TEACHERS' FORUM

Exploring the Impact of Handwriting vs. Keyboarding on L2 Assessments: Biases, Integrity, Authenticity, and Literacies

DENNIS BOGUSZ

New York University
E-mail: db538@nyu.edu

Is paper or computer better for assessing L2 students' writing? The ineluctable transition to technology might suggest this question has already been answered. However, the technology divide in L2 assessments may have indeed widened since the pandemic: whereas some teachers have fully embraced technology in assessments as in instruction, others are reluctant to eliminate paper, owing to concerns about the reliability, integrity and authenticity of L2 production on computer. This article shares observations from several French classes at an American high school in which assessments that were otherwise identical were offered to students on both paper and computer. These observations revealed several overlapping areas of L2 research that merit further consideration, including instructor bias between media, academic integrity of student work, and the need to align the technological literacies between instructors and students. The reflection that follows points to specific directions for further empirical research on the effects of input medium on L2 learners in K-12 and higher education.

INTRODUCTION

It has been almost a hundred years since a keyboard first appeared in classrooms. In 1926, Horace Mann School in New York was an early adopter of the typewriter in its first-grade curriculum (Sinks & Thurston, 1972). Not until the launch of the personal computer in the 1980s and widespread adoption of word processing in the 1990s, however, did questions about the difference between students' handwritten and typewritten output become an interest of education research. Since then, findings have perpetuated, rather than settled, the debate between medium of input and effect on student learning. Although the Covid pandemic catapulted many teachers into adopting computer-based assessments, others have since reverted back to those based on paper owing to equitable access to technology, lack of technological literacy, student distractions, or concerns about academic integrity and authenticity.

While teaching French at a private American high school, I found myself facing these issues without resolution. In 2021-22, I continued using both paper-based and electronic media in my instruction and expected students to do the same in their learning. When it came time for assessments, however, I decided to offer students the choice to take the same assessment on either paper or computer. My aim was simply to compare L2 assessments by medium to help discern student preference, observe any significant differences in grades between media, and thereby inform my teaching practice. To be forthcoming, I have since

discontinued this option: although my class materials and student assignments still alternate between media, assessments are now paper-only.

For starters, few of my students across two French classes opted for the electronic format when I offered it on two separate quizzes. Students reported a general comfort with paper and perceived no advantage to taking the same quiz on a computer. Just two students opted to take the first quiz on computer, and two more joined them on the second quiz. Students completed both quizzes in the time allotted, and there were no significant differences in completion per medium. Mean grades did not significantly differ across the class per medium nor between overall assessment grades per student. I had supposed more students would want to take their quizzes on computer, which prompted me to investigate further how this issue is treated in the wider field.

I began researching how scholars had taken up similar points previously and was surprised to find that although a substantial body of research exists on the difference between handwriting and keyboarding in L1 environments, far less focuses specifically on L2 assessments. Several overlapping areas emerged in my research, a synthesis of which I share below. I then began to reflect critically on what I was observing in the classroom. I also share these reflections in the hope they inspire similar critical thought for both practitioners and scholars and incite greater intentionality when it comes to medium of L2 assessment.

PREVIOUS RESEARCH

Handwriting vs. Keyboarding

For partisans of paper, evidence suggests that handwriting is superior to typing for literacy development (Longcamp et al., 2008; Longcamp et al., 2011), letter recognition and brain stimuli, among even preliterate children (James & Engelhardt, 2012). Handwriting continues to activate the adult brain in ways that neither typing nor tracing can do to a similar degree (James, 2017). In addition to overall literacy development, handwriting improves visual recognition of graphs for perceptual learning (Araújo et al., 2022). In a study by Mueller and Oppenheimer (2014), university students who handwrote notes outperformed those who did so on their laptops in conceptual questions when tested immediately after lectures, though Morehead and colleagues (2019) qualified these results by showing that it was not the taking of notes in a particular medium that improved student performance on tests, but rather studying them afterward. Haring and Kelner (2021) moved such earlier lab-based research into the classroom with a study of seventh-grade English and social studies students who took research notes either by hand or on computer. In direct contrast to Mueller & Oppenheimer (2014), they found students who compiled their research notes on computer earned higher scores from their teachers and anonymous raters than those who compiled handwritten notes. Still, other practitioners advocate note-taking by hand in order to improve organizational skills and develop focus, especially among students in secondary education who are less adept at multitasking than more mature learners (Barbour, 2021). Paper notebooks also permit drawing and the use of color with a variety of instruments as well as different tactile materials (other print items, tape, glue, stickers, etc.) on which many students rely for association and recall.

Many studies support the seemingly ineluctable transition to technology. One study of third-grade students shows that writing speed is a certain benefit of computer-based tests (Wollscheid et al., 2016). In another, sixth-grade students who had taken a keyboarding class produced more legible text by computer than on paper, suggesting that keyboard input might

be especially beneficial for students who struggle with handwriting (Rogers & Case-Smith, 2002).

Students might have different skills and comfort levels in handwriting, whether the instrument is pencil, pen, stylus, or fingertip and keystrokes. Among German kindergarten children, letter recognition and visuo-spatial skills were superior among those trained on paper and pencil than those on computer; however, reading and word writing were better among students trained to use a computer keyboard (Mayer et al., 2020). The teaching of handwriting typically diminishes after primary school, yet a survey of teachers in Germany found L1 deficiencies in legible handwriting pervade in secondary schools even more than in primary schools, and significantly more so among boys (Marquardt et al., 2016). A related difficulty exists in the distinction between manuscript and cursive writing. A history professor and former president of Harvard University found that roughly two-thirds of her class could not read her handwritten feedback and would either ask someone to explain her handwritten comments on their work or simply disregard them (Faust, 2022).

When it comes to handwriting versus keyboarding in L2 contexts, most research has been done on students of English. Examinees of the Test of English as a Foreign Language in the United States aged 15 to 55 scored marginally higher on the handwritten essay portion than on computer, particularly when they scored lower on the multiple-choice portion of the test (Wolfe & Manalo, 2004). One American university's English placement test for newly-admitted foreign language students, which was offered synchronously on paper and asynchronously on computer, revealed no difference in the quality of the essay portions between paper and computer written essays, although the latter were slightly longer (Kim et al., 2018). However, many students are not fluent in keyboarding in the L1, much less in the L2, which might involve accent marks, punctuation, symbols, new letters and ligatures, or character input features for non-Roman script languages.

More research on the effects of the range of handwriting instruction on early literacy outcomes is still needed (Hall et al., 2015), especially with regards to the effects of multiple technologies on composition in the L2 classroom (Li et al., 2017). Vasylets and Marín (2022) stress the importance of measuring the differences in writing modes and cognitive load, which depend on age, handwriting and typing skills, and knowledge of L2 as well as computer literacy and previous writing experience. Further observation from the classroom is therefore needed on the cognitive differences between handwriting and keyboarding and their comparative effects on L2 literacy development. Of immediate practical implication is the importance of aligning medium of instruction with that of assessment. Introduction of a new medium during an assessment could present additional cognitive challenges and entail adverse consequences. Similarly, when assessments are in mixed media, teachers need to be aware of students' ease with all media to ensure the validity of the testing instrument.

Instructor Bias Between Media

Earlier research demonstrates bias between evaluators who grade higher for handwritten work than for typewritten work (Powers et al., 1994; Russell & Tao, 2004), even when legibility is not an evaluative criterion (Klein & Taub, 2005): raters gave lower scores to computer-input assessments because they took a stricter view to spelling errors and punctuation in computer form, associated typed text with being a final draft and had higher expectations of quality, and because raters could relate to test-takers who handwrote and were more willing to overlook mistakes in general that they would not for computer-input tests. More recently, a meta-

analysis of studies of K-12 students and handwriting instruction to improve legibility and fluency revealed that although instruction helped students produce quality text, their writing was still implicitly evaluated during the explicit evaluation of other material (Santangelo & Graham, 2016). Some instructors might associate legibility with cognition or effort while others might claim indifference to their students' handwriting. Still, handwriting style, instrument choice, unprompted student enhancements (underlining, color, or drawing), and recognition of the student via handwriting can singularly or cumulatively bias scoring.

Not all bias is necessarily eliminated by computer-based testing, however, and new biases might even emerge. Some instructors might expect more of typewritten text than handwritten text, potentially reinforcing an assumption that students can type more, faster, and with better accuracy than they can handwrite, whether or not they have had related instruction (Canz et al., 2020). Still, other instructors maintain the belief that handwriting stimulates better language acquisition and recall (Sweedler-Brown, 1991). Instructors could demonstrate bias when creating assessments by different media, for example privileging technology for some, while insisting that only handwriting is acceptable for others. Instructors might therefore conflate assessment of language literacy with that of technological literacy. Finally, instructors' biases about distinct electronic devices, such as condoning the use of tablets and laptops but not phones, or biases about certain applications, licensing, cost, data privacy, or even attitudes about the school's technology policies, might also emerge in the design and evaluation of L2 assessments (Nikolopoulou, 2020; Støle et al., 2020).

Academic Integrity

Instructors might assume that many students, when given the opportunity, resort to online translation or generative artificial intelligence to compose text in the target language as opposed to learning how to compose in the target language themselves. Students might grow accustomed to doing this on regular assignments and could be caught unprepared for a paper-based assessment or a computer-based one with restricted use. A more fruitful approach could be to create dialogue between instructors and students on the appropriate use of computer-mediated technologies in such circumstances and jointly develop related classroom policies (Hellmich & Vinall, 2023). Some teachers rely on in-person, synchronous, paper-based assessments exclusively in order to uphold academic integrity and minimize cheating on assessments. Yet the expansive growth of computer-based learning requires adjustments to computer-based assessments to prevent academic dishonesty (Surahman & Wang, 2022). In a survey of ten languages taught in higher education institutions across the United Kingdom, Europe, and Australia, writing assessments that moved online during the pandemic were among the worst impacted by academic dishonesty (Polisca et al., 2022).

Disabling notifications, restricting access to non-essential applications or live messaging, and surveillance are among the most salient solutions to prevent students from soliciting unauthorized help on an assessment. Still, these can introduce new challenges that do not resolve the underlying problem of students feeling unprepared for a given assessment regardless of medium. In the face of these technological challenges, some L2 instructors might resort to synchronous, paper-based assessments exclusively.

Evidence suggests that an honor pledge or similar promise of academic integrity at the start of an assessment primes the student to resist cheating (Tatum & Schwartz, 2017). However, less known is whether handwriting such a pledge and manually signing it are as effective in upholding academic integrity as typing a pledge and one's signature on keyboard.

Observing how students engage with their language learning resources and assessment tools is also just as important as collecting data on their assessment scores (Chun et al. 2016). As a corollary, normalizing observation and informing students that they will be measured for their interaction with the technology in addition to that of their language learning could either reduce or enhance anxiety associated with surveillance and should be explored in further L2 research.

Assessing Authentic L2 Knowledge

The moment a student employs any technology, their communication is mediated to some degree. However, it is important to distinguish between authenticity as representation or reconstruction of real-life situations (e.g., Gilmore 2007) for L2 instruction and student language production and assessment thereof. Technologies, be they paper and pen or computer keyboards, can serve simply as writing instruments in which no linguistic assistance to the student is available. Technology with more advanced features can also serve to help a student produce language they would not otherwise have done alone. The technologies I briefly outline below show the range in which technologies can assist or even replace L2 production. For the sake of simplicity, I refer to as "authentic" the language a student produces without linguistic assistance from technology. I use this notion of authenticity synonymously with natural language production.

In between the absence of any technology-mediated L2 production and complete replacement of it is a range of technologies that impact L2 assessment. Computer-based corrective feedback, for instance, offers students certain benefits that depend on the assessment objectives. Students tend to make low-level revisions to the form and typography of their composition without making the high-level revisions related to content or organization (Barkaoui, 2016; New, 2002). Sedor (2022) found that computer-based corrective feedback was limited to interpretive and presentational tasks. Thomas (2018) introduced analogy-based corrective feedback in which L2 students were presented an example of an error they must identify and decode, allowing them to deduce corrections and apply them to revisions. However, this form of feedback must allow errors to exist in the initial composition, which paper does naturally and the computer does only if so configured. Computer-assisted corrective feedback can be of particular benefit when spelling is not being assessed as part of the composition's quantity or quality. The immediacy of feedback is another feature of many electronic assessments, though that might not always be necessary or desired depending on the assessment objective.

Consider further the range of technologically mediated language from text completion, voice-to-text input, translation, editing, and generative text with artificial intelligence. Kern (2015) asks how medium alters the communicative ecology: paper and computer constrain language distinctly, and how students compose in each creates not just written language but also social identity. L2 instructors must question to what extent they are assessing their students' language production between technologies, as these can range from simple assistance to wholesale generation. Practitioners and researchers alike should therefore seriously consider the wider implications for instruction as for assessment given the advent of generative artificial intelligence and its emergence as a writing tool more generally. AI has the clear capability to follow students' instructions to produce L2 text without students needing to produce any of that text in the target language themselves.

Additional Technological Literacies in L2

Closing the gap between students' tech literacies inside and outside the classroom is a promising means to support L2 writing skills (Thorne & Reinhardt, 2008; Jenkins et al., 2009; Zheng & Warschauer, 2017). Though potentially helpful in expanding formative assessments, technologies could require additional literacies that might exacerbate the burdens of assessment design and grading instead of alleviating them. Moreover, applications and the data they maintain could become proprietary, leaving teachers with budgetary constraints and less flexibility, similar to online textbook subscriptions. Whereas these technologies hold promise for the programmers that develop them and demonstrate impressive results in experimental studies, thorough analyses of instructor capabilities are crucially needed to understand how to further implement them in the L2 classroom.

What is gained or lost in L2 literacy development through the replacement of paper by technological tools needs to be measured rather than assumed. As Elola & Oskoz (2017) remind, interweaving L2 literacy with digital literacy requires deliberation as to technological tool and pedagogical objective. Of equal importance is measuring the direction of how such L2 learning technologies are adopted, from adult learners to children, or vice versa, which could reveal whether technological replacement is coterminous with replacement of paper, and which artifactual or pedagogical gains are achieved in the direction of adoption among different age cohorts.

Meanwhile, L2 educators must be alert to what other tests their students will be taking and evaluate how their assessments complement or substitute these other tests. The College Board, for example, requires handwriting on paper for the writing portions of the Advanced Placement exams in French, German, Italian, and Spanish Language and Culture, as well as AP Spanish Literature and Culture exams, yet the AP Chinese and Japanese Language and Culture Exams are entirely computer-based. The HSK test of Chinese language fluency used in mainland China, however, remains completely paper-based. Likewise, the Spanish DELE and French DALF, like other exams measuring language fluency according to the Common European Framework of Reference for Languages, are currently paper-based. However, the National French Contest that students from my school also took separately is entirely computer-based. So, too, are ACTFL's Writing Proficiency Test (for adults) and Assessment of Performance toward Proficiency in Languages (K-12 students). After completing a series of field tests in 2023, the National Assessment of Educational Progress (The Nation's Report Card) moved to shift entirely online in 2024 and will move to remote administration and proctoring in 2026. Therefore, L2 educators ought to consider how their students are being concurrently assessed in other subjects in order to identify any complementarities or deficiencies, and justify particular divergences.

A related feature on some assessments involves computer-adaptive tests (CATs), which use algorithms to adjust question difficulty as students progress through an assessment. They are common in multiple-choice questions on certain standardized tests. The Nation's Report Card, for example, will become adaptive in 2028. CATs in the diagnosis of original writing might be limited, at least presently, so teachers might have limited means to alter question format in a typical L2 classroom. CATs are also unlike many paper- and otherwise computer-based assessments that allow students to view all of the questions on an assessment, or at least portions thereof, at once. Technology can indeed help produce communication with greater efficiency and accuracy but also alter its social and psychological conditions.

REFLECTIONS

Although I do not teach handwriting in either longform or cursive, I do expose my students to examples of French cursive handwriting since it is still pervasive in many French-speaking countries. Cursive handwriting has already been dropped from Common Core standards in the United States, yet it is almost exclusively taught among primary school children in France. L2 teachers might want to consider the cross-cultural competencies that handwriting in the target language, whether cursive or manuscript, might entail. Also related to the topic of handwriting, as much as I try to randomize my handwritten assessments and ignore a student's name on the cover sheet, it is often the case that I can recognize a given student's handwriting while grading, which can introduce bias.

I make fair use of digital tools for L2 instruction and would like to do more. I teach my students how to compose text in the target language, using accent marks, different punctuation and other symbols that are new to them, using keystrokes on a standard American QWERTY keyboard. I also show them how to use authorized assistance, such as online dictionaries and the spell check and grammar check features in the target language in Google Docs, for instance. Students seldom use laptops during class and very rarely tablets. Phones are prohibited during class as per school policy unless they are used for certain learning games like Kahoot!. Some students report using phone-based applications, such as Duolingo, outside class to supplement their in-class L2 learning. Although some instructors assume students are ahead of them in terms of technological prowess, I have often found this to not be the case. While I do learn quite a bit from my students in terms of technology, I nonetheless continuously find myself needing to impart technological literacy along with L2 literacy to my students. It also remains up to me to explain which media are suited to specific language productions and how.

I encourage students to produce as much original language as possible on their own or in small groups. Nevertheless, I have been alarmed at the extent to which students increasingly depend on technology in less than beneficial ways, from using applications to complete worksheet questions to copying generated text onto writing assignments and presentations. In my observations, students tend to over-rely on corrective feedback in their computer-based work without engaging in deeper learning. Where both my students and I are jointly neophytes is in regard to generative artificial intelligence. Some of my colleagues liken it to a pencil and paper or a computer as just another tool. I remind them, however, that a student still creates their own language expression on paper or on computer whereas AI creates it for them. Each iteration with AI certainly helps to improve AI's language production, not necessarily students' natural language production.

As for assessments, I find that students generally prefer whichever medium, paper or computer, that they perceive positively correlates with higher scores. In the quizzes I described in the introduction, students in all my classes were required to transcribe the school's honor pledge at the start of their assessments, either on paper or into a related field in Google Forms. In addition to my visual monitoring during the computer assessments, I also used Dyknow, an application that monitors students' devices while they are logged onto the school's network. Students were aware of this surveillance, which might have also prevented cheating. As in the paper-based quizzes, I encountered no suspected cases of cheating on the electronic ones either.

Since conducting those quizzes, I have since resumed paper-based only assessments. I have had to reconcile my wish for greater technological integration in L2 assessment with

the reality that students are over-relying on it for their regular assignments. I have witnessed firsthand the importance of aligning media of instruction with that of assessment, especially when I use both paper and computer in my classes. I also realized that students are looking to me to guide them in both their L2 and technological literacies. At times, students fully uphold academic integrity, such as during synchronous, in-person assessments on both paper and computer. On the other hand, when they are unsupervised, such as in their homework, they readily fall into the trap that technology can produce L2 faster and supposedly better for them, compromising not only their honesty, but also L2 learning.

CONCLUSION

These observations of research and current practice will hopefully motivate other L2 instructors to make conscientious decisions about the design and implementation of assessments in different media and reflect on their own practice. Ultimately, even if there is no difference in preference or scores between paper- and computer-based assessments, L2 instructors must still question what is gained or lost in the use of one medium over another before finding superiority in either, even when they use both in class. They must also question to what extent they are assessing their students' natural language production versus technology-assisted language production.

In order to further analyze the relative benefits and challenges of paper- and computer-based assessments, I offer a few directions for researchers. First, future studies need to measure students' literacies of diverse technologies at different age cohorts and L2 levels and compare them with their L2 instructors' technological literacies and preferences. Second, comparative studies across languages could consider media of input that students use in their other classes and measure effects that are L2-dependent. Data on different forms of electronic assessments as well as on students' technological literacy between L1 and L2 environments would promote even richer comparison. Similarly, comparison of L2 student notes taken in multiple modes could yield deep textual, idiographic, logographic, and semiotic analyses as well as tests for their effectiveness in later recall. Third, increased collaboration between researchers and practitioners in evidence-based pedagogy could prove effective for both parties (Sato & Loewen, 2022, 2019). The benefits of classroom-based studies have been shown to inspire future Second Language Acquisition or SLA theories (Zéphir, 2008) as well as to inform practices that are mutually beneficial to students and teachers (Thompson, 2012). Ethnographies of the teacher-student relationship reveal a great deal about motivation for L2 learning (Henry & Thorsen, 2018); how L2 teachers engage with different technologies themselves and reconcile any disparities with their students' engagement in the target language are also important for future research.

Finally, critical studies on the use of artificial intelligence in composition are crucial to understanding what exactly is lost or gained in L2 learning. Practitioners and researchers alike must therefore seriously consider the wider implications for instruction as well as for assessment given the rapid advances of AI. As technologies move from merely corrective to fully generative, deeper understanding as to how students grow in their L2 literacy along with their technological literacy is of timely import. Chapelle's (2009) claims ring true today: "Technology can bring to language learning materials a novelty and expense, which create an opportunity for multiple forms of rich input and interaction as well as a data collection capacity unknown to authors of paper materials." Based on my observations, however, it would be premature to write off, so to speak, paper-based assessments in the L2 classroom.

REFERENCES

Araújo, S., Domingues, M., & Fernandes, T. (2022). From hand to eye: A meta-analysis of the benefit from handwriting training in visual graph recognition. *Educational Psychology Review*, 34(3), 1577–1612. https://doi.org/10.1007/s10648-021-09651-4

Barbour, B. (2021, July 13). *On the advantages of paper Notebooks*. Edutopia. https://www.edutopia.org/article/advantages-paper-notebooks

Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Journal*, 100(1), 320–340. https://doi.org/10.1111/modl.12316

Chapelle, C. A. (2009). The relationship between second language acquisition theory and computer-assisted language learning. *The Modern Language Journal*, 93(s1), 741–753. https://doi.org/10.1111/j.1540-4781.2009.00970.x

Chun, D., Kern, R., & Smith, B. (2016). Technology in language use, language teaching, and language learning. *The Modern Language Journal*, 100(S1), 64–80. https://doi.org/10.1111/modl.12302

Canz, T., Hoffmann, L., & Kania, R. (2020). Presentation-mode effects in large-scale writing assessments. *Assessing Writing*, 45, 100470. https://doi.org/10.1016/j.asw.2020.100470

Faust, D. G. (2022, September 16). Gen Z never learned to read cursive. *The Atlantic*. https://www.theatlantic.com/magazine/archive/2022/10/gen-z-handwriting-teaching-cursive-history/671246/

Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, 40(2), 97–118. https://doi.org/10.1017/S0261444807004144

Hall, A. H., Simpson, A., Guo, Y., & Wang, S. (2015). Examining the effects of preschool writing instruction on emergent literacy skills: A systematic review of the literature. *Literacy Research and Instruction*, 54(2), 115–134. https://doi.org/10.1080/19388071.2014.991883

Haring, D., & Kelner, T. (2021, May 1). The duel between the pen and keyboard continues. ASCD, 78(8). https://www.ascd.org/el/articles/the-duel-between-the-pen-and-keyboard-continues

Hellmich, E. A., & Vinall, K. (2023). Student use and instructor beliefs: Machine translation in language education. *Language Learning & Technology*, 27(1), 1–27. https://hdl.handle.net/10125/73525

Henry, A., & Thorsen, C. (2018). Teacher-student relationships and L2 motivation. *The Modern Language Journal*, 102(1), 218–241. https://doi.org/10.1111/modl.12446

James, K. H. (2017). The importance of handwriting experience on the development of the literate brain. *Current Directions in Psychological Science*, 26(6), 502–508. https://doi.org/10.1177/0963721417709821

James, K. H., & Engelhardt, L. (2012). The effects of handwriting experience on functional brain development in pre-literate children. *Trends in Neuroscience and Education*, 1(1), 32–42. https://doi.org/10.1016/j.tine.2012.08.001

Jenkins, H., Purushotma, R., Weigel, M., Clinton, K., & Robison, A. J. (2009). Confronting the challenges of participatory culture: Media education for the 21st century. The MIT Press. https://doi.org/10.7551/mitpress/8435.001.0001

Kern, R. (2015). *Language, literacy, and technology* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139567701

Kim, H. R., Bowles, M., Yan, X., & Chung, S. J. (2018). Examining the comparability between paper- and computer-based versions of an integrated writing placement test. *Assessing Writing*, 36, 49–62. https://doi.org/10.1016/j.asw.2018.03.006

Klein, J., & Taub, D. (2005). The effect of variations in handwriting and print on evaluation of student essays. *Assessing Writing*, 10(2), 134–148. https://doi.org/10.1016/j.asw.2005.05.002

Li, Z., Dursun, A., & Hegelheimer, V. (2017). Technology and L2 writing. In C. A. Chapelle & S. Sauro (Eds.), *The Handbook of Technology and Second Language Teaching and Learning* (1st ed., pp. 77–92). Wiley. https://doi.org/10.1002/9781118914069.ch6

Longcamp, M., Boucard, C., Gilhodes, J.-C., Anton, J.-L., Roth, M., Nazarian, B., & Velay, J.-L. (2008). Learning through hand- or typewriting influences visual recognition of new graphic shapes: Behavioral and functional imaging evidence. *Journal of Cognitive Neuroscience*, 20(5), 802–815. https://doi.org/10.1162/jocn.2008.20504

Longcamp, M., Hlushchuk, Y., & Hari, R. (2011). What differs in visual recognition of handwritten vs. printed letters? An fMRI study. *Human Brain Mapping*, 32(8), 1250–1259. https://doi.org/10.1002/hbm.21105

Marquardt, C., Diaz Meyer, M., Schneider, M., & Hilgemann, R. (2016). Learning handwriting at school – A teachers' survey on actual problems and future options. *Trends in Neuroscience and Education*, 5(3), 82–89. https://doi.org/10.1016/j.tine.2016.07.001

Mayer, C., Wallner, S., Budde-Spengler, N., Braunert, S., Arndt, P. A., & Kiefer, M. (2020). Literacy training of kindergarten children with pencil, keyboard or tablet stylus: The influence of the writing tool on reading and writing performance at the letter and word level. *Frontiers in Psychology*, 10, 3054. https://doi.org/10.3389/fpsyg.2019.03054

Morehead, K., Dunlosky, J., & Rawson, K. A. (2019). How much mightier is the pen than the keyboard for note-taking? A replication and extension of Mueller and Oppenheimer (2014). *Educational Psychology Review*, 31(3), 753–780. https://doi.org/10.1007/s10648-019-09468-2

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science*, 25(6), 1159–1168. https://doi.org/10.1177/0956797614524581

New, E. (2002). Computer-aided writing in French as a foreign language: A qualitative and quantitative look at the process of revision. *The Modern Language Journal*, 83(1), 81–97. https://doi.org/10.1111/0026-7902.00007

Nikolopoulou, K. (2020). Secondary education teachers' perceptions of mobile phone and tablet use in classrooms: Benefits, constraints and concerns. *Journal of Computers in Education*, 7(2), 257–275. https://doi.org/10.1007/s40692-020-00156-7

Polisca, E., Stollhans, S., Bardot, R., & Rollet, C. (2022). How Covid-19 has changed language assessments in higher education: A practitioner's view. In C. Hampton & S. Salin (Eds.), *Innovative Language Teaching and Learning at University: Facilitating Transition from and to Higher Education* (pp. 81–91).

Powers, D. E., Fowles, M. E., Farnum, M., & Ramsey, P. (1994). Will they think less of my handwritten essay if others word process theirs? Effects on essay scores of intermingling handwritten and word-processed essays. *Journal of Educational Measurement*, 31(3), 220–233.

Rogers, J., & Case-Smith, J. (2002). Relationships between handwriting and keyboarding performance of sixth-grade students. *The American Journal of Occupational Therapy*, 56(1), 34–39. https://doi.org/10.5014/ajot.56.1.34

Russell, M., & Tao, W. (2004). The influence of computer-print on rater scores. *Practical Assessment, Research, and Evaluation*, 9(10), 1–3. https://doi.org/10.7275/2EFE-TS97

Santangelo, T., & Graham, S. (2016). A comprehensive meta-analysis of handwriting instruction. *Educational Psychology Review*, 28(2), 225–265. https://doi.org/10.1007/s10648-015-9335-1

Sato, M., & Loewen, S. (Eds.). (2019). *Evidence-based second language pedagogy: A collection of instructed second language acquisition studies* (1st ed.). Routledge. https://doi.org/10.4324/9781351190558

Sato, M., & Loewen, S. (2022). The research–practice dialogue in second language learning and teaching: Past, present, and future. *The Modern Language Journal*, 106(3), 509–527. https://doi.org/10.1111/modl.12791

Sedor, N. (2023). Integrated performance assessment (IPA): Implementation, task types, and feedback. *Foreign Language Annals*, 56(1), 170–190. https://doi.org/10.1111/flan.12661

Sinks, T. & Thurston, J. (1972). Effect of typing on school achievement in elementary grades. *Educational Leadership*, 29, 344–348.

Støle, H., Mangen, A., & Schwippert, K. (2020). Assessing children's reading comprehension on paper and screen: A mode-effect study. *Computers & Education*, 151, 103861. https://doi.org/10.1016/j.compedu.2020.103861

Surahman, E., & Wang, T. (2022). Academic dishonesty and trustworthy assessment in online learning: A systematic literature review. *Journal of Computer Assisted Learning*, 38(6), 1535–1553. https://doi.org/10.1111/jcal.12708

Sweedler-Brown, C. O. (1991). Computers and assessment: The effect of typing versus handwriting on the holistic scoring of essays. *Research and Teaching in Developmental Education*, 8(1), 5–14. https://www.jstor.org/stable/42801814

Tatum, H., & Schwartz, B. M. (2017). Honor codes: Evidence based strategies for improving academic integrity. *Theory Into Practice*, 56(2), 129–135. https://doi.org/10.1080/00405841.2017.1308175

Thomas, K. E. (2018). Comparing explicit exemplar-based and rule-based corrective feedback: Introducing analogy-based corrective feedback. *The Modern Language Journal*, 102(2), 371–391. https://doi.org/10.1111/modl.12470

Thorne, S. L., & Reinhardt, J. (2008). "Bridging activities," new media literacies and advanced foreign language proficiency. *CALICO Journal*, 25(3), 558–572.

Vasylets, O., & Marín, J. (2022). Pen-and-paper versus computer-mediated writing modality as a new dimension of task complexity. *Languages*, 7(3), 195. https://doi.org/10.3390/languages7030195

Wolfe, E. W., & Manalo, J. R. (2004). Composition medium comparability in a direct writing assessment of non-native English speakers. *Language Learning & Technology*, 8(1), 53–65.

Wollscheid, S., Sjaastad, J., Tømte, C., & Løver, N. (2016). The effect of pen and paper or tablet computer on early writing – A pilot study. *Computers & Education*, 98, 70–80. https://doi.org/10.1016/j.compedu.2016.03.008

Zheng, B., & Warschauer, M. (2017). Epilogue: Second language writing in the age of computer-mediated communication. *Journal of Second Language Writing*, 36, 61–67. https://doi.org/10.1016/j.jslw.2017.05.014