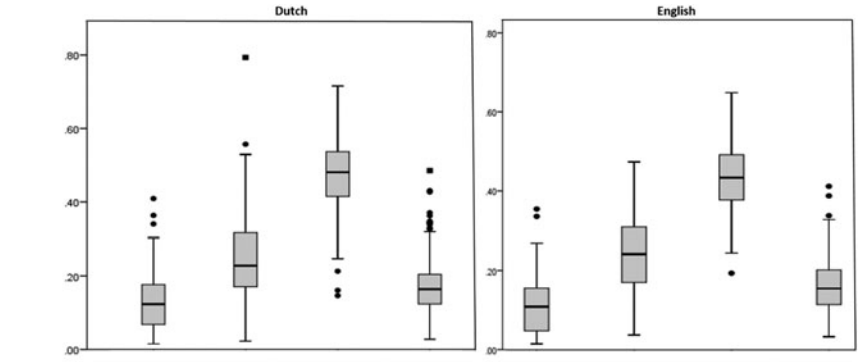
Content (19%)	Background information on the topics (e.g., report on the refugee crisis in a Flemish news program)
Language (64%)	Language: general (53%) (e.g., spelling of *crisis* vs. *crises*) Language: synonyms (11%) (e.g., other word for *regulate*)
Other (12%)	Activities that cannot be classified as content, language, search, or task (e.g., opening Media Player)
Search (4%)	General searches online or on the computer (e.g., opening Google)
Task (1%)	Searches related to the task (e.g., what is a synthesis?)

collection of variables provided by keylogging to a useful and manageable set that enables us to describe and explain the variance in source use as well as possible. That is why we analyze the data here at a text level and not at a personal level. We realize that this could create a certain degree of dependency in the analysis. However, the texts that the participants wrote concerned various topics and there were six months in between the two test moments.

In this study our analyses mainly relied on CFA and structural equation modeling (SEM). The analyses were conducted in R (version 3.4.3 using the Lavaan library). We opted for these analyses because we wanted to verify whether the component structure and the underlying variables that were found in a previous (smaller scale) study by Leijten et al. (2017) could be confirmed and further elaborated in this new larger scale study. Finally, to analyze the differences between the sessions (beginning vs. end of academic year) and/or languages (L1 vs. L2), we used a generalized linear model (repeated measurements, multivariate analysis). We will briefly explain the stepwise procedure we followed in the analysis of the data.

The CFA allowed us to examine the nature of and the relations among latent constructs, derived from principal component analysis (PCA) reported in a previous study (Leijten et al., 2017; which were based on the data of the first cohort $N = 60$). CFA explicitly tests *a priori* hypotheses about relations between observed variables (i.e., process measures describing source use during writing) and latent components (i.e., dimensions that are not directly observed but inferred from underlying, highly clustered variables). In the current study, the latter were taken from the exploratory PCA that is presented and discussed in Leijten et al. (2017). So as to avoid of sample dependence, we conducted the CFA on a reduced dataset excluding the logfiles used in the PCA analysis and limiting it to a single observation per participant. The CFA model contained three components and eight underlying variables (see Figure 2). Following an iterative procedure, we optimized the model, striving for an optimal balance between "goodness-of-fit" (i.e., how well the model fits the observed data) and "parsimony" (i.e., limited number of variables that significantly contribute to the model). We used the following fit indices. We started with the chi-squared test indicating the difference between the observed and the expected covariance. Next, we looked at the comparative fit index (CFI) that compares the resulting model with a so-called zero model (without any relations).

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press



	Initial Reading Time	Source Time	Active Writing Time	Pause Time
Mean	10.6%	24.5%	47.5%	17.3%
Std. Deviation	8.1%	10.4%	9.5%	7.2%
Lower Bound	9.6%	23.3%	46.4%	16.4%
Higher Bound	11.5%	25.7%	48.6%	18.1%

	Initial Reading Time	Source Time	Active Writing Time	Pause Time
	10.8%	26.1%	45.8%	17.2%
	7.8%	10.2%	8.5%	7.3%
	9.4%	24.4%	44.4%	16.0%
	12.1%	27.8%	47.2%	18.4%

FIGURE 2. Overview of the initial reading time, source time, active writing time, and the pause time in percentages (100%) for Dutch (N = 280) and English (N = 138) (including lower and higher bounds).

The literature recommends a CFI score higher than .90 (or even .95; Hu & Bentler, 1999). Finally, we used the root mean square error of approximation (RMSEA), that is, the square root of the discrepancy between the sample covariance matrix and the model covariance matrix. A value lower than .008 refers to a good fit, values between 0.08 and 0.10 to a moderate fit (Hooper, Coughlan, & Mullen, 2008).

We first applied the CFA to the Dutch data (L1) and then extended the model to the English data (L2), using the multiple group measurement invariance analysis. This enabled us to test whether the variables used in the analysis are indeed comparable constructs for both languages: we checked the fit respectively for the loadings, the intercepts, and the variables' means. Based on this assessment, parameters in the CFA model can be set equal or they can vary across languages. The same fit indices were used to evaluate the measurement invariance.

In the final step we included text quality in the analysis, using SEM. This allowed us to estimate the relation between text quality and the three latent components that were used to define the process characteristics of source interaction in the CFA model.

RESULTS

GENERAL INFORMATION ON SOURCE USE DURING WRITING

As stated before, the students were given a maximum of 40 minutes to write the text. The average time they needed in Dutch (L1) was 32:48 minutes (SD = 7:07) and 35:01 minutes (SD = 6:16) in English (L2). As shown in Figure 2, the students spent an average of 64.8% (SD = 10.1%) of their time writing the text in Dutch and 35.1% (SD = 10.1%) reading the sources (see Figure 2). For English the results are quite similar: on average 63% was spent in the text document and 37% in the sources. Of the total process time, the students reserved about 10% to orient themselves and get acquainted with the given

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

source materials in the initial reading phase. An additional quarter of the time was used to consult sources later on in the writing process. This means that students reserved 60 to 65% of their process time to the text, of which more than a quarter is used for pausing (pause threshold > 2,000 ms). An analysis of variance showed no main effect between both languages ($F(4, 415) = 1.071$; $p = .371$; $\eta^2 = .010$), nor was there a significant effect of language on initial reading time, source time, active writing time, and pause time.

As explained in "Materials and Tasks," the students were provided with three sources. As shown in Table 4, the writers spent on average three and a half minutes on the report, two and a half minutes on the web text, and three minutes on the newspaper article for the Dutch writing task. Almost three minutes were spent consulting additional sources (most of them theme related; a few of them language related). In English the writers spent on average four and a half minutes on the report—which is a minute longer than in the Dutch task, two and a half minutes on the web text, and almost three minutes on the newspaper article. Therefore, in English the use of additional sources increased by 19%. They spent about the same time in the given sources, only the text type differed: the report was read 16% longer and the reading time of the newspaper article decreased by 16%. Finally, in the English task students spent about three and a half minutes on other sources, which is also almost a minute longer than in the Dutch writing task (Table 5). This difference is mainly related to an increase in consulting language-related resources in the English condition (e.g., online dictionaries, thesauri, and grammars).

DESCRIPTIVE CASE STUDY OF SOURCE USE DURING WRITING

To yield a more detailed picture of the possibilities (and limitations) that keylogging offers to describe the complexity of source interaction, we will present a case study. We will also illustrate the selected case graphically, using a visualization based on the log data (see Pajek output and Process-graphics in Inputlog; Leijten et al., 2014). This will enable us to highlight the most important indicators that characterize source use during writing and link these indicators to the four components described in Figure 1.

For this case study we selected one writer, whom we will call Frida. Frida wrote a high-quality text in both languages. In this descriptive case study, we present the

TABLE 4. Overview of the mean duration (in minutes) and the mean proportion of time spent consulting the provided sources and other sources involved

	Text	Report	Web text	Newspaper article	Other (cf. Table 3)
Dutch					
Mean duration (*SD*)	21:16 (5:45)	3:38 (1:39)	2:20 (1:19)	3:09 (1:27)	2:48 (2:50)
Percentage (%)	64.1	10.9	7.1	9.5	8.4
English					
Mean duration (*SD*)	22:05 (5:11)	4:29 (2:14)	2:27 (1:71)	2:48 (1:21)	3.32 (2.40)
Percentage (%)	62.4	12.7	7.0	7.9	10.0

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

TABLE 5. Factor loadings for the CFA model in the Dutch and English language condition, including descriptive values for the underlying variables

	Dutch (N = 149)			English (N = 99)			Sign
	CFA loading (Std. all)	*M*	*SD*	CFA loading (Std. all)	*M*	*SD*	
Initial reading time							
Proportion initial reading time vs. total reading time of sources	.912	0.294	0.204	.995	0.296	0.205	.951
Number of source switches during initial reading time (per min)	–.778	4.236	5.034	–.564	3.587	5.349	.334
Number of switches between sources (per min)	–.248	4.580	2.387	–.221	5.015	3.292	.230
Source interaction							
Number of sources (per min)	1.095	0.296	0.179	.889	0.394	0.216	**< .001**
Number of sources per main categories (per min)	.691	0.157	0.054	.873	0.167	0.037	.122
Number of switches between sources (per min)	.206	4.580	2.387	.482	5.015	3.292	.230
Variance of source use							
Relative time spent on text during first interval (out of 5)	–.952	0.274	0.201	–.998	0.254	0.196	.438
Variance of time spent on text during process	.794	0.396	0.140	.737	0.422	0.140	.154

process that leads to her English text. Frida needed 40:23 minutes to complete her text. Additional analyses show that she produced 407 words, of which 232 were kept in the final text (.57 product-process ratio). She spent 26.7% of her writing time consulting source materials and made 133 transitions between her own text and the sources she consulted, which is approximately three per minute.

The visualizations in the network diagram presented in Figure 3 clearly show the distribution of her time attributed to each of these sources: the size of the circles represents the proportion of the total time spent, the arrows indicate the interaction between the sources, and the corresponding figures indicate the number of switches between the respective sources. At the bottom of the figure, we show a process bar linked to a time axis. This illustrates at what moments in the process the writer decides to consult sources (line at the top) or to work in the text (line at the bottom). The vertical lines indicate the switches that cause an S-burst.

Frida's process bar indicates that during the first 7:29 minutes (initial reading time) she concentrates on the sources provided. She divides her time to reading (part of) the report (2:00 min), the newspaper article (1:49 min), and the web text (3:40 min), in that order. At the end of this initial reading phase, she decides to copy and paste two sentences from the web text, which she immediately edits and rephrases. During the next 10 minutes she produces the first part of her text, switching frequently between her own text and the sources provided. However, she also occasionally consults an online thesaurus, for example, looking for a synonym for the verb *to provide* (she chooses *supply with*). In the

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

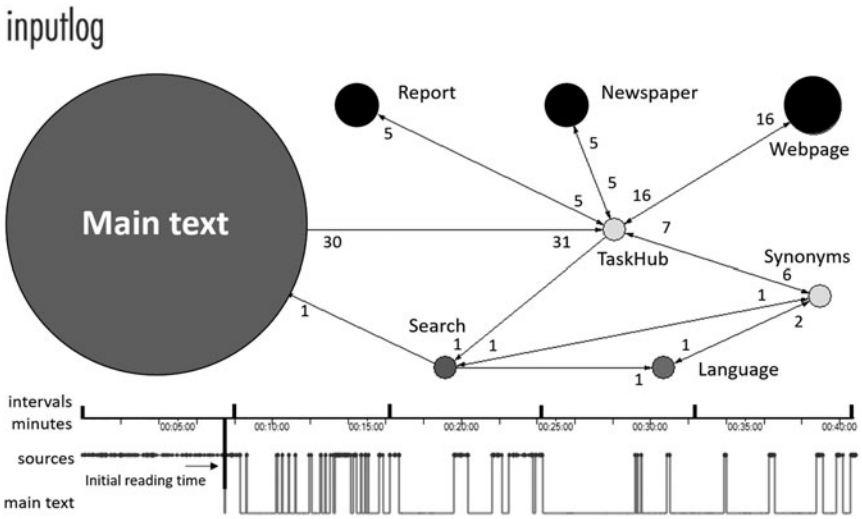


FIGURE 3. Visualization of Frida's source use. Above: The circles show the relative proportion of the time spent in the text and the sources used; the arrows (and the numbers) indicate the switches between sources. Below: A timeline is combined with a graph indicating the moment and the duration of each source switch.

next phase, until about minute 31, she is focused on her text, producing text very fluently with only a few switches. Hardly any copy-paste actions take place, and the writing bursts are getting longer (e.g., between 24 and 28 min). The newspaper article and the thesaurus/translation dictionary are the main sources Frida relies upon. She switches often between the sources by using the taskbar, which is called TaskHub in Figure 3.

Additional analyses show that Frida reserved the last 10 minutes for revision. She systematically rereads and edits her text in three rounds, going repetitively back and forth through her text from the beginning till the end. The time she spends on the provided sources is equally spread between the three texts and is very short, which can be deduced from the density of the vertical lines. She mainly uses the sources now for fast fact checks, combined with a few searches for language-related issues.

To sum up, Frida's writing process shows a pattern in which the sources are mainly used in the first interval (i.e., the first 8 min, when we divide the process into five equal parts). From then on resources are only occasionally accessed. In the final two intervals the proportion of time spent in the text even increases to more than 95%, showing the high variance of source use—and fragmentation—in her process. This case study clearly shows that the distribution of source use can vary considerably across the task process. The same applies to other descriptive indicators, like switches and proportion or type of source use.

CONFIRMATIVE FACTOR ANALYSIS OF SOURCE USE

As said, Leijten et al. (2017) proposed a model, based on PCA, which explained about 75% of the variance to describe source use in source-based writing in L1. We conducted

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

a CFA, following the procedure described in "Confirmative Factor Analysis and Structural Equation Modeling," to check these findings.

The basic CFA model, which involved the three components and the eight variables of the PCA, showed a good fit, but also provided indications for further improvement. The iterative process allowed us to remove one variable (i.e., "Proportion of time in text versus total time on task," which was part of the "Source variance" component; parsimony criterion), which makes the analysis more manageable and lean. Moreover, the model was further optimized by adding a cross-loading of the variable "Number of switches between sources (per min)" on the "Initial reading time" component. The fit indices for this final CFA model as presented in Table 6 shows a good fit of the model (chi square for Dutch $\chi^2(10) = 12.5$; $p = .253$; CFI = .993 > .95; RMSEA = .043 < .10 and English $\chi^2(12) = 22.4$; $p = .003$; CFI = 0.959 > .95; RMSEA = .098 < .10). Therefore, this model was used as the reference model for further analysis.

Interestingly, a multivariate analysis showed an overall main significant effect of the language used ($F(7, 239) = 3.363$; $p = .002$.; $\eta^2 = .991$). However, only one out of the seven variables included in the model showed a significant difference, more specifically "Number of sources" (per min). This variable belongs to the "Source interaction" component (see Table 3). This is an indication that students tend to use a greater variety of sources per minute when composing their text in L2 (English) than when writing in their L1 (Dutch). This was confirmed in the measurement invariance analysis, comparing both languages at the component level, showing that only the "'Source interaction" significantly differs between both languages ($SD = –0468$ ($SE = 0.012$); $p < .001$). For the other two components, the students' source use in L2 seemed to be in line with what they do when composing their text in L1 (Table 5).

To check whether the model can also be applied to describe and explain source use in source-based L2 writing, we carried out a multiple group measurement invariance analysis. This analysis showed a strong/scalar invariance for both language conditions: constrain item intercepts equal across both languages in combination with constrain factor loadings (see Table 6). Only the fit for the variable means across languages proved to be significant, as illustrated previously. This allowed us to fixate the model for both languages at both the loading and intercept level when setting up the final SEM-model including quality scores (see "Relationship between Source Use Components and Text Quality"), meaning that the model can be used to describe both L1 (Dutch) and L2 (English) source-based writing.

TABLE 6. Fit measures of the interlanguage measurement invariance analysis

measure[a]	Chi square	Δ *Df*	significance (*p*)	CFI	RMSEA
Loadings fit	43.622			0.971	0.081
Intercepts fit	46.572	4	.5662	0.972	0.073
Means fit	57.929	3	< .01***	0.959	0.084

Note: Fit indices: p > .05; CFI > .95; RMSEA < .10

[a]Configural fit model did not converge.

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

STABILITY OF TEXT QUALITY AT BOTH MEASUREMENT MOMENTS

As indicated in "Text Assessment," professional evaluators assessed the text quality using pairwise comparisons in Dutch (L1) and English (L2). General Linear Model (GLM) repeated measures on the Z-scores per language showed that there is no difference in the average quality of the texts in both languages at the two measurement moments (Table 7). The average Z-score of the Dutch texts at measurement moment 1 was –.07 (SD = 1.07) and at measurement moment 2 (6 months later) .11 (SD = .90). The quality of the texts proved to be very similar at the two measurement moments ($F(1, 67) = 2.204$, $p = .142$, $\eta_p^2 = .032$). The same holds for English because the average Z-score of the English texts at the first measurement moment was –.21 (SD = .80) and at the second measurement moment (6 months later) .21 (SD = .98). However, this positive trend in the quality scores did not prove to be significant: $F(1, 34) = 3.728$; $p = .062$; $\eta_p^2 = .099$ (Table 7).

Text quality may remain stable between the two test moments in both languages, but then the question arises, how does product quality relate to source use? Are high-quality writing products characterized by specific strategies in source use? And is this potential relationship identical or similar in L1 and L2?

RELATIONSHIP BETWEEN SOURCE USE COMPONENTS AND TEXT QUALITY

To assess a potential relationship between the source components and text quality (as assessed with D-pac), we extended the CFA model, discussed in "Confirmative Factor Analysis pf Source Use," to include text quality using SEM. The resulting model is shown in Figure 4.

Figure 4 shows that in the L1 (Dutch) condition, text quality is significantly determined by "Source use" with a path loading of .26. This means that when the (standardized) score for the "Source interaction" component increases with 1 SD, the (standardized) text quality score will increase with .26 SD. This significant effect was not confirmed in the English L2 condition. The two conditions (L1 and L2) differed significantly ($p < .001$), more specifically with respect to the size of the coefficient for the relationship between source use and text quality (Table 8).

DISCUSSION

This study has provided insight into the way in which master's students consult and process external sources in their L1 and L2 writing process and how source use impacts

TABLE 7. Mean quality (Z-scores) per test moment in Dutch and English (SD)

	Moment 1 (October)	Moment 2 (April)	Significance
Dutch (N = 68)			
Mean quality score (SD)	–.07 (1.1)	.11 (.9)	.142
English (N = 35)			
Mean quality score (SD)	–.21 (.80)	.22 (.98)	.062

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

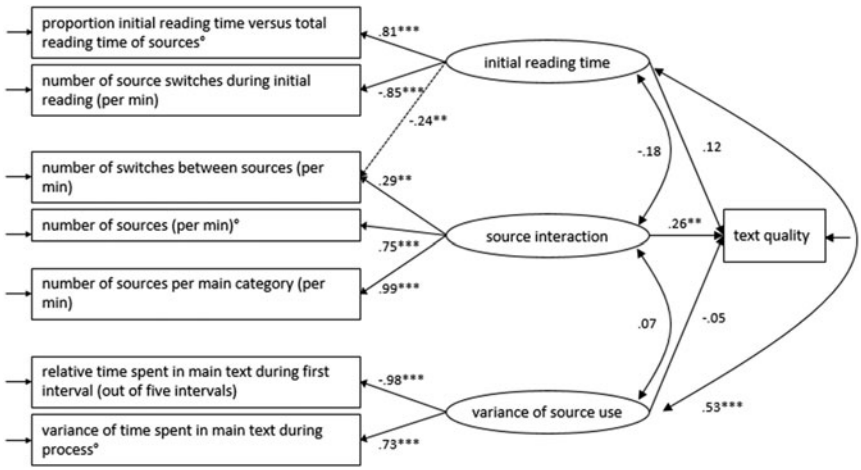


FIGURE 4. SEM model showing the effect of source interaction on text quality in the Dutch condition. °For these variables the nonstandardized factor loadings are fixed to 1.00 so as to make the model identifiable.

text quality in L1 writing. in the following text, we first discuss the major findings of the study in relation to the existing literature in L1 and L2 writing processes, before pointing out the theoretical, methodological, and pedagogical implications of these findings and new avenues for research.

Based on a CFA, we have mapped out the variables that can be used as indicators to describe and predict source use in the writing process of a source-based writing task. Three components were confirmed that hold for both Dutch (L1) and English (L2): first of all, the time writers spend in the planning phase of the writing process on reading sources; secondly, the interaction with the sources: the number of sources consulted and the number of switches between the sources; and thirdly, the extent to which source use varies during the different phases of the writing process. For further process research into source use and S-bursts, we therefore argue that it is sufficient to focus on only a few well-defined variables. When we compare the underlying activities, we observe that students perform differently in Dutch and English with respect to the source interaction component. More specifically, students consult a greater variety of sources per minute in English than in Dutch.

TABLE 8. Standardized parameter estimates of the final SEM model

	Dutch			English		
	Estimate	Std. all	sign.	Estimate	Std. all	sign.
Initial reading time	0.356	.123	.278	0.923	.185	.289
Source interaction	1.977	.261	**.005**	2.025	.129	.369
Variance source use	–0.476	–.053	.608	–0.734	–.053	.706

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

Regarding the link between text quality and source use, it was observed that there was no improvement in text quality during the academic year in Dutch (L1) and English (L2). However, text quality and source use prove to be correlated. Here the component of source interaction was influential again, albeit only in L1. In Dutch there was a relation between source interaction and quality, but no difference was found in English. Further exploration of the data on source interaction in Dutch showed that students tend to use more different sources, both categorized and noncategorized (see "Preparation and Analysis of Inputlog Data").

As shown in "Literature Review," the empirical research on source use in writing using a process perspective is rather limited, which makes it difficult to explain and contrast our findings using existing literature. However, we consider the three components that we found to be connected to research on theoretical constructs in the broader field of L1 and L2 writing process studies. Firstly, "initial reading" appears to be related to studies that focus on planning. Initial planning before writing leads to higher quality texts in L1, but might be attributed to longer time on task (Ellis & Yuan, 2004; Johnson, Mercado, & Acevedo, 2012). Furthermore, planning tends to increase the writing fluency (i.e., production rate per minute; De Smet et al. 2014; Limpo & Alves, 2018). In L2, preplanning has a beneficial effect on fluency (Ellis & Yuan, 2004; Johnson, Mercado, & Acevedo, 2012) and linguistic complexity (Ellis & Yuan, 2004). In our study, the writers were forced to think about the most optimal trade-off between initial reading of the sources, planning of how to use the sources and the time needed to complete the complex writing task. As theory suggests, initial planning is an important component in writing (Bereiter & Scardamalia, 1987; Hayes & Nash, 1996; Kellogg, 2008).

Secondly, all three components are related to the temporal distribution of the writing process. This in itself is not surprising because studies in the broader field of writing process research state the importance of the temporal distribution of the processes (Alves et al., 2008, 2015; Barkaoui, 2019; Roca de Larios, Manchón, Murphy, & Marin, 2008; Van Waes & Leijten, 2015; Xu & Qi, 2017). Although the temporal process of source use has not been studied empirically on large datasets to date, our study suggests that the interruption of the writing process caused by source use is a relevant factor comparable to interruptions caused by planning, translating, and revising. Beauvais, Olive, and Passerault (2011) state that the cognitive effort and distribution of writing processes is likely to influence the characteristics of final texts. We therefore consider it to be necessary to include these components in the understanding of writing processes: the integration of sources in L1 and L2 writing tasks will increase cognitive effort and writers need strategies to balance this load to improve their writing skills (cf. future research).

IMPLICATIONS

The findings of this study are of theoretical, methodological, and pedagogical significance. At the theoretical level we have noticed that the number and type of sources writers use are relevant to describe source-based writing. Furthermore, we have found that the use of sources in the beginning of the writing process is important. We can relate these findings to numerous planning studies in L1 and L2 (De Smet et al. 2014; Ellis & Yuan, 2004; Johnson et al., 2012; Limpo & Alves, 2018).

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

From a methodological perspective we have shown that it is important to carefully select the relevant variables as derived from the keystroke logging data collection. We also have listed the variables that reliably distinguish source use between students and that are very straightforward to incorporate in various writing process studies (also Baaijen & Galbraith, 2018; Van Waes et al., 2015). To enable researchers to carry out future research on source use more easily, we will also implement the proposed approach in Inputlog. Six out of the seven variables that provide insight into source use will be included in the automatic source analysis. The seventh variable, in which we grouped the various sources in six source categories, is a very laborious step in the process of analysis and could not be integrated in Inputlog yet. Therefore, we would like to formulate a cautious remark about the recoding possibilities of the source analysis. For this study, for example, the coding of the focus events was first precoded step by step in Microsoft Excel and subsequently the 3,104 focus events had to be grouped in Inputlog. This was a highly decisive step for the outcome of the study, and it is a step that still needs to be done manually by researchers.

For second language education, we advise teachers to raise students' awareness of the main components of source-based writing. Teachers should focus on the importance of deliberate fragmentation of the writing process, be it from a planning, pausing, revising, or source-use perspective. Doolan and Fitzsimmons-Doolan (2016) propose a sequential way to instruct source-based writing, as they argue that a shortcoming of much summary instruction is that it is exclusively centered on the summary as the end goal. They propose to include the techniques of summarizing and paraphrasing into larger written assignments, such as essays, for which the combination of independent and integrated writing is necessary. Such "sequencing" of the instruction of different writing strategies needs to be characterized by the interrelatedness of smaller and larger assignments, spread out over time.

LIMITATIONS AND FUTURE RESEARCH

We decided to approach the source use of master's students in this article from a technical process perspective (as compared to thinking aloud protocols by López-Serrano, Roca de Larios, & Manchón [2019] or stimulated recall by Révész, Michel & Lee [2019]). We are well aware that due to this approach, we do not know exactly what the students have read in the sources, nor how they have integrated the selected information into the text. In terms of both research methodology and quality approach, we can therefore certainly see some limitations and additional possibilities for further research. In the following text, we briefly point out some ideas.

An integrated writing task is by definition source based, implying that the provided information must first be understood before the new text can be created. As Cumming (2013) states, the problem with such tasks is that the writing performance is influenced not only by factors related to the writing process (including consultation and processing of sources, planning, and revision activities) but also by factors that play a role in the previous understanding of the source material (reading skills, prior knowledge of the subject, etc.). In follow-up research, it therefore seems useful to measure the students' reading skills but also their prior knowledge of the various topics, which were quite diverse in the current student sample. By testing these aspects, we can answer questions

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

such as: Do students read longer because they find it difficult to understand the sources, or is it a deliberate approach to record the information as accurately as possible? For L2, research by Plakans (2009) indeed suggests that, for L2 writers with a low proficiency, reading can sometimes hinder comprehension of the topic.

Further research is also needed to investigate how students best divide their time over the various activities. Relevant questions are whether it would be useful to have students focus on specific activities in specific intervals during the process, such as first reading of the sources (consulting dictionaries only for interpretation problems), then planning the text and in the final phase formulating the text (using sources for synonyms, findings equivalents in L2, etc.; Vangehuchten, Leijten, & Schrijver, 2018). As several recent studies on the role of the working memory in writing suggest (e.g., Zabihi, 2018 for L2; Medimorec & Risko, 2016 for L1), the cognitive load of a reading-to-write task makes it difficult to be proficient in all the required competences (content, accuracy, fluency) at once. Hence the pedagogical need to go about the instruction by parts.

We showed that there is a partial relationship between source interaction and quality: only in Dutch (L1) the text quality improved by more variation in source use. On the one hand, a lack of variation in the English sources might explain why no relation was found between source interaction and quality in English. Either repeatedly consulting the same sources or consulting a wide variety of sources can be expected to have different consequences for the richness of information, syntactic constructions and lexical choices in the final text. On the other hand, the absence of a relation between source interaction and quality can be caused by differences in the size of the L1 and L2 datasets. Overall, there was much individual variation in the data. In the smaller L2 dataset, individual variation in source use might have obscured effects of source use on text quality. This is why, in further research, we will strive to collect additional L2 data to further equalize the number of observations.

As mentioned in "Implications," this study has given several insights that may inform further research into source use in writing. First, the present study has demonstrated that the quantitative approach using keylogging is highly complementary to the more qualitative studies based on think-aloud protocols (see for example Chan, 2017; Lenski & Johns, 1997; Mateos et al., 2014; McGinley, 1992; Plakans, 2008). The added value of keylogging research is that for a large group of students very rich writing process data can be compared from different angles (e.g., source use, fluency, revisions). However, a qualitative approach, in which some students are asked to write the text in an eye-tracking lab, would certainly provide additional insights (e.g. approach of De Smet, Leijten, & Van Waes, 2018).

Second, a complementary approach aimed in particular at comparing similarities between the sources and the text might be pertinent for follow-up research. This could be done using so-called plagiarism software, which identifies identical passages and links them to the source material (such as TurnItIn or Scribbr; also advised by Nguyen & Buckingham, 2019). There are also computer algorithms for automated text comparison, which have been developed on the basis of so-called cosine similarity in the domain of journalistic science (see, e.g., Boumans & Trilling, 2016) and in computer linguistics (see, e.g., Kim & Crossley, 2018; Lin, Jiang, & Lee, 2014) and which provide richer analyses. Various studies have focused on the lexical similarities between source-based writing products and their related sources (Gebril & Plakans, 2016;

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

Kim & Crossley, 2018; Neumann et al., 2019; Plakans & Gebril, 2013). Writers do not only rely on source texts for content but also for language and text organizational support (Plakans & Gebril, 2012; Ye & Ren, 2019). These aspects of "textual borrowing" might not always be visible in the end product, but might be traceable in the writing process. Insights into these various perspectives on text similarities should therefore be combined with a process analysis of copy behavior: At what stage do students copy passages from the source texts and how do they edit these passages to integrate them into a coherent text? This would enable us to bring process and product insights together in a novel and enriching way (Cislaru, 2015; Leijten, Van Horenbeeck, & Van Waes, 2015, 2019).

Third, we would like to state the necessity to work toward a specific process-oriented pedagogical approach that will be monitored in an intervention study design. Source use turns out to be a very complex competence, which is not acquired by all students, not even at the end of a master's degree (Davis, 2013; McGinley, 1992). In this study, no improvement was observed for none of the examined languages (L1 and L2) between the two measurement points in terms of use of sources. Of course, this does not hold for each individual student. Some students do improve in the writing task at hand and have been able to integrate the knowledge. Therefore, we assume that certain writer characteristics, for example, learning style and self-efficacy are of importance in improving writing skills. This relates to the major findings in the studies by Baaijen, Galbraith, and De Glopper (2014), Baaijen and Galbraith (2018), and Limpo and Alves (2017), focusing on writing beliefs, the mediating role of goal achievement, and self-efficacy, respectively. For instance, Baaijen and her colleagues set up a planning task and distinguished transactional and transmissional beliefs. They found that "transactional beliefs are about the preference for a top-down strategy or a bottom-up strategy, while transmissional beliefs are about the content that is written about. These beliefs interact in their effects on text quality, the number and type of revisions carried out, and the extent to which writers develop their understanding" (2014, p. 81). Limpo and Alves (2017, p. 97) also found a diverging effect in their study of personal characteristics (i.e., goal achievement and self-efficacy) when modeling the relation between writing strategies and writing performance. These findings highlight the importance of accounting for writers' characteristics when studying the effect of writing (sub)processes or writing strategies on writing quality. We therefore think that adding personal characteristics as mediating variables to the modeling procedure is crucial in follow-up research because different writing attitudes and writing profiles might benefit from different source-use strategies (at different stages in the writing process of L1 and L2 writers).

Overall, this first quantitative study into source use in writing has provided a number of both theoretical and methodological insights. With regard to source-based writing processes in L1 and L2, it has shown three components to be relevant to describe source use by master's students. The study has also revealed that source use and text quality are correlated in L1. These findings open up a number of interesting research prospects regarding source-based writing processes study, for example the impact of linguistic proficiency in different L2s on source use or the rationale for choosing a specific source-use strategy.

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S0272263119000251

NOTES

[1]English is regarded in the present study as the students' L2, although English is, technically, our participants' third language: they learned the language from the age of 12, after they started learning French. However, given the omnipresence of English in Flemish popular media (e.g., Anglo-American films and TV series are subtitled, not dubbed), we can state that all of our students came into contact with English before the age of 10 in a receptive way and, for the majority, also in a productive way, e.g., during online gaming activities, on holidays. Hence, English is regarded as their L2.

[2]www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions.

[3]No students were excluded from the test based on their typing skills.

[4]The pairwise comparison data were modeled according to the Bradley–Terry–Luce model, which results in a continue variable in Z-scores (Van Daal et al., 2017).

REFERENCES

Alves, R. A., & Limpo, T. (2015). Progress in written language bursts, pauses, transcription, and written composition across schooling. *Scientific Studies of Reading*, *19*, 374–391. doi: 10.1080/10888438.2015.1059838.

Alves, R. A., Castro, S. L., & Olive, T. (2008). Execution and pauses in writing narratives: Processing time, cognitive effort and typing skill. *International Journal of Psychology*, *43*, 969–979.

Baaijen, V. M., & Galbraith, D. (2018). Discovery through writing: Relationships with writing processes and text quality. *Cognition and Instruction*, *36*(3), 199–233.

Baaijen, V. M., Galbraith, D., & de Glopper, K. (2014). Effects of writing beliefs and planning on writing performance. *Learning and Instruction*, *33*, 81–91. doi: 10.1016/j.learninstruc.2014.04.001.

Barkaoui, K. (2019). What can L2 writers' pausing behavior tell us about their L2 writing processes?. *Studies in Second Language Acquisition*, *41*, 529–554.

Beauvais, C., Olive, T., & Passerault, J.-M. (2011). Why are some texts good and others not? Relationship between text quality and management of the writing processes. *Journal of Educational Psychology*, *103*, 415–428. doi: 10.1037/a0022545.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars. *Digital Journalism*, *4*, 8–23.

Bowles, M. A. (2010). *The think-aloud controversy in second language research*. New York, NY, and London, UK: Routledge.

Byrnes, H., & Manchón, R. M. (Eds.), (2014). *Text-based language learning—Insights from and for L2 writing*. Amsterdam:John Benjamins.

Cerdán, R., & Vidal-Abarca, E. (2008). The effects of tasks on integrating information from multiple documents. *Journal of Educational Psychology*, *100*, 209–222.

Chan, S. (2017). Using keystroke logging to understand writers' processes on a reading-into-writing test. *Language Testing in Asia*, *7*, 10. doi: 10.1186/s40468-017-0040-5.

Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, *18*, 80–98. doi: 10.1177/0741088301018001004

Chenoweth, N. A., & Hayes, J. R. (2003). The inner voice in writing. *Written Communication*, *20*, 99–118.

Chukharev-Hudilainen, E., Saricaoglu, A., Torrance, M., & Feng, H.-H. (2019). Combined deployable keystroke logging and eyetracking for investigating L2 writing fluency. *Studies in Second Language Acquisition*, *41*, 583–604.

Cislaru, G. (Ed.), (2015). *Writing(s) at the crossroads: The process-product interface*. Amsterdam, The Netherlands: John Benjamins.

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing*, *18*, 119–135. doi:10.1016/j.jslw.2009.02.002.

Cuevas, I., Mateos, M., Martín, E., Luna, M., Martín, A., Solari, M, González-Lamas, J., & Martinez, I. (2016). Collaborative writing of an argumentative synthesis from multiple sources: The role of writing beliefs and strategies to deal with controversy. *Journal of Writing Research*, *8*, 205–226. doi:10.17239/jowr-2016.08.02.02.

Cumming, A. (2013). Assessing integrated writing tasks for academic purposes: Promises and perils. *Language Assessment Quarterly*, *10*, 1–8. doi:10.1080/15434303.2011.622016.

Cumming, A., Lai, C., & Cho, H. (2016). Students' writing from sources for academic purposes: A synthesis of recent research. *Journal of English for Academic Purposes*, *23*, 47–58. doi:10.1016/j.jeap.2016.06.002.

Davis, M. (2013). The development of source use by international postgraduate students. *Journal of English for Academic Purposes*, *12*, 125–135. doi: http://dx.doi.org/10.1016/j.jeap.2012.11.008.

De Smet, M. J., Brand-Gruwel, S., Leijten, M., & Kirschner, P. A. (2014). Electronic outlining as a writing strategy: Effects on students' writing products, mental effort and writing process. *Computers & Education*, *78*, 352–366.

De Smet, M. J. R., Leijten, M., & Van Waes, L. (2018). Exploring the process of reading during writing using eye tracking and keystroke logging. *Written Communication*, *35*, 411–447. doi:10.1177/0741088318788070.

Doolan, S. M., & Fitzsimmons-Doolan, S. (2016). Facilitating L2 Writers' Interpretation of Source Texts. *TESOL Journal*, *7*, 716–745. doi:10.1002/tesj.239.

Ellis, R. (2009). Task-based research and language pedagogy. In K. Van den Branden, M. Bygate, & J. Norris (Eds.), *Task-based language teaching: A reader* (pp. 109–130). Amsterdam, The Netherlands, and Philadelphia, PA: John Benjamins and others.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, *26*, 59–84.

Gebril, A., & Plakans, L. (2016). Source-based tasks in academic writing assessment: Lexical diversity, textual borrowing and proficiency. *Journal of English for Academic Purposes*, *24*, 78–88. doi:10.1016/j.jeap.2016.10.001.

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

Hayes, J. R., & Chenoweth, N. A. (2006). Is working memory involved in the transcribing and editing of texts? *Written Communication*, *23*, 135–141.

Hayes, J. R., & Nash, J. G. (1996). On the nature of planning in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 29–55). Mahwah, NJ: Lawrence Erlbaum Associates.

Hinkel, E. L. I. (2003). Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly*, *37*, 275–301. doi:10.2307/3588505.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*, 53–60.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55. doi:10.1080/10705519909540118.

Johansson, R., Wengelin, Å., Johansson, V., & Holmqvist, K. (2010). Looking at the keyboard or the monitor: Relationship with text production processes. *Reading and Writing*, *23*, 835–851. doi:10.1007/s11145-009-9189-3.

Johnson, M. D., Mercado, L., & Acevedo, A. (2012). The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. *Journal of Second Language Writing*, *21*, 264–282. doi: http://dx.doi.org/10.1016/j.jslw.2012.05.011.

Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, *1*, 1–26. doi:10.17239/jowr-2008.01.01.1.

Kobayashi, H., & Rinnert, C. (2012). Understanding L2 writing development from a multicompetence perspective: Dynamic repertoires of knowledge and text construction. In R. M. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 101–134). Berlin, Germany: De Gruyter Mouton.

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

Kim, M., & Crossley, S. A. (2018). Modeling second language writing quality: A structural equation investigation of lexical, syntactic, and cohesive features in source-based and independent writing. *Assessing Writing*, *37*, 39–56. doi:10.1016/j.asw.2018.03.002.

Leijten, M., Van Horenbeeck, E., & Van Waes, L. (2015). Analyzing writing process data: A linguistic perspective. In G. Cislaru (Ed.), *Writing(s) at the crossroads: The process-product interface* (pp. 277–302). Amsterdam, The Netherlands, and Philadelphia, PA: John Benjamins Publishing Company.

Leijten, M., Van Horenbeeck, E., & Van Waes, L. (2019). Analyzing keystroke logging data from a linguistic perspective. In K. Sullivan & E. Lindgren (Eds.), *Observing writing: Insights from keystroke logging and handwriting* (pp. 71–95). Amsterdam, The Netherlands: Brill.

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

Leijten, M., Van Waes, L., Schriver, K., & Hayes, J. R. (2014). Writing in the workplace: Constructing documents using multiple digital sources. *Journal of Writing Research*, *5*, 285–337.

Leijten, M., Van Waes, L., Schrijver, I., Bernolet, S., & Vangehuchten, L. (2017). Hoe schrijven masterstudenten syntheseteksten? Het brongebruik van gevorderde schrijvers in kaart gebracht. *Pedagogische Studiën*, *94*, 233–253.

Lenski, S. D., & Johns, J. L. (1997). Patterns of reading-to-write. *Reading Research and Instruction*, *37*, 15–38. doi:10.1080/19388079709558252.

Lindgren, E., Leijten, M., & Van Waes, L. (2011). Adapting to the reader during writing. *Written Language and Literacy*, *14*, 188–223. doi: 10.1075/wll.14.2.02lin.

Limpo, T., & Alves, R. A. (2017). Written language bursts mediate the relationship between transcription skills and writing performance. *Written Communication*, *34*, 306–332. doi:10.1177/0741088317714234.

Limpo, T., & Alves, R. A. (2018). Effects of planning strategies on writing dynamics and final texts. *Acta Psychologica*, *188*, 97–109. doi: https://doi.org/10.1016/j.actpsy.2018.06.001.

Lin, Y.-S., Jiang, J.-Y., & Lee, S.-J. (2014). A similarity measure for text classification and clustering. *IEEE Transactions on Knowledge and Data Engineering*, *26*, 1575–1590.

Lindgren, E., & Sullivan, K. (Eds.), (2019). *Observing writing: Insights from keystroke logging and handwriting* (Vol. *38*). Leiden, The Netherlands: Brill.

Liu, G.-Z., Lin, V., Kou, X., & Wang, H.-Y. (2016). Best practices in L2 English source use pedagogy: A thematic review and synthesis of empirical studies. *Educational Research Review*, *19*, 36–57. doi: http://dx.doi.org/10.1016/j.edurev.2016.06.002.

López-Serrano, S., Roca de Larios, J., & Manchón, R. M. (2019). Language reflection fostered by individual L2 writing tasks: Developing a theoretically motivated and empirically based coding system. *Studies in Second Language Acquisition*, *41*(3), 503–527.

Martínez, I., Mateos, M., Martín, E., & Rijlaarsdam, G. (2015). Learning history by composing synthesis texts: Effects of an instructional programme on learning, reading and writing processes, and text quality. *Journal of Writing Research*, *7*, 275–302. doi:10.17239/jowr-2015.07.02.03.

Marzec-Stawiarska, M. (2016). The influence of summary writing on the development of reading skills in a foreign language. *System*, *59*, 90–99. doi:10.1016/j.system.2016.04.006.

Mateos, M., Solé, I., Martín, E., Cuevas, I., Miras, M., & Castells, N. (2014). Writing a synthesis from multiple sources as a learning activity. In *Writing as a learning activity* (pp. 169–190). Leiden: Brill. doi:10.1163/9789004265011_009.

McGinley, W. (1992). The role of reading and writing while composing from sources. *Reading Research Quarterly*, *27*, 226–248.

Medimorec, S., & Risko, E. F. (2016). Effects of disfluency in writing. *British Journal of Psychology*, *107*, 625–650. doi:10.1111/bjop.12177.

Neumann, H., Leu, S., & McDonough, K. (2019). L2 writers' use of outside sources and the related challenges. *Journal of English for Academic Purposes*, *38*, 106–120. doi:10.1016/j.jeap.2019.02.002.

Nguyen, Q., & Buckingham, L. (2019). Source-use expectations in assignments: The perceptions and practices of Vietnamese master's students. *English for Specific Purposes*, *53*, 90–103. doi:10.1016/j.esp.2018.10.001.

Plakans, L. (2008). Comparing composing processes in writing-only and reading-to-write test tasks. *Assessing Writing*, *13*, 111–129. doi:10.1016/j.asw.2008.07.001.

Plakans, L. (2009). The role of reading strategies in integrated L2 writing tasks. *Journal of English for Academic Purposes*, *8*, 252–266. doi:10.1016/j.jeap.2009.05.001.

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

Plakans, L., & Gebril, A. (2012). A close investigation into source use in integrated second language writing tasks. *Assessing Writing*, *17*, 18–34. doi:10.1016/j.asw.2011.09.002.

Plakans, L., & Gebril, A. (2013). Using multiple texts in an integrated writing assessment: Source text use as a predictor of score. *Journal of Second Language Writing*, *22*, 217–230. doi:10.1016/j.jslw.2013.02.003.

Pollitt, A. (2012). Comparative judgement for assessment. *International Journal of Technology and Design Education*, *22*, 157–170. doi:10.1007/s10798-011-9189-x.

Raedts, M. (2008). *De invloed van zelfeffectiviteitsverwachtingen, taakkennis en observerend lezen bij een nieuwe en complexe schrijftaak [The influence of self-efficacy, task knowledge, and observational learning on a new and complex writing task]* (Unpublished doctoral dissertation). Universiteit Antwerpen , Antwerp, Belgium.

Révész, A., Kourtali, N.-E., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and Linguistic complexity. *Language Learning*, *67*, 208–241. doi:10.1111/lang.12205.

Révész, A., Michel, M., & Lee, M. (2019). Exploring second language writers' pausing and revision behaviors: A mixed-methods study. *Studies in Second Language Acquisition*, *41*(3), 605–631.

Roca de Larios, J., Manchón, R., Murphy, L., & Marín, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing*, *17*, 30–47. doi: https://doi.org/10.1016/j.jslw.2007.08.005.

Schriver, K. (2012). What we know about expertise in professional communication. In V. W. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 275–312). New York, NY: Psychology Press.

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. doi:10.1023/B:ESLL.0000033847.00732.af.

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. doi:10.1177/0741088312466532.

Spivey, N. N., & Nelson, N. (1997). *The constructivist metaphor: Reading, writing, and the making of meaning*. San Diego: Academic Press.

Van Daal, T., Lesterhuis, M., Coertjens, L., van de Kamp, M.-T., Donche, V., & De Maeyer, S. (2017). The complexity of assessing student work using comparative judgment: The moderating role of decision accuracy. *Frontiers in Education*, *2*, 1–44 doi:10.3389/feduc.2017.00044.

Vangehuchten, L., Leijten, M., & Schrijver, I. (2018). Reading-to-write tasks for professional purposes in Spanish as a foreign language. *Revista Espanola De Linguistica Aplicabe [Spanish Journal of Applied Linguistics]*, *31*, 638–659. doi.org/10.1075/resla.31.2.

Van Waes, L., & Leijten, M. (2015). Fluency in writing: A multidimensional perspective on writing fluency applied to L1 and L2. *Computers and Composition*, *38*, 79–95. doi: 10.1016/j.compcom.2015.09.012.

Van Waes, L., Leijten, M., Lindgren, E., & Wengelin, A. (2015). Keystroke logging in writing research: Analyzing online writing processes. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 410–427). New York, NY, and London, UK: The Guilford Press.

Van Waes, L., Leijten, M., Mariën, P., & Engelborghs, S. (2017). Typing competencies in Alzheimer's disease: An exploration of copy tasks. *Computers in Human Behavior*, *73*, 311–319. doi: 10.1016/j.chb.2017.03.050.

Xu, C., & Qi, Y. (2017). Analyzing pauses in computer-assisted EFL writing—A computer keystroke-log perspective. *Educational Technology and Society*, *20*, 24–34.

Ye, W., & Ren, W. (2019). Source use in the story continuation writing task. *Assessing Writing*, *39*, 39–49. doi:10.1016/j.asw.2018.12.001.

Zabihi, R. (2018). The role of cognitive and affective factors in measures of L2 writing. *Written Communication*, *35*, 32–57. doi:10.1177/0741088317735836.

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press

APPENDIX A

FULL PROMT OF THE WRITING TASK

Task

Write a concise and coherent synthesis text on the basis of three source texts:

- An excerpt from a European Union report
- A text from the European Union website
- A newspaper article from *De Standaard*/*The Independent*/*Le Soir*/*El Pais*/*Süddeutsche Zeitung*

The text should count minimally 200 and maximally 250 words.

Context

During this session, you write two texts for the teachers of Dutch/English/French/Spanish/German of the Royal Lyceum in Antwerp. The texts are explicitly aimed at an audience of 17- to 18-year-old students in the last year of secondary education (in the general education system). The teachers of the different languages want to adopt your text in a special edition of the high school newspaper. Your synthesis text should be a stand-alone text, such that it can be understood without the background information from the three source texts.

Sources

You are allowed to make use of additional sources, such as:

- Online dictionaries
- The Internet (except for social media and e-mail)

Pay attention: do *not* open two windows alongside each other (e.g., Word document and a source or the Internet).

https://doi.org/10.1017/S0272263119000251 Published online by Cambridge University Press