

2015

Paper-Based or Computer-Based Essay Writing: Differences in Performance and Perception

Benjamin Kohler
St. Cloud State University

Follow this and additional works at: https://repository.stcloudstate.edu/stcloud_ling



Part of the [Applied Linguistics Commons](#)

Recommended Citation

Kohler, Benjamin (2015) "Paper-Based or Computer-Based Essay Writing: Differences in Performance and Perception," *Linguistic Portfolios*: Vol. 4 , Article 13.

Available at: https://repository.stcloudstate.edu/stcloud_ling/vol4/iss1/13

This Article is brought to you for free and open access by theRepository at St. Cloud State. It has been accepted for inclusion in Linguistic Portfolios by an authorized editor of theRepository at St. Cloud State. For more information, please contact rswexelbaum@stcloudstate.edu.

PAPER-BASED OR COMPUTER-BASED ESSAY WRITING: DIFFERENCES IN PERFORMANCE AND PERCEPTION

BENJAMIN T. KOHLER

ABSTRACT

The purpose of this investigation is to gather and compare data as to students' performances, processes, and perceptions of computer-based writing assessments versus paper-based writing assessments within the context of a university ESL program in the hopes of providing further support and considerations for implementation of computer-based writing assessments for ESL programs. Other learner variables, such as computer experience and typing ability, are compared to see if they have a measurable impact on students' products, processes, or perceptions of writing in both modes. The raters also provided evidence for their reactions to scoring written texts in various modes, as scoring is also a significant factor when verifying the comparability of written assessments. Data for students' and raters' performances were gathered through essay scores; data for students' processes were gathered through online questionnaires and also stimulated recall interviews; and data for students' and raters' perceptions were gathered through online questionnaires. Though results are limited in scope due to small sample sizes, the performances of ESL students in this study infer that the majority of them are not yet proficient enough in writing computer-based essays to substantiate obligating them to utilize this medium for high-stakes testing. Providing ESL students with a choice of medium is also not recommended due to discrepancies in students' perceptions of their abilities and also due to raters' scoring of essays across both modes.

1.0 Paper-based versus Computer-based Writing Assessments

The large Midwestern university that is the location and context for this investigation uses a computer-based medium for its ESL placement testing of reading and listening skills, though it requires the same participants who take those computer-based exams to perform writing assessments (placement essays) using a paper-based medium. This university, like many others, finds itself in the middle of a transition between electronic and paper mediums for its ESL assessments. ESL instructors at this university frequently ask students to submit electronic or computer-typed assignments and essays as part of standard ESL-program curricular requirements. Students at this particular university, like students at many other US universities and other universities around the world, are increasingly expected to create and submit written assignments on the computer as opposed to handwriting them. Juxtaposition may then exist between what and how students are expected to perform in practice and what and how they are expected to perform in assessment. Nonetheless, implementation of this type of change requires a great deal of preparation and consideration in order for it to be integrated successfully into programmatic curricula (Harrington, Shermis, & Rollins, 2000).

While this change in writing mode may not be new, and it may not be consistent the whole world over, it is clear that computers, mobile technologies, and the Internet are

changing the way people today write and produce text, and this influence is quickly spreading. The creation of electronic text is starting to become a more normalized medium for writing, especially in contexts such as US universities. This trend puts into question the authenticity of using paper-based writing assessments over computer-based ones, when the majority of students' writing is now produced by typing in electronic text (Endres, 2012). More and more test administrators and evaluators will have to continue to research the complicated implications that would come with implementation of this significant change in medium of assessment, for the way in which it is implemented may also have a great impact on students and assessment.

1.1 Research Questions

This investigation hopes to provide the context of this university with further support and considerations for implementation of computer-based writing assessment for its ESL program. The research questions for this thesis paper are the following:

1. How do ESL students' writing performance and their perceptions of their writing performance differ when writing on paper and on the computer?
2. How do ESL students' writing processes and their perceptions of those writing processes differ when writing on paper and on the computer?
3. How do independent variables (computer or typing experience and typing proficiency) affect students' written products and or writing processes?
4. How do raters' scores and perceptions differ when scoring handwritten texts, transcribed texts and typed texts?

The purpose of the current investigation was to gather and compare data as to students' performances, processes, and perceptions of computer-based writing assessments versus paper-based writing assessments within the context of the university's ESL program. Other learner variables, such as computer experience and typing ability, were also compared to see if they had a measurable impact on students' products, processes, or perceptions. The raters in this study also provided evidence for their reactions to scoring written texts in various modes, as scoring is also a significant factor when verifying the comparability of written assessments.

2.0 Participants

Participating in this study were eight intermediate-level ESL students from a large university in the Midwestern United States (the primary group). All students had been placed into the same ESL course that focuses on academic reading and writing skills based on international English proficiency exams and subsequent English language placement tests taken prior to the semester, including an Accuplacer ESL Reading test and an ESL essay test.

One additional student (the stimulated recall subgroup) was also video-recorded while producing a paper-based essay and a computer-based essay. The recorded participant came from an ESL course of a different section from that of the primary group taught by me.

Six ESL instructors participated in this investigation by scoring participant essays and also answering a questionnaire based on the scoring of those essays. Background

information from these participants was not gathered, though they had a variety of different teaching experiences between them. They also included a mix of males and females and native and non-native English speakers.

3.0 Materials

For each essay, participants responded to a short prompt. There were two different prompts, each directing the participants to write an essay of the argumentative genre. The instructions in both essay prompts were exactly the same, and participants were asked to respond to whether they agreed or disagreed with a statement about a topic that was familiar to them. The topic of each argumentative essay was distinct, but both were meant to allow participants to write freely without the necessity of specific background information or experience. To that end, both essays, the first labeled the Cellphone Essay and the second the ESL Essay, were related to classroom policies.

At the time of this study, all instructors at the university had the capacity to utilize a common learning management system (LMS) in conjunction with the course that they taught. All participants in this study wrote an essay on the computer through use of LMS's quiz function. The spell-checker function of the LMS quizzes was effectively deactivated for the purposes of this investigation. Participants had the options to block cut, copy, paste, and delete text; they were also able to undo or redo completed actions. Participants were not allowed to use any external resources (online or otherwise) while completing the essays in the LMS.

Four different questionnaires were used for the purposes of this investigation. All four were created and administered with an online survey software and questionnaire tool called SurveyMonkey (<http://surveymonkey.com>). Links in this document are provided to the online SurveyMonkey questionnaires.

The first questionnaire retrospectively gathered data as to participants' perceptions of writing both their first and their second essay. This questionnaire, labeled the Essay Questionnaire (<https://www.surveymonkey.com/s/3B8LMCV>), was the same for the first essay and the second essay, whether it was a paper-based essay or computer-based essay. The Essay Questionnaire contained 28 randomized, Likert-scale items and three open-ended, short answer questions that asked students what they did before, during, and after writing their essays.

The second retrospective questionnaire gathered data with regard to a comparison and perception of writing with the two mediums. This was labeled the Comparison Questionnaire (<https://www.surveymonkey.com/s/T8XBRRB>). It contained the same 28 randomized, Likert-scale items as the Essay Questionnaire, though the true number of items was doubled, as two scales (one for the paper-based essay and the other for the computer-based essay) were provided for each question. Included were also four open-ended, short answer questions that asked students more directly which writing mode they preferred and why.

The third questionnaire, labeled the Student Profile Questionnaire (<https://www.surveymonkey.com/s/PCYW9NR>), gathered information detailing the

demographics of the participants, as well as the history of their English language education and experience and also their computer education and experience.

The fourth questionnaire, labeled the Rater Questionnaire, gathered information about raters' perceptions on the differences between scoring handwritten essays and typed essays. This short questionnaire consisted of three open-ended, short answer questions that asked raters about their perceptions of scoring essays in both writing modes. It was printed and participants completed it with pens or pencils.

Finally, participants' scores on a keyboarding test from the Internet (<http://www.speedtypingonline.com/typing-test>) were recorded based on their speed (calculated in words per minute—WPM) and accuracy.

4.0 Participant Training

The eight primary group participants were trained to perform the task of writing essays both on paper and on the computer within the context of the LMS quiz function. Two weeks prior to writing the essays for this study, at the fourth week of the semester, participants went to a university computer lab during class-time to perform 30-minute writing tasks. The participants spent two sessions of 50 minutes each (half of the class) training in the computer lab. Some participants performed writing tasks on paper while others performed them on the computer. The participants became accustomed to sessions in which half the class was writing on paper while the other half wrote on the computer. The participants familiarized themselves with the function of the LMS quizzes and also the 30-minute time limit for writing essays.

4.1 Writing the Essays

Following the two sessions of training, the two essays for this study were written by primary group participants over the course of two writing sessions (labeled Session 1 and Session 2). Each writing session took place in a separate class period with two days between them; but both class periods fell within the fifth week of the semester.

The participants were randomly divided into four groups (labeled Group A, Group B, Group C, and Group D) of two participants each. Essay writing sessions were timed, allowing participants 30 minutes to complete each essay. The LMS quiz set a visible timer for participants to see while they were performing the computer-based essay. At the end of the 30 minutes, the LMS quiz automatically terminated, saving participants' work. I personally monitored the time and collected paper-based essays after exactly 30 minutes. The participants were not allowed the use of dictionaries, spell-checkers, grammar-checkers, or any other outside resources.

On the day of Session 1, Group A and Group B both wrote essays on paper. Group A completed the Cellphone Essay and Group B completed the ESL Essay. Groups C and D both wrote essays on the computer. Group C completed the Cellphone Essay and Group D completed the ESL Essay.

On the day of Session 2, all roles alternated. Group A and Group B both wrote essays on the computer. Group A completed the ESL Essay and Group B completed the

Cellphone Essay. Groups C and D both wrote essays on paper. Group C completed the ESL Essay and Group D completed the Cellphone Essay.

The following table (Table 1) shows the Within Subject Design for Sessions 1 and 2. Under the headings ‘Paper-based’ and ‘Computer-based’ are the labeled groups which wrote either the Cellphone Essay or the ESL Essay during Session 1 or Session 2.

	Paper-based	Computer-based
Session 1	Group A Cellphone Essay	Group C Cellphone Essay
	Group B ESL Essay	Group D ESL Essay
Session 2	Group D Cellphone Essay	Group B Cellphone Essay
	Group C ESL Essay	Group A ESL Essay

Table 1: Within Subject Design

With the Within Subject Design of this investigation, each participant completed two essays. One of the essays was completed on a paper-based medium and the other on a computer-based medium. All participants also completed both the Cellphone Essay and the ESL Essay.

4.2 Scoring the Essays

All paper-based essays were transcribed by the researcher into typed text, maintaining all the errors and formatting of the original handwritten text. There were three types of essays that were scored: paper-based essays (the original handwritten essays), transcribed essays (the handwritten essays typed on the computer), and computer-based essays (the original typed essays). Transcribed essays were made to look the same as computer-based essays and had text that was of the same font and size. The primary-group participants were divided and labeled in the following groups:

Groups	Participants
Group A	5 & 6
Group B	7 & 8
Group C	1 & 2
Group D	3 & 4

Table 2: Primary-Group Participant Labels

Each primary-group participant had three essays that corresponded to them: a paper-based essay (PB), a transcribed essay (TR), and a computer-based essay (CB). The six raters scored all of the essays in one single session. For the eight primary-group participants, there were a total of 24 essays that were scored and calculated. Six

additional essays that were authored by ESL students from the same course who did not participate in the study were also scored. This was also done so that non-participant students could fulfill their class writing assignment, receive feedback on their writing, and benefit from having their essays scored by raters. The rater scores of these extra essays were not combined with essay mean scores, meaning that these essays scores are not included anywhere in this report, nor is data pertaining to these non-participant students included in any way in this investigation.

The next table (Table 3) demonstrates the 24 essays that corresponded to the eight primary group participants (labeled 1 through 8 for CB, PB, and TR) and also six extra essays that corresponded to non-participant students (PB9, TR9, PB10, TR10, PB13, and PB 14).

Essay Order	Rater #1	Rater #2	Rater #3	Rater #4	Rater #5	Rater #6
1:	TR1	CB8	TR3	TR4	CB2	CB1
2:	PB3	PB1	PB2	PB7	PB14	PB8
3:	CB7	TR2	CB4	CB3	TR5	TR6
4:	PB9	PB4	PB5	PB10	PB6	PB13
5:	CB6	CB5	TR10	TR9	TR8	TR7

Table 3: Essay Scoring

I prepared the essays in pamphlets for each rater in order to establish and maintain a specific order for essay scoring. The previous table (Table 3) demonstrates the order in which raters scored the essays so as to control for the order effect when scoring the three types of essays. The first essay all raters scored was a typed essay (either computer-based or transcribed). The second essay was handwritten; the third was typed (CB or TR); the fourth was handwritten; and the fifth was typed (CB or TR). With this particular order, and including the extra essays, each rater scored a total of five essays, three of which were typed and two of which were handwritten.

Raters scored two essays together at the beginning of the session in order to calibrate their scoring. One of these essays was handwritten and the other typed. The two essays used for calibration were also extra essays from non-participant students. Data from these two essays was not included nor calculated with primary group or rater statistics.

I explained to the raters that some of the essays they scored would be handwritten and others would be typed. Though the raters were asked to score both handwritten and typed essays, they remained blind as to the precise research questions of the investigation.

Participants' essays were analyzed by quantity of words written and also a quality score they received based on an adjusted grading rubric that is used for grading the ESL essay tests used for university English-language placement (Jacobs, Zingraf, Wormuth, Hartfiel, & Hughey, 1981). The rubric assigns a numeric grade to each essay based on five separate categories, which are 'Content,' 'Organization,' 'Vocabulary,' 'Language Use,' and 'Mechanics'. These categories combine to create an 'Overall' essay score. Each essay was scored exactly one time.

4.3 Completing the Questionnaires

The questionnaire categories were created to match those of aforementioned ESL Composition Rubric. This was done so that student perceptions could be mapped against their actual performances. The 28 Likert items were divided into four items for each of the five writing categories, as well as including an overall category and one more related to quantity of writing.

All primary group participants completed the Essay Questionnaire immediately after they finished writing their essays for Session 1, whether they wrote a paper-based or computer-based essay. The participants had approximately 15 to 20 minutes of class time to fill out their online questionnaires.

Several days later, immediately after the participants wrote their essays for Session 2, they then completed the Essay Questionnaire again. Completion of the Essay Questionnaire during Session 2 was carried out in the exact same manner as in Session 1. After Session 2, the participants had each completed two Essay Questionnaires, one of which corresponded with a paper-based essay and the other with a computer-based essay.

The first class period following Session 2 took place the next week, five days later. The participants completed the Comparison Questionnaire and also the Student Profile Questionnaire at that time, taking around 20 to 30 minutes of class time to complete these online questionnaires.

The Rater Questionnaires were completed by the six raters immediately after they scored the essays.

4.4 Completing the Keyboarding Tests

Along with the Comparison Questionnaire and the Student Profile Questionnaire, primary-group participants also completed the online keyboarding test during the first class period following Session 2. The participants were allowed a trial test or two before the researcher recorded their results on the one-minute keyboarding test.

4.5 Stimulated Recall Protocols

The subgroup participant who took part in the stimulated recall interviews also completed two essays in the same manner as the previous primary group participants who participated in Sessions 1 and 2. In this case, however, I set up a video recorder to record the paper and the computer screen while the essays were being written.

During Subgroup Session 1, after the participant had been video-recorded handwriting an essay, a stimulated recall protocol was carried out in which the participant watched the video with me on my personal laptop computer and commented on what his thought-processes were during the writing process. My questions to the participant had to do with what he was thinking when pausing and revising during the writing process.

5.0 Results and Discussion

I will base the discussion points of this investigation on the proposed research questions. Implications for the results found in this study will be drawn for the context of this particular US university and others like it. Limitations of this study and also possibilities for further research will also be included within this section.

5.1 Research Question #1

How do ESL students' writing performance and their perceptions of their writing performance differ when writing on paper and on the computer?

Like participants in other studies (Goldberg, Russell, & Cook, 2003; Harrington, et. al., 2000; Lee, 2004; Moge, Paterson, Burk, & Purcell, 2010; Whithaus, Harrison, & Midyette, 2008), ESL students in this study tended to think that they would score better (albeit barely) on computer-based essays than on paper-based ones based on their Likert responses to Essay and Comparison Questionnaires. Again, like researchers had found before (Endres, 2012), this perception of greater ability in computer-based writing seemed to be a false one, as means of ESL students' holistic handwritten essay scores were 85.75 while their typed scores were only 81. Though the difference between these scores was not statistically significant, the discrepancy between students' common perceptions and their results is important. When responding to short answer questions about personal preference, however, students were much less certain that they could achieve better results with computer-based writing. Some, in fact, admitted that with experience and training, differences in their perceptions and performances between writing modes were actually changing.

This is to say that as students gain experience and training with computer-based writing, both their performances and perceptions of that performance are improving, also narrowing whatever gap may exist between their performance and perception. Though students might never actually receive training to recognize the differences between their performances in the different writing modes, certainly further exposure will help them gain enough knowledge and experience to be confident in both (Endres, 2012; Moge, Cowan, Paterson, & Purcell, 2012). A choice, at some point, between writing modes might actually be a less significant decision for students given that they may be justifiably confident in their abilities for both. For the present time, however, it seems that many of the ESL students in this study do not yet have the experience or knowledge (as shown by their test scores) or well-placed confidence in their typing abilities (as shown by the discrepancy between their computer-based test scores and their high perceptions of them) to be able to accurately choose the writing medium that will allow them to perform to the best of their abilities.

Because ESL students' scores in handwritten essays were better than their typed essays, it can be assumed that perhaps the majority of these students produce better quality writing in a paper-based mode than a computer-based mode. Though it is difficult to generalize these results for a much wider audience, it could be reasonably theorized that ESL students' capacity for computer-based writing is still underdeveloped compared to their capacity for paper-based writing. Since, for this study, the mean difference between the holistic scores in both modes was not significant, however, we might speculate that this gap is one that is becoming less and less pronounced as students continue to improve their computer-based writing performance. A lack of statistically significant differences between analytical (or categorical) scores in both writing modes shows also that if a gap exists between the quality of ESL students' typing and handwriting abilities, it is more likely to be narrowing than widening across all writing categories.

In just one area was there nearly a statistically significant negative correlation between a paper-based categorical score and a computer-based one: that being Organization. Mean scores for computer-based Organization ($M = 16.88$, $SD = 1.55$) were consistently lower when scores for paper-based Organization were higher ($M = 17.75$, $SD = 0.71$), which implies that the quality of ESL students' organization of writing is not consistent between writing modes. One possibility could be that the quantity of words affects the Organization score of participants' essays. Quantity of words and paragraphs typed proved to be more significant positive correlations of high essay scores than were quantity of words and paragraphs handwritten. Though the difference between quantities of words in both modes was minimal, the difference between paragraphs was more considerable, with ESL students having handwritten more paragraphs ($M = 4.4$, $SD = 0.92$) than they typed ($M = 3.6$, $SD = 0.74$). This consideration will be further developed in accordance with ESL students' writing processes and Research Question #2.

5.2 Research Question #2

How do ESL students' writing processes and their perceptions of those writing processes differ when writing on paper and on the computer?

The quality scores of ESL students' essays in this study provide us with important considerations for determining a difference in their physical processes of handwriting and typing. Average essay scores for Content, Organization, and Mechanics were higher for paper-based essays, while computer-based essays scored higher in Vocabulary and Language Use. Differences between these scores may not have been statistically significant, but this minimal difference in writing performance could represent a more considerable difference in writing process between modes. Computer-based writing could have a more positive effect on writers' attention to form and accuracy, while paper-based writing could have a more positive effect on writers' attention to content and meaning.

Qualitative evidence from the stimulated recalls, for example, showed that the subgroup participant was aware that he was able to create more content and better organization with a paper-based essay (377 words and 5 paragraphs) than with a

computer-based essay (266 words and 2 paragraphs). Though quantity of typed words and paragraphs for the primary group had a strong positive correlation toward computer-based scores of Content, Organization (for word count, not for paragraph count), and Language Use, quantity of handwritten words and paragraphs did not correspond with better scores in any category. This may have been due to the increased probability that handwriters would create more paragraphs, thus eliminating any discrepancy between number of paragraphs and scores. Typists, on the other hand, were less likely to create more paragraphs, and so the difference in scores was more pronounced when they did produce more.

The fact that paper-based essays contained more paragraphs than computer-based essays could contribute toward better scores in Content and Organization. It could be that ESL students typed fewer paragraphs than they handwrote (even though they typed more words than they handwrote) because they spent more time revising their computer-based essays than their paper-based ones. The fact that they produced more handwritten paragraphs could be due to less time spent revising and more time spent developing the content and organization. Though the relationship between Organization scores and quantity of paragraphs was insignificant in both writing modes, this relationship was much stronger for computer-based essays, $r(8) = 0.572$, $p = 0.14$, than for paper-based ones, $r(8) = -0.055$, $p = 0.90$.

If ESL students' computer-based Content and Organization scores suffered because of lack of words and paragraphs, it may have been due to the fact that, while typing, they were preoccupied with improving the form and accuracy of their language use and vocabulary. In fact, participants were more likely to say that they reread and revised their writing while typing than while handwriting. The stimulated recalls demonstrated, for example, that the subgroup participant made more revisions overall while typing, especially with regard to spelling. Frequent spelling revisions were due to lower typing ability, and thus they are more likely to affect less efficient and less experienced typists (a consideration of typing ability will be discussed more completely with Research Question #3).

Computer-based essays did generate slightly better scores for Vocabulary and Language Use than paper-based essays, although not for Mechanics. The category of Mechanics contained elements of spelling, punctuation, and paragraphing. As will be discussed later in Research Question #4, although raters assumed that typed essays had the benefit of spell check, they were not able to specify whether one mode or the other had more spelling mistakes. The likely cause of a higher Mechanics score for paper-based essays seems then due to the higher number of paragraphs compared to the computer-based essays.

The fact that the subgroup participant made more revisions that were related to the content and organization of his essay while typing than he did while handwriting could have been a common trend for other participants, though we cannot be sure of this according to their survey responses. It does seem plausible, however, that the increased ability to move, add, and erase electronic text increases the probability that ESL students

will revise their essays more often and for more time while they are engaged in the process of typing as opposed to while they are handwriting. Revisions related to content and organization, though fewer in number, often took more time than those related to form and accuracy. In total, a propensity toward revisions of any type is likely to prolong the writing process, which, during a timed essay assessment, is not likely to benefit students' scores related to content and organization. As content and organization typically account for a large percentage of writers' overall scores (a combined 40% on the rubric used to score this study's participants), physical differences between handwriting and typing could also account for significant differences in writing performance and scores of that performance.

Lastly, it must be noted that the subgroup participant may have shed light on another likely probability that could account for participants' lower scores of Content and Organization for computer-based essays. It has already been established that typed essays may have allowed participants to revise for a longer time and more frequently, but it must also be described that typed revisions may take them longer because of the techniques they utilize to make the revisions. It would seem that for the subgroup participant, and possibly for other participants with similar computer-proficiency and typing levels, lack of experience may obligate them to use inefficient techniques to add or remove electronic text. The subgroup participant did not block select text in order to erase, cut, or paste at all while typing, though this is something that may actually help writers with their organization that is not as easy or clean when handwriting. The subgroup participant also scrolled through the text using the keyboard's arrow keys rather than using the mouse to move the cursor, which is something that also might have saved him time. This student was slowed down even more because he tended to write and erase text many times, which was quite inefficient when it came to erasing entire words in order to change the spelling or capitalization of one or two letters in them. The undo and redo functions, which could have assisted the student in this respect, were also not employed. If it can be supposed that other inexperienced or less proficient typists employed similar techniques with their writing processes of computer-based essays, then it would come as no surprise that their scores for Content and Organization and Mechanics (insofar as paragraphs are concerned) suffered.

5.3 Research Question #3

How do independent variables (computer or typing experience and typing proficiency) affect students' written products and or writing processes?

As explained previously, it could be reasonably assumed that the handwriting fluency of the ESL students is better than their keyboarding fluency. We might make this assumption based on several factors, including higher paper-based essay scores of quality compared with computer-based scores, a low mean WPM score for participants, and responses from participants claiming they were faster, more familiar and more comfortable with handwriting than typing. According to a 'simple view of writing', as reproduced by Connelly, Gee, and Walsh (2007), lack of fluency in lower order cognitive processes such as keyboarding or handwriting constrain higher order cognitive processes such as planning and reviewing. To this end, it might make sense that less fluent typists

would be forced to spend more time on lower order processes as opposed to higher order processes that have to do with the content and organization of their ideas in essays.

Mean results from this investigation, however, found no relationships between participants' computer-based scores of content and organization and their WPM scores. Though no significant relationship was found between mean holistic computer-based scores and WPM, a more specific analysis including the variable of keyboarding lessons led to a significant discovery. Students who had had keyboarding lessons were more likely to have higher WPM scores than were students who had not had keyboarding lessons. Those same students who had had lessons were also more likely to have scored better in their computer-based essay than were students who had not had lessons. The two students with the highest WPM scores were the only two students who achieved better holistic computer-based essay scores than paper-based ones. Not surprisingly, both of these students had also had keyboarding lessons, adding to the positive effect that this type of training had on not only the proficiency of their computer-based writing, but the quality of it as well. These results can be observed in Table 4 below.

Participant	Keyboarding lessons	WPM	Essay scores
#4	Yes (in other language)	50	Better CB
#1	Yes (in other language)	48	Better CB
#6	Yes (in English)	36	Better PB
#8	Yes (in other language)	34	Better PB
#3	No	31	Better PB
#5	No	27	Better PB
#7	No	21	Same
#2	No	18	Better PB

Table 4: Individual Representation of Participants' scores

Students who had not had keyboarding lessons were more likely to score worse on their computer-based essays than their paper-based essays. The subgroup participant, for example, also showed that a discrepancy in his typing fluency had a considerable effect on his ability to create a good computer-based essay under the time constraints. This participant's results seemed to correspond with those from researchers who found that writers with low proficiency with computers performed worse when typing than when handwriting (Chen, White, McCloskey, Soroui, & Chun, 2011; Wolfe, Bolton, Feltovich, & Niday, 1996; Wolfe & Manolo, 2004). Ironically, this student claimed to have used a computer for a similar amount of time to other participants in the study; although, he did say that he had used a computer for school for less time. Years of computer use and years of computer use for school were not found to be related to essay scores, WPM, or whether or not participants had had keyboarding lessons. This seems to imply that ESL students' experience with computers does not necessarily equate to proficiency in computer-based writing. In fact, based on the results of this study, it might be reasonable to conclude that proficiency in computer-based writing can be more

accurately measured by a student's typing fluency (WPM) and the quality of texts produced rather than by years of computer use.

5.4 Research Question #4

How do raters' scores and perceptions differ when scoring handwritten texts, transcribed texts and typed texts?

The most significant findings from quantitative rater data substantiated results from other researchers (MacCann, Eastment, & Pickering, S., 2002; Moge, et. al., 2010; Russell & Tao, 2004a; Shaw, 2003) that demonstrated that essay raters tended to score handwritten texts more positively than typed ones. This study's results are demonstrated by Table 5 below.

	Raters					
	1	2	3	4	5	6
Typed Holistic Means	85.00	76.33	81.00	78.67	83.00	84.33
Handwritten Holistic Means	90.00	78.50	90.00	84.50	85.50	85.00

Table 5: Difference between Scores of Typed and Handwritten Essays

It was conjectured that this might be due to increased visibility of errors in typed texts and also to the fact that typed texts seemed shorter than their handwritten counterparts (Endres, 2012). The six raters who participated in this study, however, did not come to a consensus about whether more errors occurred in typed essays or handwritten ones. One rater actually thought that the typed essays appeared longer than the handwritten ones, though most raters agreed that there were no differences in errors between modes.

Results showed that handwritten mean scores were greater than typed mean scores for all categories of writing. The most significant differences were between handwritten and typed scores of Content and Organization, as shown in Table 6 below.

	Handwritten Scores	Transcribed Scores	Computer-based Scores	Typed Scores (Transcribed & CB combined)
Content	18.0	17.4	17.1	17.2
Organization	18.1	17.3	16.9	17.2
Vocabulary	21.4	20.4	20.8	20.5
Language Use	23.9	22.8	22.8	22.7
Mechanics	4.2	3.8	3.8	3.7
Holistic	85.6	81.9	81.3	81.4

Table 6: Descriptive Data of Raters' Mean Scores

Possibilities that the writing process and students' typing proficiency are relevant factors that contribute toward a difference in writing performance have already been discussed. An additional possibility is that raters have biases or more tendencies toward scoring handwritten essays more positively than typed ones. The raters in this study were divided as to whether they thought that they scored essays differently or not. Most raters agreed that handwritten essays were more difficult to read and that typed essays were quicker and easier to read. Some felt that this led to quicker and easier scoring, while others felt that they had no prejudice between the modes.

It is noteworthy that while raters believed that it was quicker and easier to work with typed texts rather than the more laborious handwritten texts, they often scored typed texts lower. One possibility, from this researcher's point of view, is that by taking more time to decipher and consider the handwritten writing of ESL students, more attention is paid to the intended meaning of the handwriting and less to structures that might otherwise impede meaning. Because raters said that typed essays were quicker and easier to read, however, raters might spend less time focusing on meaning and more time focusing on language structure, whether it impedes meaning or not. An increased focus on meaning when reading handwritten essays and an increased likelihood of focus applied to form/structure in typed essays implies that raters will be more likely to score handwritten essays higher in Content and Organization and typed essays lower in Vocabulary, Language Use, and Mechanics.

The positive relationships between raters' scores of typed and handwritten Content and also scores of typed and handwritten Organization were found to be consistent, meaning that if they were to have a biased preference for handwritten essays over typed essays, then they all shared this opinion at roughly the same rate. Primary group data showed that typed scores for Vocabulary and Language Use were slightly greater than handwritten scores, while these same categories within the rater data received better handwritten scores. This could mean that the rater data, which consists of

more scores than the primary group data, is pointing toward a more accurate portrayal of performance results or that raters have a tendency to score handwritten essays more positively in all categories. It could be, as predicted, that raters may see mistakes in language form and accuracy more easily when they are typed because they are easier to read. It could also be that raters are more likely to be forgiving with handwritten mistakes because they are less obvious and subject to more issues of legibility and objectivity in interpretation.

Unfortunately the sample size of this study is too small to make strong generalizations one way or the other, and it is clear that further investigation would be required to obtain a clearer, more complete portrayal of raters' discrepancy between scoring typed texts and handwritten texts. A more complete error analysis of the ESL students' essays and raters' interpretation of them across handwritten and typed modes would also give further insight into the factors that are most pertinent to raters when scoring essays in one mode or another.

6.0 Conclusion

This study was meant to provide the context of this particular university with considerations and implications for following what is seen to be a trend in modern-day testing and assessment of ESL writing ability: switching more traditional paper-based writing assessments to computer-based ones. Results from this investigation, however, give reason to believe that the physical differences between typing and handwriting may have a significant effect on ESL students' writing performance. It is the conclusion of this researcher that the performances of ESL students in this study infer that the majority of them are not yet proficient enough in writing computer-based essays to substantiate obligating all ESL students at the university to utilize this medium for high-stakes testing. Therefore, it is my recommendation that the university that is the focus of this investigation continue to use paper-based writing assessments for the purposes of placing students in its ESL program until a higher percentage of these students are proven to have a level of computer-writing proficiency that is at least equal to that of their paper-based writing proficiency. Providing ESL students with a choice of medium is also not recommended due to discrepancies in students' perceptions of their abilities and also raters' scoring of essays across both writing modes.

Though the scope of this study was small, and its small sample size may make inferences toward larger contexts less accurate, I feel that the information in this investigation can provide administrators, instructors, writers, and raters alike with important points to consider when it comes to making choices about choosing and using a computer or a pencil and paper for writing. Hopefully the differences between these modes of writing will continue to be explored and discussed in the future, as a better understanding of the strengths, weaknesses, and differences in computer-based and paper-based writing can ultimately lead toward better utilization and performance in both. The question then changes from asking *when* we should choose paper *or* plastic to *how* we can best use both paper *and* plastic.

ABOUT THE AUTHOR

Benjamin Kohler is a recent graduate of the MA TESL & Applied Linguistics Program at SCSU and is currently an English Lecturer at Qatar University. He earned BAs in English and Spanish from St. John's University in Collegeville, MN. He has extensive experience teaching ESL and EFL to various age groups and levels in the United States, Spain, and Chile. In the future he hopes to continue to teach and perform research in areas such as computer-assisted language learning and second language acquisition. He may be reached at <http://www.linkedin.com/pub/benjamin-kohler/30/a0/87/>

Recommendation: This MA thesis summary was recommended for publication by Dr. Ettien Koffi, PhD, Linguistics Department, St. Cloud State University, St. Cloud, MN. Email: enkoffi@stcloudstate.edu

References

- Chen, J., White, S., McCloskey, M., Soroui, J., & Chun, Y. (2011). Effects of computer versus paper administration of an adult functional writing assessment. *Assessing Writing, 16*(1), 49-71.
- Connelly, V., Gee, D., & Walsh, E. (2007). A comparison of keyboarded and handwritten composition and the relationship with transcription speed. *British Journal of Educational Psychology, 77*, 479-492. doi: 10.1348/000709906X116768
- Endres, H. (2012). A comparability study of computer-based and paper-based writing tests. *Cambridge ESOL: Research Notes, 49*, 26-32.
- Goldberg, A., Russell, M., & Cook, A. (2003). The effect of computers on student writing: A meta analysis of studies from 1992 to 2002. *Journal of Technology Learning and Assessment, 2*, 1-51.
- Harrington, S., Shermis, M.D., & Rollins, A.L. (2000). The influence of word processing on English placement test results. *Computers and Composition, 17*, 197-210.
- Jacobs, H.L., Zingraf, S.A., Wormuth, D.R., Hartfiel, V.F., & Hughey, J.B. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House Publishers, Inc.
- Lee, H.K. (2004). A comparative study of ESL writers' performance in a paper-based and a computer-delivered writing test. *Assessing Writing, 9*(1), 4-26.
- MacCann, R., Eastment, B., & Pickering, S. (2002). Responding to free response examination questions: Computer versus pen and paper. *British Journal of Educational Technology, 33*, 173-88.
- Mogey, N., Cowan, J., Paterson J., & Purcell, M. (2012). Students' choices between typing and handwriting in examinations. *Active Learning in Higher Education, 13*, 117-128. doi: 10.1177/1469787412441297

- Mogey, N., Paterson, J., Burk, J., & Purcell, M. (2010). Typing compared with handwriting for essay examinations at university: Letting the students choose. *ALT-J, Research in Learning Technology*, 18(1), 29-47. doi: 10.1080/09687761003657580
- Russell, M. & Tao, W. (2004a). Effects of handwriting and computer-print on composition scores: A follow-up to Powers, Fowles, Farnum, & Ramsey. *Practical Assessment, Research & Evaluation*, 9(1).
- Shaw, S.K. (2003). Legibility and the rating of second-language writing: The effect on examiners when assessing handwritten and word-processed scripts. *Cambridge ESOL: Research Notes*, 11, 7–10.
- Whithaus, C., Harrison, S.B. & Midyette, J. (2008). Keyboarding compared with handwriting on a high-stakes writing assessment: Student choice of composing medium, raters' perceptions, and text quality. *Assessing Writing*, 13, 4-25.
- Wolfe, E. W., Bolton, S., Feltovich, B., & Niday, D. M. (1996). The influence of student experience with word processors on the quality of essays written for a direct writing assessment. *Assessing Writing*, 3(2), 123-147.
- Wolfe, E.W. & Manalo, J.R. (2004). Composition medium comparability in a direct writing assessment of non-native English speakers. *Language, Learning, and Technology*, 8(1), 53–65.