

Response processes in L2 writing tasks with Internet access

Dea Jespersen, University of Copenhagen
Slobodanka Dimova, University of Copenhagen

Technology has changed modern L2 written communication in many ways, but how these changes have affected our understanding of the L2 writing construct needs further investigation (Weigle, 2002). Given that the Internet provides access to numerous resources available to L2 writers, the Danish Ministry of Education conducted pilots to modernize the school-leaving exams by including an L2 writing assessment in French with Internet access (DAMVAD, 2013). This study is guided by questions related to (1) differences in students' writing performance with Internet access (IA) and without Internet access (WIA), (2) students' writing behavior when they have IA or WIA, and (3) students' perceptions of the writing assessment with IA. Using a balanced design, two writing tasks in a WIA and an IA version were administered to ninth-grade L2 learners of French (N=32). Scores, window tracker logs, and a student survey were used in the analysis. Results suggested that while students strongly preferred the IA tasks, the task format (IA or WIA) did not affect their scores. The students did not use online resources beyond dictionary and conjugation sites, for either the IA or the WIA task.

Keywords: Internet-based writing, digital literacy, assessment, writing from sources

1 Introduction

Information and communication technologies (ICT) have grown into a staple of obligatory education in a number of countries around the world despite the discrepancies observed among the demographic groups in some countries. Therefore, digital literacy, i.e. the ability to use the Internet for information collection and writing, has become an important skill to master (Dudeny, Gavin & Hockly, 2016). In many foreign language (L2) writing assessment contexts, however, test-takers are still required to write offline although the Internet is a common writing resource in school and other real-life situations. To overcome such a contradictory approach to teaching and assessment of L2 writing and to make exams more authentic, the Danish Ministry of Education has considered the inclusion of Internet access (IA) in L2 writing (English, French, and German). The assumption is that the Internet will bring the L2 writing exam closer to real-life communication and draw the attention to ICT in the foreign language classroom,

Corresponding author's email: slobodanka.dimova@hum.ku.dk

eISSN: 1457-9863

Publisher: University of Jyväskylä, Language Campus

© 2021: The authors

<https://apples.journal.fi>

<https://doi.org/10.47862/apples.99431>

but little is known about how to design L2 writing tasks that involve authentic uses of the Internet.

Despite the recognition of the importance of understanding the interaction between language and technology and defining the appropriate language uses in specific digital environments (Chapelle & Douglas, 2006; Douglas & Hegelheimer, 2007), a lack of empirical evidence concerning the role of digital literacies exists, and many conclusions are drawn solely on speculations (Bulger et al., 2014, p. 1582). For that reason, the purpose of the current study was to explore young test-takers' L2 response processes, i.e. their writing behavior, performance, and perception of L2 writing assessment where Internet access is part of the task. The study could help us improve our understanding of the expanded L2 writing construct, i.e. a construct including elements of digital literacy because it investigates 1) the differences in students' writing performance with and without Internet access, 2) the difference in students' writing behavior with and without Internet access, and 3) students' perceptions of the writing assessment with and without Internet access.

2 Literature review

Computer technology in L2 assessment was introduced in the mid 1980's to improve test practicality, particularly in terms of test administration, rating, data storage, and result reporting (García Laborda, 2007). Since then, computer and web-based testing has gained popularity because it is easier and cheaper to distribute on a large scale (Dooey, 2008; Jin & Yan, 2017). Discussions about design of computer and web-based items underline the importance of appropriate hardware and software for item development, as well as decisions about the computer interface, ease of navigation, page layout, and textual and visual representation on the screen (Fulcher, 2003). According to Roever (2001), the Internet can provide a highly authentic environment for language testing if the tasks involve use of the Internet (e.g., writing emails or searching for information). Moreover, the design of discrete-point items, cloze tests, C-tests, essays, listening and reading comprehension items can be facilitated by the Internet because they do not require advanced programming and are easier to administer. Depending on the assessment needs, various innovative designs can be applied to develop items that are contextualized with the use of online visual input (pictures and video) and sound files (Dimova et al., 2020). However, issues with browser incompatibility, test security, server failure, and data storage may be difficult to control for and resolve.

While the mode of delivery quickly changed from paper to computer, the measured constructs were assumed to remain the same (Binkley et al., 2012; Douglas & Hegelheimer, 2007). Although generally L2 students tend to have positive attitudes towards technology in teaching and assessment (Golonka et al., 2014; Stricker & Attali, 2010; Wang & Vasquez, 2012), concerns were raised about test validity because test digitalization inevitably changed the measured language construct by introducing additional parameters, such as digital (or computer) literacy. In other words, uncertainty exists in terms of how to decide to what degree test takers' digital literacy level or their language proficiency levels affect their test performance.

Digital literacy is defined as “a person’s ability to perform tasks effectively in a digital environment” (Jones-Kavalier & Flannigan, 2006, p. 9). Digital literacy is not just the ability to use computers or software, but it is also the ability to find, understand, evaluate, and use information available in multiple formats via computers, tablets, and smartphones (Buckingham, 2006; Jones-Kavalier & Flannigan, 2006). Van Deursen and van Dijk (2009) operationalized digital literacy as a set of different skills: operational (using devices’ hardware and software), formal (accessing networks and web environments), and informational (finding, selection, evaluation, and process information). As part of information, Ananiadou and Claro (2009) also added the ability to restructure and transform information to develop knowledge, i.e. information as a product (p. 20).

Lack of digital literacy could result in differential test performance and hence construct irrelevant variance (Chapelle & Douglas, 2006; Jin & Yan, 2017; Roeber, 2001). In L1 and L2 writing assessment, most research has focused on the effects of computers as a test delivery mode. Based on our review of previous literature, little research exists on how the Internet as a resource affects writing test performance with younger learners as studies tend to focus on writing in higher education. A meta-analysis of fifteen articles published between 1992 and 2002 that compared the quality of writing on paper and computer showed a clear positive effect of the computer on the quality of the participants’ writing (Goldberg et al., 2003). In a more recent study, Laurie, Bridglall and Arseneault (2015) found that the different modes of delivery (computer- or paper-based) of an L1 writing test administered to 302 primary-school students had no effect on their test scores but had an effect on different writing traits (e.g., syntax, punctuation, ideas, and spelling). Syntax, punctuation, and ideas were better in the students’ pen and paper essays whereas spelling was best in the computer-written versions. In a similar manner, Jin and Yan (2017) found that despite the facilitative effect of test-takers’ computer familiarity, the test format (computer-based versus pen-and-paper) had no effect on scores, text complexity, number of language errors, and writing processes in an English L2 writing test administered to Chinese test-takers (n=116).

In L2 writing, if the writers have relevant digital literacy, the digital environment could facilitate their writing process because it could help reduce some of the cognitive load L2 students experience (Stapleton, 2010, 2012). To compensate for a breakdown in written communication, computer technology could provide tools that bridge linguistic or topical information gaps (Kozlova & Presas, 2013). This means that writers do not need to rely solely on their knowledge but can make use of the web as their external memory (Sparrow et al., 2011). L2 writers can transform a copied text by using the thesaurus in the word processor (Stapleton, 2012), use spelling and grammar checks, borrow formulaic phrases (Pecorari & Petrić, 2014), and use search engines, such as Google, as corpora and concordancer (Geluso, 2011).

Teachers often believe that the use of these applications and machine translation programs, such as Google Translate, equals cheating although these programs are found to assist communication for L2 beginners (Garcia & Pena, 2011). Similarly, questions of authorship are real threats that can compromise test security (Chapelle & Douglas, 2006), and it is believed that Internet access causes increased plagiarism (Evering & Moorman, 2012a; Pecorari & Petrić, 2014). Internet connection gives students the possibility to access social platforms and/or copy large amounts of text. Cheating by sharing can be enhanced as it is

difficult to monitor test-takers' accesses to social platforms. Despite these validity threats, test-takers are more often warned against plagiarism than trained how to avoid it (Evering & Moorman, 2012; Pecorari & Petrić, 2014). Furthermore, it is often unclear whether borrowing of formulaic phrases should be frowned upon (Pecorari & Shaw, 2012), especially at a lower level as Cumming et al. (2005) found that borrowing in integrated reading-writing tasks happened more often than at more advanced levels.

Others argue that information problem solving with digital tools, also called information literacy, can be time-consuming and take up a lot of cognitive resources because test-takers need to "Recognize when information is needed and have the ability to locate, evaluate, and use effectively the needed information" (American Library Association, 2000). This is a complex skill that is different from the traditional reading to write skill (Coiro & Dobler, 2007). For young learners, searching can take up a lot of time with unsuccessful or irrelevant searches (Zhang et al., 2009). Digital writers need skills and knowledge to locate the information, judge how quickly it could be accessed, decode the information and judge its relevance, and then go back to the writing task and incorporate the information in the text to be written (Leijten et al., 2014, p. 326).

Leijten et al. (2014) argued that earlier writing models needed an update to include changes technologies bring in the cognitive processes involved in writing. In a case study based on keylogging and interviews with a professional communicator, they showed that expert writing included thorough searches for inspiration in multiple sources, construction of verbal and visual contents, and management of attention and motivation. This led to additions to Hayes' 2012 model of writing and a proposition of a model of the information search process. Focusing on traditional writing, Hayes' (2012) writing model helps to "understand writing as the interaction among subprocesses, each of which does part of the writing job but not the whole job" (p. 375). Leijten et al. (2014) suggested inclusion of subprocesses that account for the search process, the motivation management, and construction of verbal and visual content (design schemas) (pp. 324-326). Bulger et al. (2014) also used keylogging of student activity during a timed writing test with access to online sources to investigate the role of digital literacy in the writing process. They defined digital literacy as "the ability to read and write using online sources, and includes the ability to select sources relevant to the task, synthesize information into a coherent message, and communicate the message with an audience" (p. 1567). They found that academic experience and ability to integrate sources into the written text were more likely to predict digital literacy than technical knowledge or the process of accessing many sources.

The writing expected from the beginner L2 student in a digital environment is, however, more likely to be based on knowledge telling rather than what Scardamalia and Bereiter (cit. Weigle, 2002, pp. 28-35) call knowledge transformation with sources. The cognitive processes involved in beginner L2 writing are mainly characterized by word-to-word retrieval (Weigle, 2002, p. 36) since the vocabulary is limited and access to it is not automated yet. Most of the working memory is dedicated to translating ideas into the L2, leaving little time for planning and revision. Since attention resources are limited, low proficiency writers spend less time on the macro-processes such as structure of content or revising (Roca de Larios et al., 2008). The writing is slower and more fragmented because information gaps at the macro (ideas of contents) and micro levels

(translate an idea into a word) appear more often than in more expert writers (Manchón et al., 2007). In other words, traditional writing research suggests that the L2 beginner writer's cognitive resources are scarce because these writers principally focus on communicating the message. In language assessment, however, the cognitive resource allocations during writing tasks must be analyzed to gain insight into the fit between the construct and the actual performance, and therefore the inferential score validity (American Educational Research Association et al., 2014; Zumbo & Hubley, 2017).

In his review of computer-assisted testing, García Laborda (2007) called for rethinking the language constructs in technologically-mediated environments by reconsidering how to develop items germane not only to computer-assisted but also Internet-based tests (p. 8). He also argued that the shift to Internet-based language testing would inevitably have consequences on language instruction styles, test preparation activities, and enjoyment. While studies investigating washback effects on teaching and test preparation activities are scarce, a number of studies focus on test-takers' perception of the use of technology in language testing.

3 Context

The majority of students in Danish schools grow up in a highly digitalized society. A 2020 survey (Danmarks Statistik, 2020), representative of the population in Denmark, revealed that nearly all families in Denmark had Internet connection at home (97%), a number close to households in Norway and Sweden, and above the European average. Ninety-three percent of young people in Denmark (16-24 year-olds) access the Internet daily. Most lower secondary school students are used to accessing the Internet on different types of devices in their private life and at school. In 2013, there was one computer per 4,9 students in schools (Børns Vilkår, 2019, p. 50). In 2018, 97% of students reported that they use a computer or a tablet often or very often to solve tasks in class (Børne- og Undervisningsministeriet, 2018, p. 24).

Following this trend, the Danish Ministry of Education is expecting all schools to offer obligatory education to provide a "well-functioning IT-infrastructure for students" (Undervisningsministeriet, 2017a). For that reason, all public schools have a wireless network (Digitaliseringsstyrelsen, 2016), digital devices are essential part of the daily instruction in Danish schools, and teachers are expected to include ICT their teaching as an integrated part of the curricula (European Commission/EACEA/Eurydice, 2019). To address the inequality that digitalization of education could bring, schools offer free access to digital devices and software, and students can choose to bring their own device (BYOD, bring your own device) or borrow one from the school for the entire school year regardless of their social background (Børne- og Undervisningsministeriet, 2019; Koldborg & Pam, 2014). The free access to digital devices in schools however may not be enough to fight inequality in digital competences. In a report by Children's Welfare (Børns Vilkår, 2019), 52% of school principals and 37% of teachers express concerns that students who use their own devices have better prerequisites to use digital media because they tend to come from families with greater resources and interest in digital information resources (p. 50). For example, these families are likely to invest in devices that are more up to date than the ones provided by the schools (p. 50). Therefore, in the Danish context, students need to develop

informational digital literacy, i.e. confident, critical and creative uses (Redecker & Punie, 2017, p. 90), rather than operational digital literacy, i.e. how to operate the digital devices (van Deursen & van Dijk, 2009).

Given the presence of ICT across the curriculum, introduction of digital exams with Internet access was proposed in 2009, first as part of the obligatory school-leaving exam in Danish (Christensen, 2017) and then as part of the exams in the three taught foreign languages, English, German, and French. The new exam format was presumed to elicit L2 writing behavior that is closer to real-life writing experiences, which was expected to motivate students (Danmarks Evalueringsinstitut, 2016; Warnich, 2013). To guarantee a higher level alignment between the curriculum, teaching practices, and language assessment, the integration of digital literacy to language assessment makes sense. The introduction of the new format is expected to have positive washback in the classroom in that teachers will implement similar writing activities to prepare students for the exam.

3.1 School-leaving exams: L2 writing section in French

Students take obligatory school-leaving exams at the end of grade 9, when they are about 15 years old. The present study focuses on the writing section of the French school-leaving exam. As part of the written section of the exam, the students have two hours to complete a series of discrete point items related to language use and to write a 100 to 150 word response to each of the two writing tasks. During the performance-based part of the exam, students have access to digital and/or paper aids (dictionaries, grammar overviews, conjugation tools, as well as different correction and helping tools available in their word processor of choice). The written production is scored holistically with the use of an assessment grid.

The L2 school-leaving exams used to be relatively low-stakes as the results had no consequences for students' future education or work. However, this changed in 2019, when final exam grades started playing a role in high school admission (Undervisningsministeriet, 2017b). As the stakes have risen, ensuring the validity of the test has become an imperative.

Previous official validity analyses of the L2 exam pilots with IA center only on face validity, as they are primarily grounded in teachers', students', and raters' perceptions and impressions. In the report on the L2 with IA exam trials in the 2015-2016 school year (Danmarks Evalueringsinstitut, 2016), 84% of the teachers (n=92) whose classes took part in the test trials and the students who participated in focus group interviews (n=9) responded positively to the new mode of examination. Also, 44% of teachers (n=16) reported that they thought Internet access improved their students' writing performance. In the interviews, teachers of English (n=4), external raters of English, French, or German (N=4), and some students who took the English exam (n=9) had divided opinions about the impact of the new exam mode on the quality of responses. While one of the teachers and the interviewed students noted that exam responses seemed more fluent because students used complex vocabulary and found inspiration for topical knowledge, an external rater thought that Internet access did not affect scores. Raters and teachers argued that the exam with Internet access tested two different sets of skills: writing skills and Internet skills, i.e. skills in finding and using online information.

4 Method

The purpose of the current study was to investigate the effect of the inclusion of Internet access on an L2 writing test administered to 9th graders in Denmark. Given the students' experience with computers and the Internet as part of their schooling, the study did not focus on whether the operational or the formal skills of digital literacy would affect students' L2 writing responses, as defined by van Deursen and van Dijk (2009). The focus was rather on the degree to which the students would apply their digital informational skills to solve the two writing tasks and whether this would influence their performances. Therefore, the use of the Internet and other sources (dictionary, conjugation table, class materials) and the product (scores and responses) were examined. The research was guided by the following research questions:

- 1) Are there any differences in students' writing performance with and without Internet access?
- 2) Are there any differences in students' writing behavior when they write with and without Internet access?
- 3) What are students' perceptions of the writing assessment with and without Internet access?

4.1 Participants

Forty-nine 9th grade students took the two pilot writing items of the French foreign language school-leaving exam. Due to problems ranging from technical issues (incomplete log files, lost responses) to cheating (one student used Google Translate extensively, another copied longer excerpts of a textbook verbatim), only 32 responses out of 49 were used. Recruitment of participants was rather difficult as in many cases only one class (10-15 students in 9th grade) studied French. The students were L1 speakers of Danish in their third year of studying French as their second foreign language. Thus, they were expected to have an elementary level of French or A2 according to the Common European Framework of Reference for languages (CEFR) (Council of Europe, 2001). All the students had English as their first foreign language. This study was conducted only a month before the official French exam, which means that the students' writing proficiency would not change much before the official exam. The participants were from different classes, taught by different French teachers.

4.2 Data

Three types of data were collected: task data (holistic scores and student responses to two tasks from the French writing exam), log files from a window tracker, and a student survey. The written responses were used to compare students' writing performance with and without Internet, while the window tracker allowed us to follow the use of sources during writing. The survey was administered to examine students' perceptions of the writing tasks with and without the Internet.

4.2.1 Task data

Two writing tasks (Task 1 and Task 2) were administered under two conditions (IA=with Internet access; and WIA=without Internet access). In both conditions, students typed their answers on a computer. A balanced design was used, where

Group 1 (n=15) had Task 1 (IA) and Task 2 WIA, and Group 2 (n=17) had Task 1 (WIA) and Task 2 (IA).

In both Tasks 1 and 2, the students had to imagine they were tourists in Paris. In Task 1, they were given information about two different restaurants in Paris. They were instructed to choose one of the restaurants and write an email to a friend explaining which restaurant they selected and why. Those who had Task 1 (IA) could access the restaurant websites through hyperlinks, while those who had Task 1 (WIA) received the restaurant descriptions (menus, pictures, contact details, history) in a booklet. Task 2 required the students to write an email to a friend where they plan an afternoon in Paris by describing what the students would recommend to visit. Those who had Task 2 (IA) could search the Internet for information, while those who had Task 2 (WIA) referred to their class material, which was sufficient to complete the task.

The prompts were in French, but the students could use a mouse-over function on the website to get the Danish translation. In the WIA version, the instructions were given both in French and Danish. In order to control for all variables except for Internet access, in both test conditions, all students used the school computers and wrote in LibreOffice documents that did not have a French spelling checker or thesaurus. In accordance with the standard test procedures, the students had access to an electronic Danish-French bilingual dictionary and a conjugation webpage throughout the test. With IA, the students could use all webpages, as long as these did not enable them to share or receive information (social platforms) or translate whole chunks of text for them (machine translation websites). This was monitored with the window tracker (see section 4.2.2.). The students were required to write 125-150 words per task.

The students had two hours to complete the two tasks, just as in the regular exam situation, with two adjustments: they did not have to complete a series of discrete-point items about language use, which are included in the actual exam, and they had to write longer texts (25 words more per text). The decision to remove the discrete-point items and to give more time for writing longer texts was based on the current discussions about changing the format of the exam's written section. The log files were used to determine the time spent on each task. The first task the participants answered started when they opened a window on the computer (e.g., a web browser or a word processor). Timing for the second task ran from the moment they handed in the first task on the school's Intranet to the moment they handed in the second task on the same electronic platform, which was registered in the log files. To crosscheck the timing data, the time registered in the log file that showed when the responses were submitted was compared to the time registered by the electronic platform.

The students' responses were anonymized and independently rated on a 7-point scale by three experienced raters. The raters were French teachers and examiners who were familiar with the French writing test and the rating scale. The raters independently assigned one holistic score for each response based on four main criteria: communication of ideas, grammar and language use, use of resources, and cultural knowledge. The descriptors for the criterion "communication of ideas" included relevance of content and coherence of the response, as well as students' ability to paraphrase and describe information and to express opinions and attitudes. "Grammar and language use" was related to students' ability to use vocabulary, grammar, and syntax effectively. While "use of resources" referred to students' ability to use visual and textual information

from the prompt and the course material, “cultural knowledge” dealt with students’ ability to relate cultural information from their own knowledge or the Internet to their personal experiences. The responses were randomly assigned to the raters so that they would not know whether students accessed the Internet to complete the task. The Pearson correlation coefficient was lower for raters 1 and 2 ($r = 0.57$), but acceptable for raters 1 and 3 ($r = 0.7$) and raters 2 and 3 ($r = 0.72$).

A mixed general linear model (GLM) was used to analyze the data with one within-subjects factor Mode (IA, WIA) and one between-group factor Group (Group 1=Task 1 IA, Task 2 WIA; Group 2= Task 1 WIA, Task 2 IA). The averages of the three raters’ holistic scores were used as dependent variables.

The responses from Task 2 (recommend places to visit in Paris) were also analyzed in terms of number of ideas (places, monuments, art, restaurants) mentioned in the text. The purpose of this analysis was to examine whether those who had IA generated more ideas because they had the opportunity to find information on the web.

4.2.2 Window tracker log files

The students’ behavior when writing under the different test conditions was analyzed using log files provided by a window tracker that registered all the windows students opened on the classroom computers in real time (Tek911 Inc., 2013). The window logs were completely anonymous as no keystroke logs or personal data (e.g., names, passwords) were registered. The raters who rated the students’ written responses were unaware of the existence of these data to prevent any influence on their scoring decisions. The students were given numbers for data-processing purposes, and the log files were retrieved immediately after the test.

The window tracker was used to monitor the information student accessed on the computer. It created a log with the names and the times of the applications or the websites that the students visited. The logs were analyzed to find out which words the students looked up in the dictionary, which Google searches they conducted, which websites they visited, and how much time they spent on the writing prompts and the word processing document. The log files were transferred to spreadsheet documents (Excel), anonymized, and analyzed.

4.2.3 Survey instrument

The survey was designed in Google Forms (Google, n.d.) and was administered online. It consisted of 32, mostly yes/no or multiple-choice questions written in Danish. The questions related to the students’ perceptions of task difficulty, task navigation, use of digital or paper-based sources (e.g. dictionary, conjugation table, Wikipedia), clarity of instructions, and students’ levels of interest. The survey was distributed the day after the test, and it took 15 minutes to complete for logistic reasons. The answers provided by the 32 retained participants were anonymized and gathered in a spreadsheet.

5 Results

5.1 Score variation

The descriptive statistics of the exam data presented in Table 1 show that the means across the two tasks with or without Internet were similar. The Pearson

correlations between the Task 1 and the Task 2 scores for each group were significant (Group 1, $r = .75$; Group 2, $r = .82$). All skewness and kurtosis indices were within ± 2 , which indicates that the data reflect normal distributions. The Shapiro-Wilk tests confirmed that the assumption of the normality was met by the two groups with different mode conditions ($p > .05$). The Levene's tests confirmed that the assumption of the homogeneity was satisfied by the two groups ($p > .05$). The Mauchly's test of Sphericity was not performed because the dependent variable had only two levels.

Table 1. Descriptive statistics for Task 1 and Task 2 for Group 1 and Group 2.

	Group 1		Group 2	
	Task 1 (IA)	Task 2 (WIA)	Task 1(WIA)	Task 2(IA)
Number of students	15	15	17	17
Mean	4.57	4.71	4.38	4.3
Standard deviation	1.19	1.01	.95	.81
Skew	.44	.52	-.56	-.65
Kurtosis	-.36	.45	.97	-.56

Results suggested no statistically significant effects at the .05 level. The within-subjects effect of Mode (IA or WIA) yielded an F ratio of $F(1,30)=.01$, $p=.93$, indicating that the mean score for the IA task ($M=4.33$, $SD=1.01$) was not significantly different from the mean score for the WIA task ($M=4.32$, $SD=.86$). The between-subject effect of Group (Group1, Group2) yielded an F ratio of $F(1, 30)=1.41$, $p=.24$, indicating that the mean score was not significantly different regardless of whether Task 1 was IA or Task 2 was WIA ($M=4.51$, $SD=1.1$) than when Task 1 was WIA and Task 2 was IA ($M=4.15$, $SD =.81$). The interaction effect Mode and Group was non-significant, $F(1, 30) = 1.27$, $p=.26$, which suggested that allowing students to use the Internet, regardless of task, did not have effect on their scores. Table 2 presents the results from the GLM analysis.

Table 2. GLM results for task scores.

source	df	SS	MS	F
Mode	1	.16	.16	.06
Between-subjects	30	52.15	1.73	
Group	1	.016	.016	.76
Mode X Group	1	.187	.187	.78
Within-subjects	30	7.28	.24	

5.2 Students' writing behavior

As mentioned earlier, the log files recorded the time students spent on each site they visited while completing the writing tasks. Results suggested that all students, regardless of whether they accessed the Internet, spent around 20% of their time on the electronic dictionary as the main source, while the conjugation website was not as popular. When they had IA, they spent about 8% of their time on the prompt and only 9% for web search.

Google, a search engine, was the most common Internet tool as 38% of students ($n=12$) looked up (1) words, expressions, and phrases in Danish, English, and French (e.g., "eiffel", "foie gras", "j'ai décide") or (2) searched for specific

information (“best shopping street in paris”). The Google searches ranged from one to 10 occurrences in students’ log files, which means that the students did not use Google search extensively. Some students opened Wikipedia (n=3), Google Maps (n=2), or online folders, such as Dropbox (n=1) or OneDrive (n=1). Three students translated the Paris Tourism Office website into Danish, and one student chose English instead of French at the restaurant website.

Actions such as opening a browser, downloading the responses, logging onto the electronic dictionary, or error messages (e.g., when the browser was “temporarily off-line”) were registered as “Other time on the Internet.” The time registered as “Other time,” on the other hand, was information about opened windows including use of word count applications in the word processor, saving documents, choosing files to download, computer update information (e.g., battery update, automatic opening of “Program Manager”), and going back to the desktop. Table 3 provides specific information about time spent on each type of activity.

Table 3. Time allocation for different activities during task completion.

	Mean time IA		Mean time WIA	
Time on word processor	00:23:42	45%	00:33:12	60%
Time on electronic dictionary	00:10:53	21%	00:12:21	22%
Time on conjugation website	00:01:42	3%	00:02:52	5%
Time on online prompt	00:04:19	8%	00:00:00	
Time on information search	00:04:54	9%	00:00:00	
Other time on the Internet	00:05:16	10%	00:03:59	8%
Other time	00:02:17	4%	00:02:40	5%
Total time	00:53:03		00:55:03	

5.2.1 Time on the word processor

The greatest difference found was that the students spent almost 10 minutes longer on the document when they wrote WIA than when they had IA. When writing WIA, the total time on opened word processor window was an average of 33.12 minutes, or 60% of the assignment time. When given IA, the total time they students spent on the word processor was 23.42 minutes, or 45 % of the total task time. However, long time stretches on the word document were rare in both tasks and conditions (IA or WIA) because the students tended to shift between windows frequently. The maximum time the word processor was opened continuously was 6-9 minutes. This was recorded in seven log files when students had WIA tasks and only two when they had IA tasks. These stretches occurred either at the beginning or at the end of task completion, which meant that they used the time to go over the prompt and to check their texts. On average, the maximum consecutive time the students spent on the document was 2.2 minutes, while they shifted between windows 310 times. The students spent even shorter time on the other windows (websites, dictionary, prompt, etc.) they opened.

5.3. Students' perceptions of a writing test with Internet access

Although the log files revealed sporadic uses of the Internet to search for additional content related to the tasks, the vast majority of the students (85%) preferred to have IA. Some believed the Internet provided more opportunities for finding ideas and information. For instance, one student claimed, "I think that the Internet was useful for "Texte 2" because it is easier to find information." Students also thought that the Internet access made "things clearer" and "gave more freedom."

The Internet gave them access to more "helping tools," even though they were aware that they had used mainly the dictionary and the conjugation website. A student wrote, "I like to have the option. I didn't use the Internet so much, but it was great that it was there." Students also emphasized that the electronic dictionary was essential for their task completion.

Three test-takers (10%) preferred tasks WIA. One student felt it was easier to find the information on paper, and another noted that searching the Internet was time-consuming. It is worth noting that the same student spent 29 minutes 41 seconds on the Internet with the IA task. Three respondents (10%) felt Internet access did not express any preference, and one claimed that, although allowed, he did not use the Internet, which was confirmed with his log file.

In terms of task difficulty, in each of the two groups (Group 1=Task 1 IA, Task 2 WIA; Group 2= Task 1 WIA, Task 2 IA), 52% of the students found Task 2 (recommend places to visit in Paris) more difficult, 28 % found Task 1 more difficult, and 20% thought the tasks were equally difficult. In other words, Task 2 seemed more difficult regardless of Internet access permission. Those who found Task 1 IA more difficult claimed that web searching was time-consuming, and WIA students thought that searching for information in the booklet was inconvenient. IA students who found Task 2 more difficult stated that they were unsure about how to solve the task, while WIA students thought that finding sufficient information was difficult – IA could have helped. Figure 1 summarizes these findings.

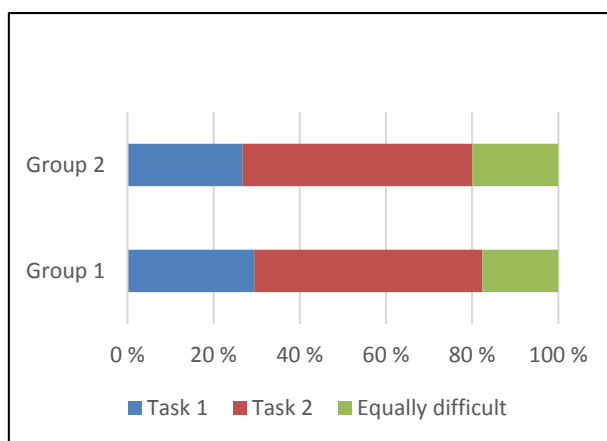


Figure 1. Which task was more difficult?

When asked how they found the ideas for Task 2, most students reported using more than one source. They knew about different landmarks in Paris from their coursework (60%), 47% had been in Paris, and 30% made up some of the information. Forty-one percent of those who had IA (N=7) said that they used it (which was confirmed with the window tracker logs), but only 23% (N=4) used

the Internet as the only source (see Figure 2). The difference in the number of ideas in Task 2 responses (references to places, art, monuments, museums, restaurants) between the IA and WIA group was not significant ($t=0.244$, $p=0.4$).

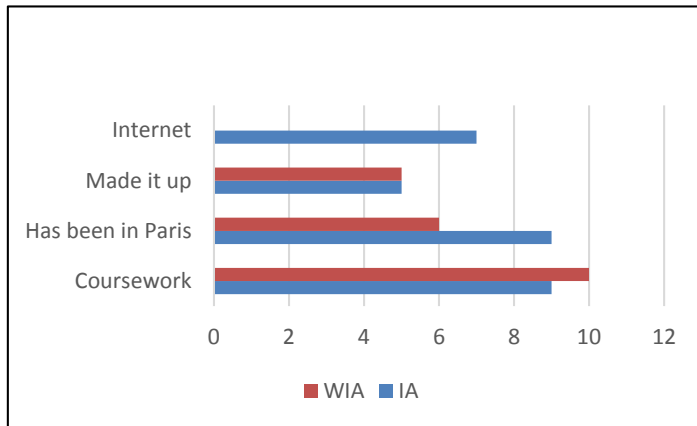


Figure 2. Sources of information for Task 2.

Finally, Internet access in a test situation caused some uncertainty about what was acceptable use of online information, and what was considered cheating. In an informal conversation after the test, a student revealed that she did not dare to copy-paste chunks from the online texts because of fear of being accused of cheating.

6 Discussion

Results from this study corroborate findings from previous research in many ways. They confirm the Danish Evaluation Institute's conclusions that students endorse inclusion of Internet access in the school-leaving L2 writing exams (Danmarks Evalueringsinstitut, 2016) and resemble previous findings in L2 testing with technology in other contexts (Goldberg et al., 2003; Golonka et al., 2014; Wang & Vasquez, 2012). In other words, modern technologies and the Internet are available for students in the Danish school, so their inclusion strengthens the exam's face validity.

Results also substantiate Cumming et al. (2005), Manchón et al. (2007), and Roca de Larios et al.'s (2008) assertions that students at lower proficiency levels lack ability and attention to use sources and time in transformative ways because of controlled and limited vocabulary access. In the current study, most students heavily relied on the bilingual dictionary as their main source because they either had to find the word they wanted to use in French or because they tried to confirm the correct word spelling. Students' attention was placed mostly at word level, which left little time to locate, retrieve, and judge information from the Internet. For that reason, they either avoided Internet searches and relied on their own knowledge or spent too much time on irrelevant searches (Zhang et al., 2009). In other words, notwithstanding Internet access allowance, students seemed to exhibit similar writing behavior, and, hence, no significant effects of Internet on students' scores were found.

The only large difference found between student IA and WIA writing behavior was the time students spent on the document, i.e. on writing the texts. Two

possible interpretations of this finding could be offered. One interpretation may be that the Internet reduced students' cognitive load by helping them deal more easily with some of the topical and linguistic gaps they experienced, as has been argued in previous discussions (Kozlova & Presas, 2013; Stapleton, 2010). Another interpretation may be that the word processor was active while WIA students were reading their paper-based prompts, so the window tracker recorded longer word processor times. Since the monitor lacks keystroke logging capability, no data about the actual writing was available for analysis.

Another consequence of Internet access inclusion test is maintaining test security. If certain applications and websites are not permitted, they need to be blocked as the findings suggested that some students failed to comply with instructions and requirements. Students' attempts to cheat or plagiarize may not have been purposeful but a result of instruction misunderstanding or uncertainty about how to avoid plagiarism. As Evering and Moorman (2012) and Pecorari and Petrić (2014) stressed, students tend to be warned against it, but they remain unclear about what constitutes cheating or plagiarism, especially in the online environment. For example, machine translation programs, like Google Translate, were not allowed in the current study, so their use would be considered cheating. The question nonetheless remains as whether students should be trained regarding how to avoid them or how to use them effectively as support tools.

To address García Laborda's (2007) questions regarding the design of items that are appropriate for an Internet-based environment, this study suggested that task characteristics influenced test-takers' online behavior. If the two tasks are compared, although the prompt of Task 1 (comparison of two restaurants) provided hyperlinks to two restaurant websites the students needed to visit, it failed to promote web searches for additional information that could be found on the Internet. Task 2 (planning a sightseeing afternoon), on the other hand, generated more web searches because the students needed to find tourist information to complete the task. In other words, Task 2 seemed more engaging with the Internet environment. Although students found Task 2 more difficult than Task 1, they performed equally well on both.

Designing relevant L2 French writing tasks (IA) was a challenge because of students' low proficiency and because of Internet source availability in different languages. Students' age (15-16 years) and proficiency level (A2) limited the selection of appropriate topics, genres, and audiences. Concerning the languages in which sources were available, results suggested that some students accessed the English versions of the websites, which raises the question whether and how students' multilingual repertoire should become part of their French writing construct.

Designing IA tasks does not mean transferring paper-based writing items to an online interface. Search for and use of relevant information on the Internet as part of task completion should be part of the design. Adding hyperlinks to different webpages or websites may be insufficient because clicking on the provided hyperlink does not constitute web search and can easily be simulated with an offline digital application. There is no need to compromise the security of the test by allowing IA if test-takers are not required to perform web search.

Finally, although the implementation of the new exam format is expected to promote integration of ICT in foreign language learning through positive washback effects, it may lead to inequality in student exam results if informational digital literacy is not an explicit goal of the foreign language

curriculum. The inequality would arise from the difference in teachers' interpretation of the exam tasks, their informational digital literacy, their access to resources, and their level of engagement with ICT in class. The variation in teachers' perceptions, experiences, and practices would lead to differences in student learning, and that would affect students' results on the exam.

Another concern is the BYOD trend, where students can bring their own devices not only to classes but also to the exams. In this study, all students had to use the school's computers, but if an exam with internet access is introduced, BYOD might reinforce inequalities in students' test performances because schools cannot control the functionalities, settings, applications, and software installed on students' own equipment. In other words, students who bring their own devices will have an advantage over those who use school devices.

7 Conclusion

Rather than investigating the effects of the response or delivery mode on test-takers' performance, the current study's focus was on the Internet as a task input mode. In other words, the study analyzed test-takers' usage of the Internet as an information source and its effect on their writing performance. Many test-developers may argue against allowing Internet access in formal assessment due to possible issues with test security and difficulties understanding the interaction between digital information literacy and L2 writing proficiency. However, given that young L2 writers inevitably use the Internet as a writing platform and as a resource to retrieve information and relevant linguistic structures in and out of school, perhaps assessing their L2 writing with Internet access should be given more consideration, especially regarding task design and comprehension of the expanded L2 writing construct.

The decision about what Internet resources should be allowed depends not only on the focal construct but also on the level of test security. For example, to avoid security breaches, some Internet resources were not allowed in this study (e.g. social platforms and machine translation). Access to cloud solutions (e.g., Dropbox, OneDrive, Google Drive) and social platforms (e.g., Facebook, Snapchat, Discord) allow students to be in touch with other people who could write the test responses for them. Monitoring students' access to these websites is difficult unless a window tracker is applied to flag access to such web services.

If the effective application of machine translation is included in the writing construct, another challenge is to define to what degree students should use machine translation as a resource to support writing (e.g., translating or double-checking a word, chunk, or phrase) and identify when they use it to translate the entire response written in L1. Machine translation evolves constantly given the introduction of neural machine translation which uses artificial intelligence (Briggs, 2018).

Findings suggest that students may engage with the Internet more if there is an information gap that they need to fill through a web search. Simply allowing students to use the Internet does not necessarily mean that they will do it – the task may need inclusion of requirements for finding, evaluating, and using of online information. The Internet also allows for mediation of information, which means students have an opportunity to use all their linguistic resources to complete the task. More importantly, the students feel more comfortable and more

motivated to write with Internet in the testing context. Since the current study involved only students at a beginner level, further research that includes proficiency and information literacy levels as factors is necessary to confirm the findings.

The results of this study may need to be interpreted with some caution due to the limited number of participants. Moreover, the low-stakes setting of the study did not elicit the same behavior that could have been seen at high-stakes exams, such as cheating. However, the results remain useful as they provide initial insights into students' response processes, including their writing behavior and motivation, which represent essential validity evidence (American Educational Research Association et al., 2014; Zumbo & Hubley, 2017).

Another limitation is that although the window tracker recorded all activities on the computer, it did not provide any information about how the students constructed the text as detailed information regarding the writing process (e.g., editing, proofing, structuring, adding/deleting information, moving text around) was unavailable. The logs could be applied only to track when and how long the word processor was active. The window tracker used in this study proved undependable in some instances. An application developed for research purposes, such as Inputlog (Leijten & Waes, 2013), might have ensured a larger pool of data. Future studies need to include monitoring applications that track keystroke in order to obtain more precise data regarding the actual writing process.

Finally, when developing writing tasks with Internet access, given that integrating digital literacy in language assessment tasks could affect the language performance, it would be important to examine the interaction between digital literacy and language use within the local context. Implementation of such tasks would only make sense if relevant access to digital devices and software is available to all students, and digital literacy is explicitly taught in relation to L2. Test fairness could be at stake if unequal access to digital devices and Internet connection exists in the local community. The purpose of the study was an initial step to understand the extended L2 writing construct by obtaining information about what writing behavior and performance tasks with Internet access elicit. The findings inform the design of relevant tasks with Internet access and the establishment of appropriate testing conditions. In order to ensure lack of bias against certain demographic groups (e.g., ethnic, linguistic, gender), once developed, the tasks need to be validated with a large scale study that includes the different student groups.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- American Library Association. (2000). *Information literacy competency standards for higher education*. The Association of College and Research Libraries, A division of the American Library Association. <http://www.ala.org/acrl/sites/ala.org/acrl/files/content/standards/standards.pdf>
- Ananiadou, K., & Claro, M. (2009). *21st century skills and competences for new millennium learners in OECD countries* [Working Paper]. OECD. <https://doi.org/10.1787/218525261154>
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In *Assessment and Teaching of 21st Century Skills* (pp. 17–66). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-2324-5_2
- Briggs, N. (2018). Neural machine translation tools in the language learning classroom: students' use, perceptions, and analyses. *JALT CALL Journal*, 14(1), 2–24. <https://journal.jaltcall.org/articles/221>
- Børne- og Undervisningsministeriet. (2018). *Evaluering af it i folkeskolen* [Evaluation of ICT in primary in lower secondary public school]. <https://www.uvm.dk/publikationer/2018/180619-evaluering-af-it-i-folkeskolen>
- Børne- og Undervisningsministeriet. (2019). *Statuspublikation: Digitalisering med omtanke og udsyn* [Status report: Thoughtful and forward looking digitalization]. <https://www.uvm.dk/publikationer/2019/190313-digitalisering-med-omtanke-og-udsyn>
- Børns Vilkår. (2019). *Skolebørns liv med digitale medier hjemme og i skolen* [School children's life with digital media at home and at school] (No. 3; Digital dannelse i børnehøjde [Digital education at children's level]). https://www.medieraadet.dk/files/docs/2019-02/Digital_Dannelse%20Del%203.pdf
- Buckingham, D. (2006). Defining digital literacy: What do young people need to know about digital media? *Nordic Journal of Digital Literacy*, 1(04), 263–277. https://www.idunn.no/dk/2006/04/defining_digital_literacy_what_do_young_people_need_to_know_about_digital
- Bulger, M., Mayer, R., & Metzger, M. (2014). Knowledge and processes that predict proficiency in digital literacy. *Reading & Writing*, 27(9), 1567–1583. <https://doi.org/10.1007/s11145-014-9507-2>
- Chapelle, C. A., & Douglas, D. (2006). *Assessing language through computer technology*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511733116>
- Christensen, E. (2017, May 18). Danskprøve-kommission: Iprøven har været undervejs siden 2009 [Exam commission for Danish: The I-exam has been on its way since 2009]. *Folkeskolen*. <https://www.folkeskolen.dk/608253/danskproeve-kommission-iproeven-har-vaeret-undervejs-siden-2009>
- Coiro, J., & Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Reading Research Quarterly*, 42(2), 214–257. <https://doi.org/10.1598/RRQ.42.2.2>
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press. <https://www.coe.int/en/web/common-european-framework-reference-languages/home>
- Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing*, 10(1), 5–43. <https://doi.org/10.1016/j.asw.2005.02.001>
- DAMVAD. (2013). *Evaluering af forsøg med folkeskolens prøver skoleåret 2012-2013* [Evaluation of trials with the school leaving exams school year 2012-2013]. DAMVAD. <http://www.emu.dk/sites/default/files/Evaluering%20af%20forsog%20med%20folkeskolens%20proever%202012%202013.pdf>
- Danmarks Evalueringsinstitut. (2016). *Evaluering af forsøg med adgang til internet ved skriftlige prøver i fremmedsprog* [Evaluation of written exam trials with internet access in Foreign Languages]. Danmarks Evalueringsinstitut. <https://uvm.dk:443/publikationer/folkeskolen/2017-eva-rapport-evaluering-af-forsog-med-adgang-til-internet-ved-skriftlige-proever-i-fremmedsprog>

- Danmarks Statistik. (2020). *It-anvendelse i befolkningen 2020 [Use of ICT in the population in 2020]*. <https://www.dst.dk/Site/Dst/Udgivelser/GetPubFile.aspx?id=29450&sid=itbef2020>
- Digitaliseringsstyrelsen. (2016). *Digitaliseringsstrategi-2011-15. 3.2.a Trådløst netværk på skolerne frem mod 2014. Afsluttende rapport for initiativ 3.2a og 3.2b, Bilag 2.c.a. [Strategy for digitalization-2011-15. 3.2.a Wireless network in schools towards 2014. Final report for initiatives 3.2a and 3.2b, Appendix 2.c.a.]* <https://digst.dk/Strategier/Digitaliseringsstrategi-2011-15>
- Dimova, S., Yan, X., & Ginther, A. (2020). *Local language testing: Design, implementation, and development*. Routledge. <https://doi.org/10.4324/9780429492242>
- Dooley, P. (2008). Language testing and technology: problems of transition to a new era. *ReCALL*, 20(1), 21–34. <https://doi.org/10.1017/S0958344008000311>
- Douglas, D., & Hegelheimer, V. (2007). Assessing language using computer technology. *Annual Review of Applied Linguistics*, 27, 115–132. <https://doi.org/10.1017/S0267190508070062>
- Dudenev, Gavin, & Hockly, N. (2016). Literacies, technologies and language teaching. In Farr, Fiona & Murray, Liam (Eds.), *The Routledge handbook of language learning and technology* (pp.166-181). ProQuest Ebook Central <http://ebookcentral.proquest.com>.
- European Commission/EACEA/Eurydice. (2019). *Digital education at school in Europe*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2797/763>
- Evering, L. C., & Moorman, G. (2012). Rethinking Plagiarism in the Digital Age. *Journal of Adolescent & Adult Literacy*, 56(1), 35–44. <https://doi.org/10.1002/JAAL.00100>
- Fulcher, G. (2003). Interface design in computer-based language testing. *Language Testing*, 20(4), 384–408. <https://doi.org/10.1191/0265532203lt265oa>
- Garcia, I., & Pena, M. I. (2011). Machine translation-assisted language learning: writing for beginners. *Computer Assisted Language Learning*, 24(5), 471–487. <https://doi.org/10.1080/09588221.2011.582687>
- García Laborda, J. (2007). On the Net – Introducing Standardized EFL/ESL Exams. *Language Learning & Technology*, 11(2), 3–9. <http://dx.doi.org/10125/44097>
- Geluso, J. (2011). Phraseology and frequency of occurrence on the web: Native speakers' perceptions of Google-informed second language writing. *Computer Assisted Language Learning*, 26(2), 144–157. <http://dx.doi.org/10.1080/09588221.2011.639786>
- Goldberg, A., Russell, M., & Cook, A. (2003). The effect of computers on student writing: A meta-analysis of studies from 1992 to 2002. *The Journal of Technology, Learning and Assessment*, 2(1), 3-51. <https://ejournals.bc.edu/ojs/index.php/jtla/article/view/1661>
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27(1), 70–105. <https://doi.org/10.1080/09588221.2012.700315>
- Google. (n.d.). *Google Forms*. Retrieved August 20, 2017, from <https://docs.google.com/forms/>
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388. <https://doi.org/10.1177/0741088312451260>
- Jin, Y., & Yan, M. (2017). Computer Literacy and the Construct Validity of a High-Stakes Computer-Based Writing Assessment. *Language Assessment Quarterly*, 14(2), 101–119. <https://doi.org/10.1080/15434303.2016.1261293>
- Jones-Kavalier, B. R., & Flannigan, S. L. (2006). Connecting the Digital Dots: Literacy of the 21st Century. *Educause Quarterly*, 2, 8–10. <https://er.educause.edu/articles/2006/1/connecting-the-digital-dots-literacy-of-the-21st-century>
- Koldborg, S., & Pam, S. (2014, October 1). Fakta: Hvad er "Bring your own device"? [Facts: What is "Bring your own device"?]. *Danmarks Radio*. <https://www.dr.dk/nyheder/penge/fakta-hvad-er-bring-your-own-device>
- Kozlova, I., & Presas, M. (2013). ESP students' views on online language resources for L2 test production purposes. *Teaching English with Technology*, 13(3), 35–52. <https://www.tewtjournal.org/issues/past-issue-2013/past-issue-2013-issue-3/>
- Laurie, R., Bridglall, B. L., & Arseneault, P. (2015). Investigating the effect of computer-administered versus traditional paper and pencil assessments on student writing achievement. *SAGE Open*, 5(2), 1-8. <https://doi.org/10.1177/2158244015584616>

- Leijten, M., & Waes, L. V. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication*, 30(3), 358–392. <https://doi.org/10.1177/0741088313491692>
- Leijten, Mariëlle, Van Waes, L., Schriver, K., & Hayes, J. R. (2014). Writing in the workplace: Constructing documents using multiple digital sources. *Journal of Writing Research*, 5(3), 285–337. <https://doi.org/10.17239/jowr-2014.05.03.3>
- Manchón, R. M., Murphy, L., & Larios, J. (2007). Lexical retrieval processes and strategies in second language writing: A synthesis of empirical research. *International Journal of English Studies*, 7(2), 149–174. <https://revistas.um.es/ijes/article/view/49041>
- Pecorari, D., & Petrić, B. (2014). Plagiarism in second-language writing. *Language Teaching*, 47(3), 269–302. <https://doi.org/10.1017/S0261444814000056>
- Pecorari, D., & Shaw, P. (2012). Types of student intertextuality and faculty attitudes. *Journal of Second Language Writing*, 21(2), 149–164. <https://doi.org/10.1016/j.jslw.2012.03.006>
- Redecker, C., & Punie, Y. (2017). *European Framework for the Digital Competence of Educators: DigCompEdu*. <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/european-framework-digital-competence-educators-digcompedu>
- Roca de Larios, J., Manchon, R., Murphy, L., & Marin, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing*, 17(1), 30–47. <https://doi.org/10.1016/j.jslw.2007.08.005>
- Roever, C. (2001). Web-based language testing. *Language Learning & Technology*, 5(2), 84–94. <http://dx.doi.org/10125/25129>
- Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, 333(6043), 776–778. <https://doi.org/10.1126/science.1207745>
- Stapleton, P. (2010). Writing in an electronic age: A case study of L2 composing processes. *Journal of English for Academic Purposes*, 9(4), 295–307. <https://doi.org/10.1016/j.jeap.2010.10.002>
- Stapleton, P. (2012). Shifting cognitive processes while composing in an electronic environment: A study of L2 graduate writing. *Applied Linguistics Review*, 3(1), 151–171. <https://doi.org/10.1515/applirev-2012-0007>
- Stricker, L. J., & Attali, Y. (2010). *Test takers' attitudes about the TOEFL iBT[TM]. TOEFL iBT research report. RR-10-2*. Educational Testing Service. <https://files.eric.ed.gov/fulltext/ED512343.pdf>
- Tek911 Inc. (2013). *TEK911 [At.exe and monitorviewer.exe]*. <http://www.tek911.com/readme-monitor.htm>
- Undervisningsministeriet. (2017a). *Velfungerende it i undervisningen [Well functioning ICT in teaching]*. <https://www.uvm.dk:443/folkeskolen/laering-og-laeringsmiljoe/it-i-undervisningen/velfungerende-it-i-undervisningen>
- Undervisningsministeriet. (2017b, August 17). *Optagelse til de gymnasiale uddannelser [Admission to upper secondary school level]*. UddannelsesGuiden. <https://www.ug.dk/uddannelser/artikleromuddannelser/omgymnasialeuddannelser/optagelse-til-de-gymnasiale-uddannelser>
- van Deursen, A. J. A. M. van, & van Dijk, J. A. G. M. van. (2009). Using the internet: Skill related problems in users' online behavior. *Interacting with Computers*, 21(5–6), 393–402. <https://doi.org/10.1016/j.intcom.2009.06.005>
- Wang, S., & Vasquez, C. (2012). Web 2.0 and second language learning: What does the research tell us? *CALICO Journal*, 29(3), 412–430. <https://dx.doi.org/10.11139/cj.29.3.412-430>
- Warnich, A. (2013, December 13). Succes med internet og gruppearbejde til afgangsprøven [Internet access and group work was successful at the final exams]. *Folkeskolen*. <https://www.folkeskolen.dk/537977/succes-med-internet-og-gruppearbejde-til-afgangsproeven>
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511732997>
- Zhang, Y., Robins, D., Holmes, J., & Salaba, A. (2009). Understanding internet searching performance in a heterogeneous portal for K-12 students: Search success, search time,

strategy, and effort. *Journal of Web Librarianship*, 3(1), 15–33. <https://doi.org/10.1080/19322900802660235>

Zumbo, B. D., & Hubley, A. M. (Eds.). (2017). *Understanding and investigating response processes in validation research* (Vol. 69). Springer International Publishing. <https://doi.org/10.1007/978-3-319-56129-5>

Received November 9, 2020
Revision received February 4, 2021
Accepted February 24, 2021