Measuring the Power of Learning.®

Research Memorandum
ETS RM–18-06

Behavioral Differences Between Retyping, Drafting, and Editing: A Writing Process Analysis

Paul Deane
Amanda Roth
Anna Litz
Vishal Goswami
Fred Steck
Mahlet Lewis
Theresa Richter

July 2018

ETS Research Memorandum Series

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Memorandum series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Memorandum series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

Behavioral Differences Between Retyping, Drafting, and Editing: A Writing Process Analysis

Paul Deane, Amanda Roth, Anna Litz, Vishal Goswami, Fred Steck,
Mahlet Lewis, and Theresa Richter
Educational Testing Service, Princeton, New Jersey

July 2018

Corresponding author: P. Deane, E-mail: pdeane@ets.org

Suggested citation: Deane, P., Roth, A., Litz, A., Goswami, V., Steck, F., Lewis, M., & Richter, T. (2018). *Behavioral differences between retyping, drafting, and editing: A writing process analysis* (Research Memorandum No. RM-18-06). Princeton, NJ: Educational Testing Service.

Action Editor: Don Powers

Reviewers: Randy Bennett and Peter van Rijn

Abstract

Writing process logs (keystroke logs) provide an excellent source of information about how writers have distributed their time and attention across the course of a writing task. However, relatively little is known about how the features that can readily be collected from such logs vary as a result of changes in writing task demands. In this study, we contrast the writing behavior of 463 8th-grade students in a school with low socioecomonic status in the western U.S. under 3 conditions: when they were copy typing (retyping an article), when they were drafting an essay, and when they were editing the essay they had drafted in a previous session. We observed striking differences in the characteristics of the resulting keystroke logs, reflecting differences in the mix of writing processes emphasized in each task. Copy typing was characterized by relatively slow typing, little time spent on long pause behaviors (except between words, when writers would have been scanning the text they were copying), and strictly local editing events. Drafting was characterized by relatively fluent typing, significant amounts of backspacing (reflecting false starts and sentence-level revision and editing), significantly longer pauses at sentence and word boundaries (reflecting idea generation and the process of translating ideas into words), and a moderate amount of time spent jumping to points within the most recently produced sentence to sentence in order to make edits to phrasing. Editing was characterized by very long pauses before jumping to another location to the text to make an edit, with very little time spent typing out individual words, phrases, or sentences. These differences indicate the importance of interpreting keystroke logs in the light of task demands and suggest that different features will be provide significant information about writers' performance, depending on the writing processes most emphasized in a particular literacy task.

Key words: writing, assessment, keystroke, keystroke log, writing process, copy typing, retyping, drafting, editing, revision

Table of Contents

List of Figures

Page

List of Tables

Page

Writing is a complex skill that requires authors to coordinate multiple and often competing writing processes (Alamargot & Chanquoy, 2001; Alamargot & Fayol, 2009; Becker, 2006; Flower & Hayes, 1980; Hayes, 2012; Hayes & Flower, 1980). The literature has suggested that novice writers tend to apply simple "knowledge-telling" strategies in which they retrieve knowledge about a topic from memory and write what they retrieve, whereas expert writers follow a more complex writing process that employs more sophisticated "knowledge-transforming" strategies in which the ideas they present, the way they are organized, and even the knowledge they have about their subject matter develop and change during the course of composition (Bereiter & Scardamalia, 1987; Hayes, 2011, 2012; Kellogg, 2008). However, it may be difficult for struggling writers to adopt more complex writing strategies because simple text-production processes (i.e., translation of ideas into words, transcription of words into text via handwriting or keyboarding) compete with other processes for working memory; as a result, less fluent writers may not be able to access the cognitive resources they need to succeed at complex writing tasks (Berninger, 1999; Kellogg, 1996a, 2001; McCutchen, 1996, 2000, 2011; Olive & Kellogg, 2002; Torrance & Jeffery, 1999). However, use of self-directed writing strategies that address specific subgoals, such as idea generation, organization of ideas, revision of content, or editing for language, can help novice writers improve the quality of their work by separating different parts of the writing process and enabling writers to focus their attention successively on different aspects of the writing process (Fayol, 1999; Graham & Harris, 2012; Harris & Graham, 2009; Harris, Graham, & Mason, 2006; Kellogg, 1987; Limpo & Alves, 2013).

In previous work, we have reviewed the literature on cognitive writing processes in depth and have identified several dimensions that define the skills needed to develop writing expertise (Deane, 2011, 2013; Deane et al., 2008). In particular, we would distinguish between writing processes that focus on *planning*, *execution*, and *monitoring* (Harris et al., 2006; Hayes & Nash, 1996; Kellogg, 1996b, 2001; Limpo & Alves, 2014). However, we would also include *evaluation* as a separate process type, because revisiting a text to evaluate it and, if necessary, revise or edit it is quite distinct from online monitoring during text production. We also distinguish between five types of cognitive representation that writers coordinate during the course of the writing process: *social* representations that capture writers' understandings of audience, purpose, and task demands; *conceptual* representations that capture writers' knowledge of the topic about

which they are writing; discourse (or *text structure*) representations that capture writers' understanding about how information is organized and presented in text; *verbal* or linguistic representations that capture writers' abilities to express ideas in words, phrases, and sentences; and *orthographic* (though, also, phonological) representations that capture ideas and words in some physical expression that can be shared with others. These distinctions are implicit in most existing models of writing, though they are not always fully distinguished. For instance, social and conceptual representation corresponds to rhetorical and conceptual problem spaces as described in Bereiter and Scardamalia (1987) and to the proposer in the writing process model outlined in Hayes (2012), whereas the verbal and orthographic levels of representation correspond to the translator and transcriber. Table 1 shows how this conceptualization plays out, defining an array of specific categories that help characterize the practices of expert writers.

Table 1. Classification of Writing Processes by Activity Type and Mode of Cognitive Representation

Activity type	Plan	Execute	Monitor	Evaluate
Social	Create a rhetorical plan.	Control overall writing process to keep content, organization, style, and tone consistent with one's rhetorical purpose.	Track progress toward rhetorical goals and adjust subgoals as needed.	Evaluate overall writing success; rethink purpose, audience, and task, and revise accordingly.
Conceptual	Set goals for information gathering.	Recall or gather necessary information and organize what one knows about the subject.	Track how well content fits overall goals; as needed, generate additional content or delete existing material.	Evaluate and revise for accuracy, completeness of information presented, validity of arguments, etc.
Text-structural	Choose an organizational template appropriate to the task and genre.	Organize the text appropriately, providing appropriate cues to help the reader keep track of the intended structure.	Monitor and correct text; structure cues as needed.	Revise text globally to improve overall structure.
Verbal	Plan how to present ideas, sentence by sentence, clause by clause, and phrase by phrase.	Express one's ideas in words.	Monitor and correct expression as needed to make sure that the text accurately conveys intended message.	Evaluate and edit language to make sure that it is clear, unambiguous, and precise.
Orthographic	Access words; retrieve spellings and other orthographic details.	Produce each word, letter by letter, whether on a keyboard or by hand.	Monitor and correct typos, spelling errors, and other orthographic errors as needed.	Proofread an existing text and correct it as needed.

One of the challenges of teaching and assessing writing is that specific writing tasks vary in complexity and require different combinations of skills, depending on a variety of factors. For example, many forms of everyday writing, such as writing shopping lists or sending a text message, require very little planning or monitoring and depend primarily on the ability to translate ideas into words and write the words down. Alternatively, in professional contexts, writing can become a highly collaborative endeavor, with well-defined roles (i.e., first and second author; reviewer, editor, copy editor) that turn the process of publishing a text into a complex process in which the author must cooperate with an entire ecosystem of specialists (Beaufort, 2008). As students move into the upper grades and into college, the focus of academic writing instruction shifts from everyday writing to enabling students to function first in an academic setting and ultimately in a professional one, where they may be called on to participate in more complex forms of disciplinary writing (Englert, Mariage, & Dunsmore, 2006; Hyland, 2004; Shanahan & Shanahan, 2008).

The great variety of writing tasks constitutes one of the fundamental problems in developing or assessing students' writing skill. We are not concerned here only with the wide variety of specific topics, writing purposes, genres, and disciplinary settings in which writers may be called to write, even though this can have a major impact both on how well they write and on how accurately one can estimate writing skill from a limited number of assessment tasks. We are more concerned with the very different writing practices a writer may be called on to enact over the course of a single, complex writing project (Berninger, 2009). As writers work through a large-scale writing task, they may be called on to do something as simple as copying a quote from a text or as complex as critiquing the arguments in a source or completing a reorganization of their content to match the expectations of a specific discipline or genre.

Faced with such a challenge, skilled writers may switch strategically between writing processes, serially or in parallel, to produce a seamless, fluent writing product. Without the fluency, metacognitive knowledge, or working memory resources to match that performance, struggling writers are likely to execute much simpler writing strategies, which means that different writers may produce qualitatively different writing performances in response to the same writing prompt. To understand a writer's needs, it would be very helpful to be able to describe his or her writing process in detail; to identify when he or she is in planning, execution,

monitoring, or evaluation mode; and to determine when the writer is focusing more on rhetorical goals, conceptual content, text structure, language, or orthography.

Keystroke Log Analysis: A New Source of Evidence About Writing Processes

Given the value of understanding the writing process, there have been attempts to include it in assessments of writing; however, such attempts have historically been difficult to implement (Cho, 2003; DiStefano & Killion, 1984; Wiggins, 1994). But the migration of writing to digital platforms enables a more direct study of writing processes, because it is possible to record a writer's actions and describe his or her timing in detail through the use of a keystroke log (Baaijen, Galbraith, & de Glopper, 2012; Leijten & Van Waes, 2013; Spelman Miller & Sullivan, 2006; Wengelin, 2006). In fact, analysis of keystroke logs is turning into a major field of study in its own right.

Keystroke logs can be used to identify characteristic patterns associated with keyboarding (as opposed to handwriting) and with specific writing processes. In particular, there have been analyses of modality effects (Johansson, Johansson, & Wengelin, 2012; Van Waes & Schellens, 2003); analyses of the distribution of pauses during composition, especially in relation to their position in the text, such as between words, between clauses or sentences, and at other linguistic junctures (Fayol, 2012; Medimorec & Risko, 2017; Schilperoord & Sanders, 1999); analyses of pause patterns during revision and editing (Quinlan, Loncke, & Leijten, 2012; Roscoe, Snow, Allen, & McNamara, 2015; Stevenson, Schoonen, & de Glopper, 2006); and analyses of cognitive processes during translation from one language to another (Carl, 2012; Carl, Kay, & Jensen, 2010; Dimitrova, 2006; Heilman & Neumann, 2016; Jakobsen, 2006a, 2006b).

In addition, keystroke logs can be used to characterize performance, including individual and group differences. The resulting studies have included analyses of the relationship between fluency in text production and demands on working memory (Allen, Perret, & McNamara, 2016; Hayes & Chenoweth, 2006, 2012; Kellogg, Turner, Whiteford, & Mertens, 2016); analyses of individual and developmental differences in keystroke patterns and their relation to fluency and text quality (Allen, Jacovina, et al., 2016; Connelly, Dockrell, Walter, & Critten, 2012; Deane, 2014; Torkildsen, Morken, Helland, & Helland, 2015; Zhang & Deane, 2015); analyses of group differences characterizing struggling writers and writers with disabilities (Prunty, Barnett, Wilmut, & Plumb, 2014; Sumner, Connelly, & Barnett, 2013); analyses of group differences among more and less skilled second-language learners or between first- and second-language

writers (Mikulski & Elola, 2011; Revesz, Kourtali, & Mazgutovab, 2017; Spelman Miller & Lindgren, 2011; Sullivan, Kollberg, & Palsson, 1997; Thorson, 2000); analyses of the connection between keystroke patterns and other individual traits, such as engagement, affect, and personality (Allen, Mills et al., 2016; Bixler & D'Mello, 2013); and analyses of the potential application of keystroke logging to support instruction (Flinn, 1987; Lindgren, Stevenson, & Sullivan, 2008; Sabbaghan, 2013; Spelman Miller & Lindgren, 2011).

Finally, some scholars have examined ways to improve the analysis of keystroke logs or to combine keystroke log analysis with other methodologies to capture a richer description of the writing process. Major lines of work have included attempts to provide empirical evidence about the pause duration threshold that distinguishes execution from planning or evaluation behaviors (Chukharev-Hudilainen, 2014; Rosenqvist, 2015; Spelman Miller, 2002), exploration of more advanced (particularly statistical) methods for analyzing keystroke logs (Caporossi & Leblay, 2011; Chenu, Pellegrino, Jisa, & Fayol, 2014; Leblay & Caporossi, 2015; Perrin & Wildi, 2008; Van Waes & Leijten, 2015; Wallot & Grabowski, 2013), explorations of methodologies that combine keystroke logging with linguistic analysis (Macken, Hoste, Leijten, & Van Waes, 2012), and explorations of methodologies that combine keystroke logging with eye tracking (Andersson et al., 2006; Beers, Quinlan, & Harbaugh, 2010; Johansson, Wengelin, Johansson, & Holmqvist, 2010; Wengelin et al., 2009).

Implications of Task Contrasts for Keystroke Analysis

As our discussion indicates, keystroke log analysis has been applied to many different aspects of writing. But writing may vary considerably from one occasion to the next, depending on task demands. We would expect, for example, to see a very different profile for a purely typing task, where students have access to the content and need only type it out, than for a composition task, where students must express their own ideas. We would expect to observe a third distinct profile for an editing/proofreading task, where students have already produced a draft but must scan it to find and correct errors. There are a wide array of distinct writing tasks, each of which requires writing processes to be coordinated in slightly different ways. Typing, drafting, and editing are extreme cases, but even within the same broad category, we can expect different strategies for coordinating writing processes to come into play. Consider the different processes needed when someone drafts an essay as compared to the processes the same person might use when writing a letter or drafting a poem.

The variability of writing tasks potentially poses a problem for keystroke log analysis. Consider, for instance, a common measure of writing fluency, *burst length*, which is the number of words or characters a writer produces quickly without a relatively long pause. In the context of a straightforward drafting task, burst length has a relatively straightforward interpretation: It is the amount of text that a person can generate and then type in a single cycle from translation (putting ideas into words) to transcription (expressing those words as a sequence of characters). But if the writer is actually copying text, which could easily happen in the middle of a drafting session, or interrupting a drafting session to edit part of the existing text, some proportion of the bursts identified in that writing session may have an entirely different interpretation. The variability of keystroke features is potentially an issue, because there is already evidence that keystroke features can vary across quite similar tasks. In particular, Deane and Zhang (2015) examined evidence from pairs of essays written to similar prompts but to different topics and obtained a wide range of cross-task correlations. In their study, some features seemed relatively stable across tasks. For instance, burst length registered cross-task correlations between .48 and .85 with a median of .69, and duration of word-internal pauses registered cross-task correlations between .39 and .74 with a median of .61. Other features were much less stable. For instance, the duration of between-sentence pauses registered cross-task correlations between .09 and .22 with a median of .21. Thus it is important to understand exactly how features behave across tasks and to begin to understand how those differences are linked to differences in the writing process that each requires.

As Grabowski (2008) has argued, retyping is primarily a transcription task and hence, in a digital environment, a keyboarding task. Students do not need to plan content or to translate that content into words. There is a role for monitoring, because students need to read the source article, hold parts of it in working memory, and check whether their output matches the source. However, this role is fairly limited—mostly, students need to monitor the accuracy of keyboarding and make whatever small corrections are needed to deal with keyboarding errors. Classic accounts of copy typing (Shaffer, 1978; Shaffer & Hardwick, 1968) posit that copy typing is driven by an outer control loop, which loads the next word to be typed into working memory after it has been visually scanned, and an inner loop, which drives the motor processes that implement each keystroke. Correction actions happen very quickly, typically during keystroke implementation or immediately after a word has been typed, reflecting a monitoring

process (Shaffer, 1975). We would therefore expect the retyping task to have a preponderance of text-production actions (insertions of characters between, within, and at the start of words and sentences) and relatively few editing actions. Those editing actions that take place ought to occur locally as immediate corrections when a typing error is detected and should happen fairly quickly. Similarly, cognitive accounts of typing suggest that copy typing will alternate between phases in which the typist is reading the source text and phases in which the typist is keyboarding individual words. Pauses within a word are likely primarily to reflect pure keyboarding (the inner loop) and so should be relatively short. Pauses just before a word is typed are likely to reflect the outer loop, when the typist is reading the source text in order to load it into working memory. Interword pauses are likely to be relatively long, compared to the pauses that happen word internally, but they should only last long enough for the typist to find the right place in the text, if needed, and read the next words to be typed.

Conversely, drafting is primarily a verbal expression task, though it entails a heavy role for keyboarding. It differs from retyping in that it also requires text generation and thus implicates translation as well as transcription. In an extended writing project, drafting happens after prewriting and planning, which should minimize the role of the proposer, and before a separate revision and editing stage, which should limit the kinds of monitoring the writer needs to accomplish during drafting. In other words, drafting is likely to alternate between a translation phase (when the writer is generating words and phrases to express the next idea in the text he or she intends to write) and a transcription phase (when the writer actually types out the planned text).

We can therefore expect that typing patterns within individual words will not differ very much between copy typing and drafting. The same motor programs are being engaged after words have been loaded into working memory. Most of the differences will happen between words and at other major linguistic and conceptual junctures. Much longer pauses can be expected to where the drafter has to plan the next clause or sentence, and possibly even longer pauses may appear at junctures between longer text units, where ideas may need to be retrieved from an existing plan or generated if drafting and planning are interleaved. This is precisely what psycholinguistic accounts of pause patterns during sentence processing posit (Medimorec & Risko, 2017; Olive, Alves, & Castro, 2009; Wengelin, 2006). During drafting, monitoring processes should focus on detecting when the text produced does not match the intended

meaning and making changes to the current clause or sentence to improve the expression of ideas.

When writers edit and proofread, their task requires them to read and evaluate the current draft, identify changes that need to be made, and implement them. Achieving these goals necessarily minimizes the need for idea generation, translation, and transcription and emphasizes the role of the monitor, which will drive switches in attention and adoption of different goals as the writer moves through the text (Quinlan et al., 2012). We would expect writers to scan the entire text, making changes throughout, rather than working sequentially from beginning to end, as they are likely to do when retyping or drafting. In other words, during editing and proofreading—but not during copy typing or drafting—we expect to observe a sequence of carefully considered changes distributed throughout the text with relatively little insertion of new material.

Previous Task-Contrast Studies

The fundamental issue we have just identified—how to interpret a keystroke log in relation to the task that the writer seeks to perform—is critical if we want to support inferences from keystroke log features. If the same feature means something very different from one writing context to the next, or if the distribution of a feature changes fundamentally between tasks, we need to understand the parameters that drive such shifts. However, relatively few keystroke studies have administered contrasting writing tasks using a within-subject design. Two studies are worth highlighting: Grabowski (2008) and Wallot and Grabowski (2013).

Grabowski (2008) reported on a relatively small study involving 32 female education students at the University of Heidelberg. In this study, the subjects completed three distinct writing tasks. In the first task, they typed a text from memory (12 repetitions of the first sentence of a well-known nursery rhyme). In the second task, they did copy typing (they retyped a straightforward historical text describing a vote in the United States about choosing an official language). In the third task, they wrote from memory by composing short written text (they had to recall personal knowledge about the route from their residence to the university campus). They completed all three tasks in a single 30-minute session. The first two tasks are essentially transcription tasks, distinguished by whether the text to be retyped is presented orally or visually. The third task is essentially a translation and transcription task with a minimal role for idea

generation, because the writer should be able to generate the content by recalling information from working memory.

Grabowski's (2008) keystroke log analysis focused on a range of features, including counts of major events (number of keystrokes, deletions, and cursor movements), transition times within and between words, proportion of time spent in long pauses, and keyboarding efficiency (the proportion obtained by dividing the total number of characters in the final text by the total number of keystrokes). She linked these to various quality measures, including the number of typographic errors, length of the output in characters (for the writing task), and the completeness with which texts were copied. Her major finding was that there was a strong intertask correlation for the word-internal transition time feature (the mean length of pauses inside a word). For all three tasks, the correlation for this feature was in excess of .90, though she also observed small to moderate but significant correlations for several other features. She also observed significant differences between means for many of the features ($p < .001$ on a paired-sample *t*-test), most notably for the two transition time features, relative pause time, and keyboard efficiency. She also conducted principal component analyses for each of the three tasks and examined correlations between factors. The only factor that correlated across tasks was a typing speed factor to which the word-internal transition time feature made the strongest contribution.

Wallot and Grabowski (2013) replicated Grabowski's (2008) study with a slightly larger pool of 64 trainee teachers from the University of Heidelberg. However, they applied more complex forms of statistical analysis (fractal and recurrence analysis; cf. van Orden, Kloos, & Wallot, 2011; Wallot, O'Brien, & Van Orden, 2012) to try to characterize the temporal characteristics of the keystroke log. In particular, they applied power spectral density analysis to obtain an index that measures the presence of recurrent temporal patterns in the data. They applied multifractal analysis to obtain indices of time series complexity and stability/instability. These measures supported the conclusion that typing from memory was the simplest of the three tasks and writing from memory the most complex, though in this study, interkey intervals were significantly slower for copy typing, though not significantly different between the writing task and retyping from memory.

These studies support the conclusion that there can be variability in the properties of keystroke logs when task requirements vary. However, the studies were based on very small

samples and a relatively small set of features. It is not clear in either study how the features they extracted can be linked to substantive differences in the writing processes engaged to perform each task. And yet, it is critical to examine how tasks differ on features that capture meaningful differences between writing tasks, reflecting differences in the cognitive processes on which each task draws. That was our goal in this study.

Predicted Differences Between Copy Typing, Drafting, and Editing/Proofreading

The cognitive accounts of the writing process that we have just reviewed provide us with a principled basis for hypothesizing significant differences between copy typing, drafting, and editing. The three tasks activate different combinations of writing processes and therefore should produce different profiles if we extract features from the writing process log that provide evidence about each component process. The discussion that follows attempts to outline expected cognitive differences between the three tasks and to identify writing process features that provide evidence about specific cognitive processes activated during writing.

Action of the Proposer

The literature has indicated that idea generation can be a relatively time-consuming process and that it primarily occurs at the beginning of the writing process, though it also happens at text junctures where it makes sense for writers to generate content for the next piece of writing that they intend to produce. Because the primary evidence for action of the proposer is negative in nature (the absence of typing actions), it is difficult to identify features that necessarily reflect only the action of the proposer.

The following features are likely to provide evidence about the action of the proposer during the drafting process: (a) *start time*, or the proportion of total time spent pausing before starting to type, and (b) *long pauses*, or the proportion of time spent on long pauses (where long pauses are defined as pauses at least four standard deviations greater than the median interkey interval for each writer). Conversely, during editing/proofreading, these measures should primarily reflect the action of the evaluator and represent times during which writers were scanning the text to be edited to determine what errors might need to be corrected. And finally, during copy typing, these measures cannot reflect the action of the proposer (because no new text is being generated), and they are unlikely to reflect the evaluator (because the writer is supposed to reproduce the source text without needing to evaluate it), so pauses are most likely to occur

when writers glance at the source text to refresh their memory of the content they are reproducing. Any contrasts in the distribution of start time and long insertion pauses across the three task types are therefore likely to reflect differences in the intensity of the effort required for each nonwriting activity. Of the three, scanning text for copy typing is likely to require the least effort, because it requires neither idea generation nor evaluation. Scanning text to identify errors is likely to require the most effort, because it requires writers to hold the orthographic form of the text in tension with the intended meaning, to support evaluation, and to plan changes to either form or content. We therefore expect copy typing to require relatively little start time and time for long insertion pauses, drafting to require more, and editing/proofreading to require the most.

One potential issue for the interpretation of long pauses is that they may reflect disengagement with the task (due to lack of motivation, distraction, or frustration). In this data set, we do not have a good measure of student motivation and engagement, and it is therefore difficult to evaluate the extent to which motivation (or the lack of it) is affecting student performance on the retyping, drafting, and editing/proofreading tasks. However, we would expect disengagement to be reflected in other ways, such as students quitting a task early. In our data, most students expended a significant amount of time on each task (nearly an hour each), consistent with the in-class implementation, where the writing task was integrated with instruction and thus may have been taken more seriously by students as a result.

Action of the Translator

The literature has indicated that the action of the translator produces bursts of text production, with the pauses between bursts usually happening at word boundaries. We therefore selected the following features as likely to provide evidence about the action of the translator: (a) *burst length*, or the number of words in a burst of text production; (b) *word-initial pauses*, or the duration of pauses before typing the first character in a newly inserted word, and also the percentage of total time taken up by pauses before word-initial pauses (percentage of total time will be more sensitive to extreme events, reflecting how often long pauses happen specifically with this particular type of keystroke event and the others listed later); and (c) *pauses on white space and punctuation marks between words and sentences*—both the duration of pauses between words and sentences and the percentage of total time taken up by these events. During drafting, the length of bursts should be based on the size of the content units that the translator is able to generate within the writer's current working memory constraints and should therefore be

relatively long and preferentially associated with linguistic boundaries, such as word and sentence boundaries.

By contrast, the literature has suggested that copy typing is driven by the outer loop, one word or short phrase at a time. Similarly, editing/proofreading should be driven by a search for errors, followed by edits intended primarily to fix small chunks of existing text. These considerations imply that bursts that occur during drafting should be significantly longer than bursts during copy typing and editing/proofreading.

Similarly, during drafting, longer word-initial pauses should primarily occur when the translator is engaged in generating new text, though they may also be affected by lexical access processes or spelling difficulties. During drafting, the number of especially long word-initial pauses will depend primarily on the difficulty writers experience while generating content. By contrast, word-initial pauses should mean something very different during copy typing and editing/proofreading and primarily represent the need to visually inspect (and/or evaluate) the text to be copied or corrected. We expect such pauses for visual inspection to be relatively long, because they require visual transfer of attention followed by lexical access, whereas the primary time constraint that applies during translation is lexical access. We therefore predict that word-initial pauses during drafting will be shorter than those that occur during copy typing or editing/proofreading.

However, we also expect word-initial pauses during drafting to be quite variable in duration. Word-initial pauses that reflect extensive idea generation may occupy a significant proportion of total time. Thus, although individual word-initial pauses may be quite short, we expect that the percentage of total time occupied by word-initial pauses during drafting will be relatively large compared to the amount of time spent typing the characters of individual words.

During drafting, pauses at major junctures, such as pauses at sentence boundaries, are also likely to reflect the action of the translator and/or proposer. By contrast, during copy typing and editing/proofreading, it is unclear whether these pauses will happen quickly (because they do not contain content but rather only white space and punctuation) or more slowly (because the writer is scanning ahead to build up a working memory representation of the next phrase to be typed). But it seems plausible that we may observe differences in the properties of between-sentence pauses that distinguish drafting from the other two tasks.

Action of the Transcriber

The literature has indicated that the actions of the transcriber occur when writers already know which word they intend to type and are successively typing out the characters that compose each word. We therefore selected the following feature as being the most likely to provide information about the transcriber: *word-internal pauses*, or the duration of any keystroke that involves insertion of an alphanumeric character after the first character in a word; also, the percent of total time taken by word-internal pauses.

During drafting and copy typing, transcription forms part of a hierarchical process driven by the translator. As long as a writer has achieved a minimal level of translation and transcription skill, translation and transcription should happen with a high degree of automaticity without much involvement by processes that require executive control. Conversely, copy typing and editing/proofreading require the writer to switch actively between rereading an existing text and producing text output with an emphasis on accuracy. This additional processing should impose an additional load on working memory and should therefore reduce the efficiency of transcription. Thus we predict that word-internal pauses during drafting will be shorter than word-internal pauses during copy typing and editing/proofreading.

Action of the Evaluator

To the extent that writers are actively evaluating the text they produce and making changes to existing text as a result, we expect to see backspacing, character deletions or replacements, jumps from one position in the text to make an edit in another location, and larger cuts and pastes. We therefore postulate the following major features as likely to provide us with information about these kinds of processes: (a) the number, median duration, and total time spent on *backspace* events; (b) the number, median duration, and total time spent on *local edit* events, where text is replaced; (c) the number, median duration, and total time spent on *jump-to-edit* events, where a change happens in a different place in the text than the previous keystroke action; (d) the median *jump distance* traversed by jump events; (e) the number of *multiword deletion* events (e.g., cuts, repeated backspacing); (f) the *proportion of words edited* in some way, where no spelling correction is involved; and (g) the *proportion of words spell-corrected* while they were being typed. However, we expect to find that the kinds of edits a writer makes should be very different in each of the three task types.

During copy typing, the writer's attention should be focused almost exclusively on duplicating existing content. This implies that most actions are not edits and that most of the time and effort put into copy typing will primarily involve text insertion events (word-initial, word-internal, or between-word keystroke events) appended to the end of the current text. Edits should mostly happen when something goes wrong with the copying process so that writers detect mismatches between what they originally typed and what was actually produced. Theories of copy typing imply that the copying process mostly happens word by word. As a result, edits should happen within a very narrow scope, mostly within a few words of the word currently being produced, and should mostly take the form of small spelling corrections and edits. We do not expect copy typing to involve long sequences of backspaces, multiword cuts and pastes, or long jumps to edit portions of the text far from the end. And because the corrections to be made are mostly orthographic in nature, we do not expect the pauses that happen before editing events to be especially long.

During drafting, the evaluator should be focused primarily on evaluating the current content unit being produced, whether phrase, clause, or sentence. The act of monitoring and evaluating the drafted text may also sometimes lead to writers changing their minds about their content or phrasing, which may cause them to backspace over or delete multiple-word sequences before replacing them with new or modified content. We therefore expect a larger percentage of keystroke events to consist of backspacing and editing events as compared to copy typing and for jump events to be longer, though still usually within the same sentence. Because monitoring drafting involves evaluation of content and phrasing, not just spelling and other orthographic features, we also expect the pauses before editing events to be longer in drafting than in copy typing, to take up a larger proportion of total time, and to involve relatively fewer edits focused on spelling correction.

During editing/proofreading, the writer's attention should be focused (as in copy typing) primarily on the surface form of the text, because evaluation that leads to major changes in content is more properly characterized as revision. The critical difference is that during editing/proofreading, writers are evaluating the whole text rather than monitoring quality as they go along. So we expect much more of a focus on scanning the whole text and jumping to whatever point in the text where an issue is identified to make corrections. Editing/proofreading is a highly deliberative activity requiring writers to hold the actual text firmly in mind while they

compare it to an imagined text, the text that ought to be present, if it were to comply fully with the conventions of written English. We therefore expect much longer pauses between edit actions in editing/proofreading than in drafting or copy typing and therefore time spent pausing before an edit to be a larger proportion of total time, especially time for edits that involve jumps to a different location in the text.

Research Questions

In this study, we began by identifying a set of features that seemed likely to provide evidence about each of the underlying writing processes, that is, what Hayes (2012) called the proposer, translator, transcriber, and evaluator. We sought to answer the following research questions:

1. Are the task differences that we observe consistent with our theoretical predictions? That is, do we see stronger evidence of transcription processes in keyboarding, stronger evidence of idea generation and translation processes in drafting, and stronger evidence of whole-text evaluation processes in editing/proofreading?
2. Where two tasks share a common process component (such as transcription skill), do these features provide reliable measurement as evidenced by strong correlations across tasks?

Because transcription skills play a critical role in facilitating higher order writing, the second research question focuses on a critical baseline issue. If we can measure any skill reliably across tasks using writing process logs, it should be transcription. Moreover, because differences in writing patterns among individuals may be mediated in part by differences in transcription fluency (Berninger, 1999), it is important to establish how stably features measuring transcription behaviors perform when task demands are varied.

Method

Participants

We collected data from an urban middle school in a state in the American West. This school has a population of students who identify primarily as being from minority groups and from households with low socioeconomic status; 72.1% free and reduced-price lunch; 20% limited English proficiency; 60.3% Hispanic, 27.7% Black, 4.8% White, 2.6% Asian). We

recruited four eighth-grade English language arts (ELA) teachers with a total of 463 students in their classes, corresponding to the entire eighth-grade cohort.

Materials

This study represented our first opportunity to collect writing process data using a research version of the *CRITERION*® online writing evaluation service, an online formative writing tool developed by Educational Testing Service (ETS) that allows students to submit multiple drafts of an essay and to receive automated feedback, view teacher comments, and conduct peer review. CRITERION was used as the environment within which students planned, drafted, reviewed, and revised essays for their ELA classes. It was also made available for teachers to use as part of their normal writing instruction for the rest of the school year. The specific version of CRITERION that we deployed captured keystroke logs for the essay entry window, which was encoded as HTML-formatted text. Adapting the keystroke logging system to include HTML formatting was a significant development in its own right. That development effort overlapped with the time frame during which we collected data using CRITERION, because we also intended to use student data to validate the keystroke logging process.

The CRITERION online writing tool was combined with a formative task set developed as part of the *CBAL*® learning and assessment tool research initiative at ETS. The particular module we employed is focused on the topic "Should schools be allowed to sell junk food to their students?" This CBAL Junk Food task set is designed to support a 2- to 3-week classroom unit focused on teaching argument analysis and argument writing. A professional development session was held with the teacher team to determine the appropriateness of all tasks for their classes. Because the target school had limited availability of laptop computers and a limited photocopying budget, the research team produced paper versions of key tasks in this module that led up to the essay task and made copies of the materials for all participants. The paper packets prepared for this study are included in Appendix A (the student packet) and Appendix B (the teacher packet). The materials teachers were provided included a series of exercises focused on analyzing argumentation; articles to read and summarize; and support for the writing process, including planning and peer review. We collected all completed materials at the end of the implementation, but because the primary purpose of these CBAL materials was formative (to support teachers as they prepared students for the essay writing task), ETS did not score student responses.

In addition, students were given paper copies of the passage "Mark Twain's Huckleberry Finn" from the America's Story Web site produced by the Library of Congress[1] and were asked to retype it into the CRITERION essay window as part of their familiarization with the CRITERION tool.

Study Implementation

Although the legal agreement between ETS and the participating school precluded collection of individual demographic data or of individual grades assigned to student essays, the following activities were conducted.

In September 2015, participating teachers attended a collaborative professional development session that introduced teachers to CRITERION as a writing tool and to the CBAL Junk Food module, which gives students the opportunity to research an issue and then write an argument essay taking a stand on that issue. Teachers were provided with a teacher guide, a student guide, rubrics, and benchmark essays and were walked through the steps needed to implement an extended writing task in which students spent 10 class periods learning about the issue, writing an initial draft, reviewing one another's work, revising their work, and producing a final written essay. Teachers were invited to provide comments on and/or change the sequence or substance of the student packet during this session, and student packets were not printed until after the session.

As part of their introduction to the CRITERION platform and the Junk Food unit, students were asked to retype the article "Mark Twain's Huckleberry Finn" as a CRITERION submission. They were allowed about half an hour to complete this task, which served two purposes: to familiarize students with the interface and to collect baseline data about students' typing skills.

In October 2015, students spent a week working through a sequence of preparatory tasks, using paper versions of the CBAL Junk Food formative module. In the second week of the implementation, they completed a series of writing assignments in the following sequence: (a) create a plan for their essay; (b) peer review one another's plans; (c) create an initial draft; (d) peer review one another's drafts; and (e) revise, edit, proofread, and finalize their essays. Keystroke logs were collected for the retyping task and for initial and final essay drafts.

Between October 2015 and June 2016, teachers continued to use CRITERION with their students, primarily to support embedded writing assessments already planned as part of their

curriculum. However, this study focused on student writing behavior during the initial Junk Food module.

Classification of writing process logs. Although students were supposed to perform specific tasks on specific days of the implementation, student behavior was not always consistent with expectations. For instance, some students only partially drafted their essays by the due date for a first draft and used the remaining sessions to extend that draft without actually doing any revisions or edits. Given these complexities, we manually classified student essay logs by examining a replay of their behavior during the writing process. Student logs were assigned to the following categories:

- *article retype (copy typing)*: Logs were created during the initial session, in which students tested out CRITERION by retyping an article.
- *article continuation*: A student logged back in to correct errors in the initial retyping of the article or to add to it. This was not an expected behavior, and these logs were excluded from further analysis.
- *plan record*: A student followed directions and pasted the text of his or her outline into CRITERION. We did not analyze the keystroke logs from these sessions, as they mostly consisted of a single paste event in which the plan text was inserted into the CRITERION essay submission box.
- *draft*. Students created or modified their text by adding significant amounts of new content.
- *edit*: Students modified the existing text in ways that suggested they were focusing on proofreading, error correction, and language, but not content.

After excluding plan records and article continuations, we obtained a three-way comparison among article retyping, essay drafting, and essay editing/proofreading. However, there were missing data, reflecting two different issues: student behavior (almost all students completed retyping and drafting activities, but not all students completed a full drafting or editing/proofreading session) and technical issues (some keystroke logs had to be discarded owing to errors in the data collection process).

Data cleaning. This study took place while we were still conducting quality assurance for the writing process log collection process and therefore also served as our primary vehicle for identifying and correcting errors in the writing process logs. We tested every log for every session to confirm that the log correctly recorded student actions. If the final result of the keystroke log did not match the submitted text, we worked to identify and fix the observed errors. Because teachers in the study used CRITERION throughout the year, we were able to conduct several rounds of testing, during which we fixed errors observed in the previous round and then collected fresh student data. However, the data set we analyzed in this study was collected early in the year, when students retyped an article and wrote an essay on the Junk Food topic. We were able to detect and correct some of the errors in the logs, but where we were unable to do so, we excluded the logs from the study. After repairing errors in the logs, we obtained a final data set containing records for 371 students. That data set contains 341 article retyping logs, 294 drafting logs, and 168 editing/proofreading logs. We have pairs of retyping/drafting logs for 266 students, pairs of drafting/editing logs for 116 students, and pairs of retyping/editing logs for 152 students. All three logs are available for 102 students. Thus, given the missing data, the final data set is best suited to pairwise comparisons between particular tasks, contrasting retyping to drafting or drafting to editing/proofreading.

Planned Analyses

We conducted a series of comparisons across the features we identified as relevant for each of the four major cognitive processes in writing to answer our first research question. In particular, we planned the following analyses to answer the first research question: (a) a cross-tabulation of event frequencies, calculating the chi-square statistic and standardized residuals for major keystroke event types across the three tasks; (b) an examination of feature distribution histograms to identify salient cross-task contrasts; and (c) the administration of Wilcoxon signed-rank tests (rather than analyses of variance, because the features we compared were not normally distributed) to determine whether there were differences in median feature values across tasks. Because there was a total of 17 comparisons between any pair of tasks across this set of analyses, we applied a Bonferroni correction of .05/17 = .0029 to set an appropriate significance threshold for each component analysis.

To answer the second research question, we examined the correlation between pause durations for word-internal pauses across the three task types and compared the strength of this

correlation with the strength of correlations for other event types, many of which differ significantly in interpretation across tasks. Transcription is the major process shared across tasks (particularly between copy typing and drafting), and the strength of this correlation and how much more reliable it is than less stably interpretable features should indicate how reliably the keystroke log measures transcription skill.

Results

Differences in Event Frequencies

A chi-squared test was conducted to assess the relationship between writing task (article retyping, essay drafting, essay revision, and essay editing/proofreading) and the frequency of different categories of keystroke action (see Table 2; we examine the pause times associated with each keystroke action type in subsequent sections).

Table 2. Cross-Tabulation, Number of Keystroke Actions by Type of Writing Task

Writing task	Keystroke action	Frequency	Expected frequency	Standardized residual
Article retype	in-word insertion	317,742	294,341	+43.1
	between-word insertion	94,.127	92,485	+5.4
	word-initial insertion	87,668	85,236	+8.3
	between-sentence insertion	12,834	10,582	+21.9
	between-paragraph insertion	76	218	−9.6
	backspace	51,296	76,830	−92.1
	jump to edit	5,592	8,811	−34.3
	edit	1,372	2,205	−17.7
Essay draft	in-word insertion	247,191	266,476	−37.4
	between-word insertion	81,108	83,729	−9.1
	word-initial insertion	75,631	77,167	−5.5
	between-sentence insertion	7,328	9,581	−23.0
	between-paragraph insertion	307	198	+7.8
	backspace	92,852	69,556	+88.3
	jump to edit	9,656	7,977	+18.8
	edit	2,607	1,996	+13.7
Essay edit/proof	in-word insertion	11,170	15,287	−33.3
	between-word insertion	5,782	4,803	+14.1
	word-initial insertion	3,530	4,427	−13.5
	between-sentence insertion	551	550	.1
	between-paragraph insertion	44	11	+9.7
	backspace	6,228	3,990	+35.4
	jump to edit	1,998	458	+72.0
	edit	337	115	+20.8

The test was found to be statistically significant, $\chi^2(14, N = 1{,}116{,}704) = 30{,}672.03$. The results indicate that there are significant differences in the relative frequency of keystroke actions between tasks (see Figure 1). In particular, the article retyping task appears to contain a greater

proportion of text-production events (e.g., insertion keystrokes appearing word initially, within words, between words, between sentences, and between paragraphs) and a smaller proportion of backspacing and editing actions compared to the other two tasks. Essay drafting sessions appeared to display somewhat more editing actions and a much larger proportion of backspacing events as compared to the other two tasks. Finally, essay editing/proofreading sessions were distinguished by a much larger proportion of jump-to-edit actions as compared to the other two tasks.

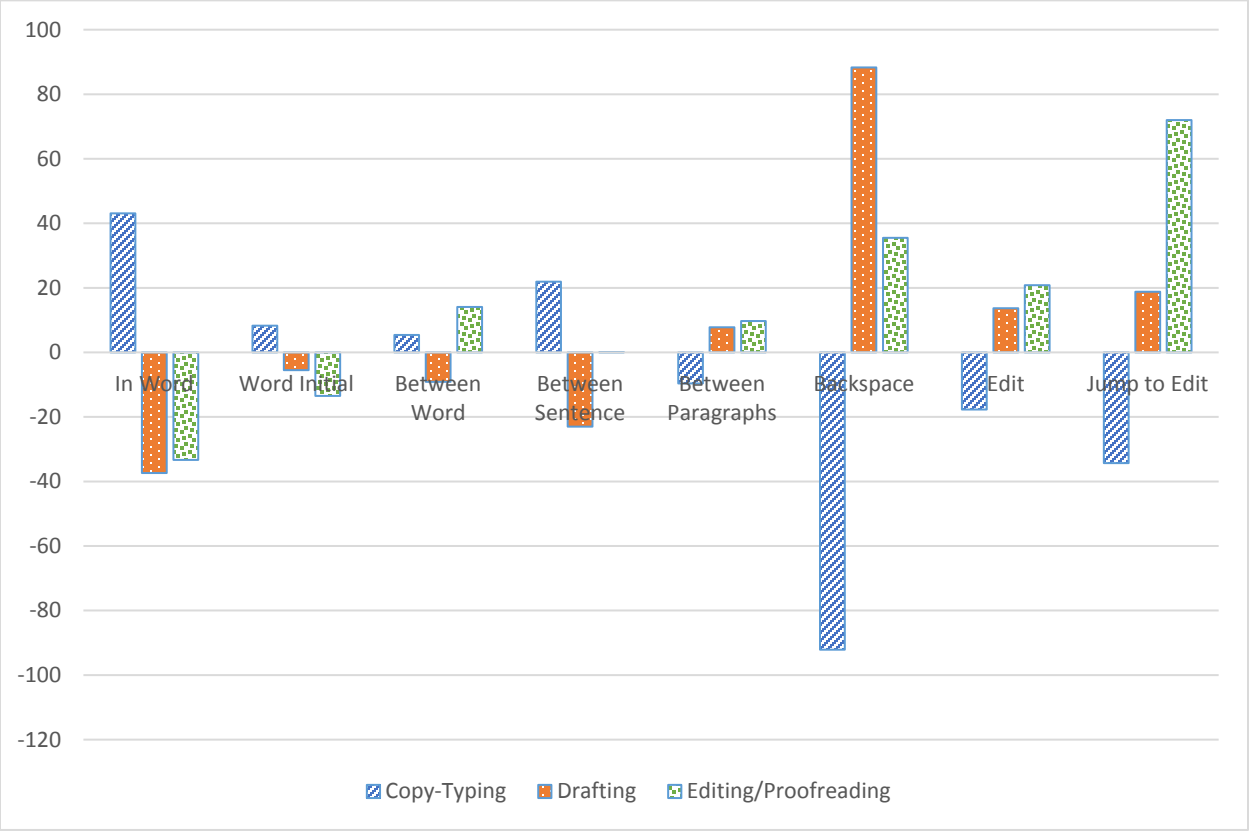

Figure 1. Standardized residuals for frequency of keystroke event types by type of writing task.

Differences in Event Durations

We examined event durations for the keystroke event types *word internal*, *word initial*, *between word*, *backspace*, *local edit*, and *jump to edit*, though not between paragraphs, as the total number of between-paragraph keystrokes was too small for further analysis. For each keystroke event type, we initially examined graphs of the distribution of pause durations for both the actual temporal values and logged values for each keystroke event type by task. Because our

goal was specifically to examine differences between tasks at the level of individual students, we used the logged median pause duration as a measure of central tendency for each individual and applied the Wilcoxon sign-rank text to determine whether, for each event type, the population median over all individual medians differed by task.

Figure 2 shows the differences we observed. Overall, the results indicate that keystroke durations are significantly shorter during drafting for pauses associated with continuous text production, for example, between-word, word-initial, and word-internal pauses. In addition, two other significant orderings were observed. First, median log backspace pauses are significantly longer for copy typing than for drafting or editing/proofreading. Second, median log word-initial pauses and median log jump-to-edit pauses are significantly longer for editing/proofreading than for drafting and significantly longer in turn for drafting than for copy typing. The evidence supporting these conclusions is presented in the following pages.

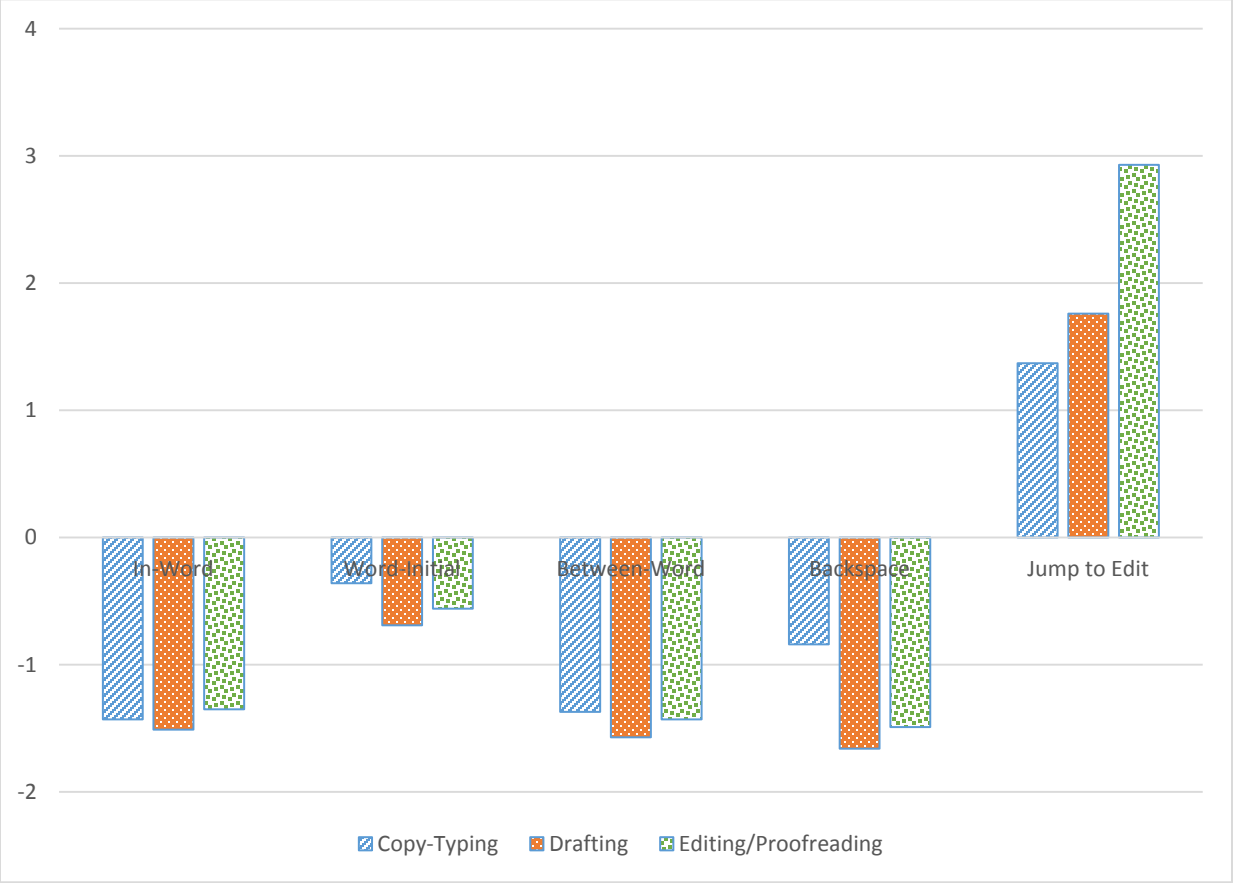

Figure 2. Median of median log pause durations for five event types (in log seconds).

Word-internal pauses. Figures 3–5 show differences in median log word-internal keystroke durations in copy typing, drafting, and editing/proofreading across all students. The three tasks show distributions of generally similar shapes, with a sharp peak representing a modal response involving very short pauses (copy typing median = .24 seconds or −1.43 log seconds, drafting median = .22 seconds or −1.51 log seconds, and editing/proofreading median = .26 s or −1.35 log seconds). An examination of random individual samples demonstrated similar patterns.

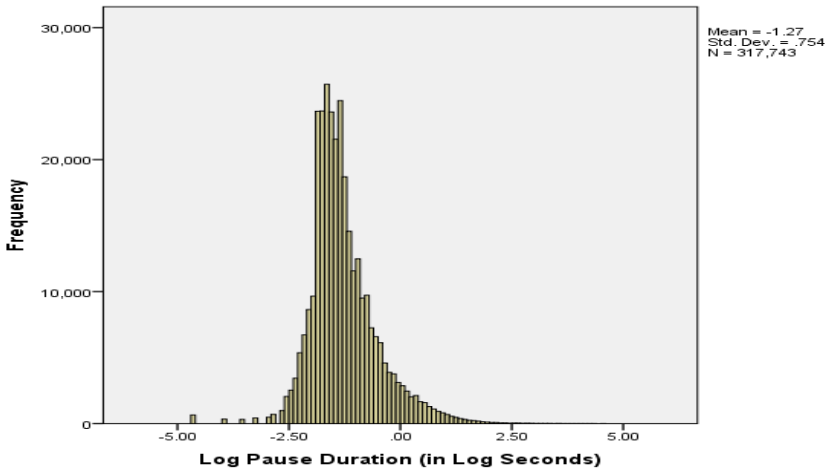

Figure 3. Word-internal pauses during copy typing, all students.

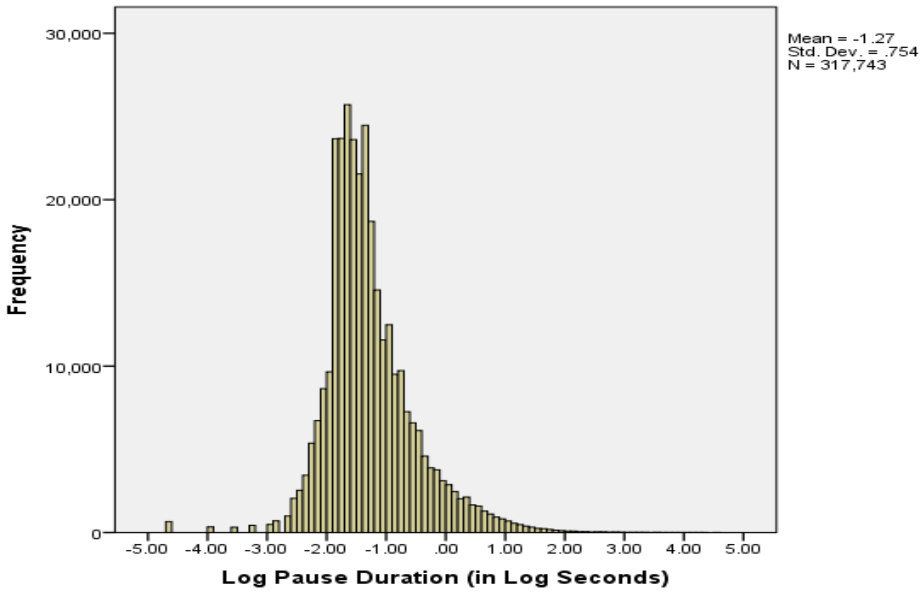

Figure 4. Word-internal pauses during drafting, all students.

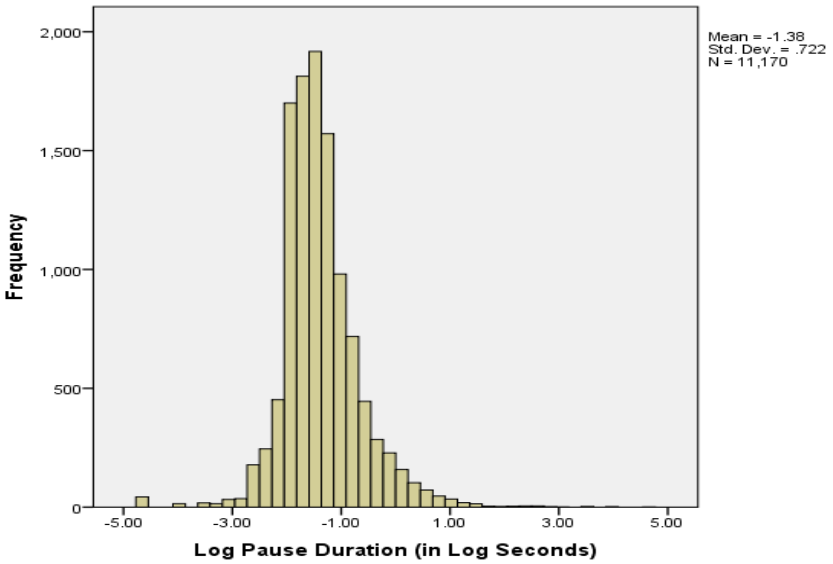

Figure 5. Word-internal pauses during editing/proofreading, all students.

A Wilcoxon signed-ranks test indicated that the grand median (i.e., median over all individual medians) was greater during copy typing than during drafting, $Z = -10.52$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the grand median of word-internal pause durations was smaller during drafting than during editing/proofreading, $Z = 3.61$, $p < .001$. A third Wilcoxon signed-ranks test indicated that there was not a significant difference in the grand median of log word-internal pauses during retyping and editing/proofreading, $p = .53$. The distribution of individual median log word-internal pause durations is shown in Table 3 by quartiles.

Table 3. Medians and First and Third Quartiles of Individual Median Word-Internal Pause Times for Copy Typing, Drafting, and Editing/Proofreading, Reported in Log Seconds

Period	Copy typing	Drafting	Editing/proofreading
Median	−1.43 (0.24)	−1.51 (0.22)	−1.35 (0.26)
1st quartile	−1.56 (0.21)	−1.66 (0.19)	−1.61 (0.20)
3rd quartile	−1.24 (0.29)	−1.35 (0.26)	−1.11 (0.33)

Word-initial pauses. Figures 6–8 show the distribution of word-initial keystrokes in copy typing, drafting, and editing/proofreading across all students. Once again, we observe a strong central peak, though there is a visible tail of longer pauses for drafting and editing/proofreading (median for copy typing = .70 seconds or −.36 log seconds, median for drafting = .50 seconds or −.59 log seconds, median for editing/proofreading = .57 seconds

or −.56 log seconds). Similar distributions are observed when individual student responses are randomly sampled.

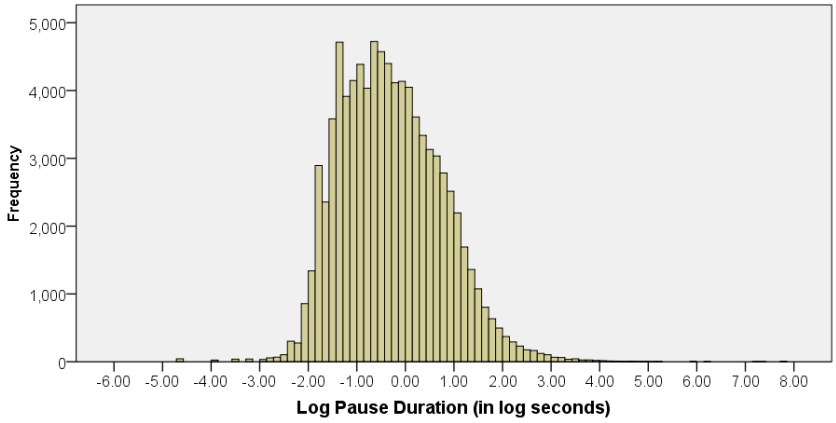

Figure 6. Word-initial pauses during copy typing.

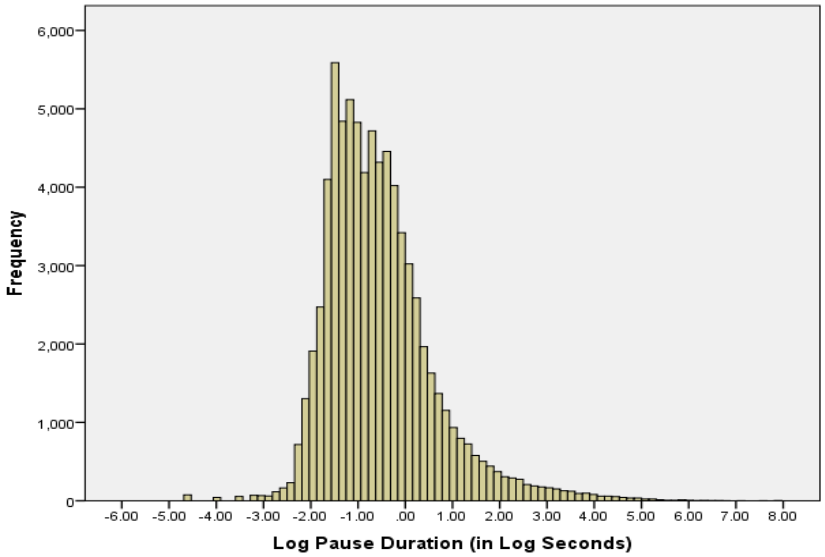

Figure 7. Word-initial pause durations during drafting.

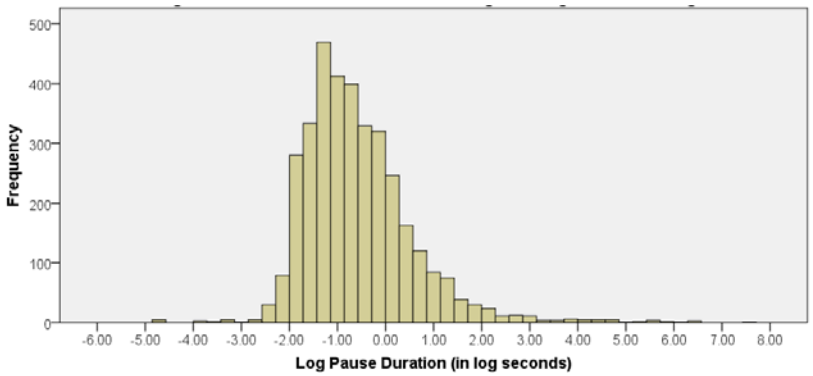

Figure 8. Word-initial pauses during editing/proofreading.

A Wilcoxon signed-ranks test indicated that the grand median (i.e., median over all individual medians) of log word-initial pause durations was greater during copy typing than during drafting, $Z = -12.11$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the grand median of log word-initial pause durations was significantly smaller during drafting than during editing/proofreading, $Z = 2.91$, $p = .004$. A third Wilcoxon signed-ranks test indicated that the grand median of log word-initial pause durations was significantly greater during copy typing than during editing/proofreading, $Z = -3.86$, $p < .001$. Table 4 shows the median and the first and third quartile values for each task.

Table 4. Medians and First and Third Quartiles of Word-Initial Pause Times for Copy Typing, Drafting, and Editing/Proofreading, Reported in Log Seconds

Period	Copy typing	Drafting	Editing/proofreading
Median	−.36 (0.70)	−.69 (0.50)	−.56 (0.57)
1st quartile	−.63 (0.53)	−.94 (0.39)	−.87 (0.42)
3rd quartile	−.01 (0.99)	−.43 (0.65)	−.09 (0.91)

Between-word pauses. Figures 9–11 show the distribution of between-word pauses in copy typing, drafting, and editing/proofreading across all students. The distribution of between-word pauses for copy typing in Figure 9 in particular suggests a mixture model where a sharp peak of long pauses much like that observed with drafting in Figure 10 and editing/proofreading in Figure 11 is mixed with another distribution of much longer pauses (median for copy typing = .26 seconds or −1.37 log seconds, median for drafting = .21 seconds or −1.57 log seconds, median for editing/proofreading = .24 seconds or −1.43 log seconds). Similar patterns were observed in random individual samples.

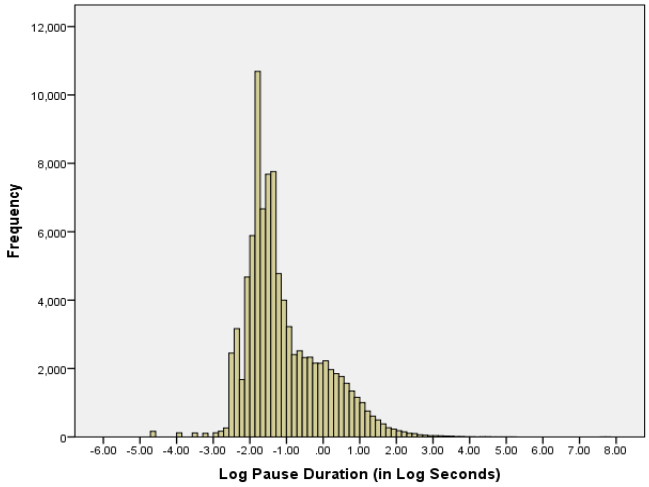

Figure 9. Between-word pauses during copy typing.

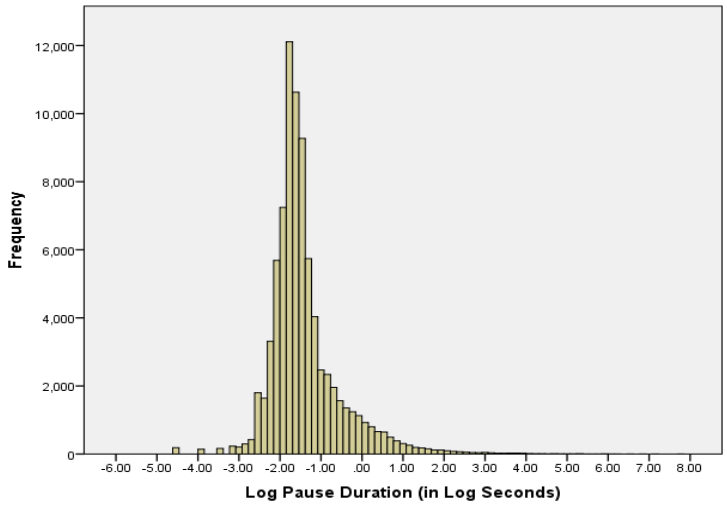

Figure 10. Between-word pauses during drafting.

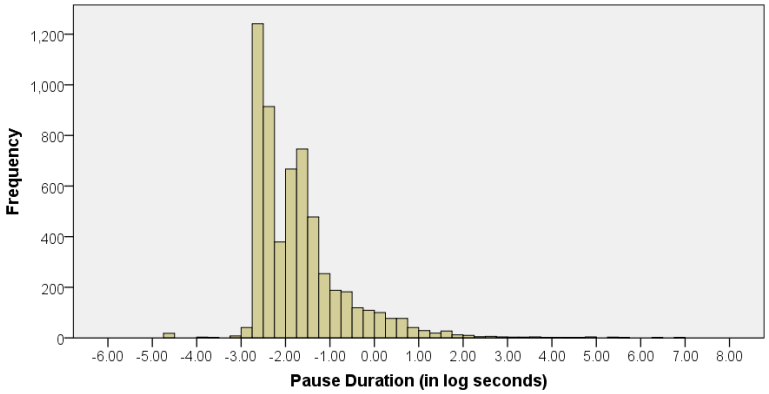

Figure 11. Between-word pauses during editing/proofreading.

A Wilcoxon signed-ranks test indicated that the grand median of log between-word pause durations was greater during copy typing than during drafting, $Z = -10.30$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the grand median of log between-word pause durations was significantly less during drafting than during editing/proofreading, $Z = 4.91$, $p = .004$. A third Wilcoxon signed-ranks test indicated that the grand medians of log between-word pause durations during editing/proofreading and copy typing were not significantly different, $p = .52$. (See Table 5).

Table 5. Median plus First, and Third Quartiles of Between-Word Pause Times for Copy Typing, Drafting, and Editing/Proofreading, Reported in Log Seconds

Period	Copy typing	Drafting	Editing/proofreading
Median	−1.37 (0.26)	−1.57 (0.21)	−1.43 (0.24)
1st quartile	−1.61 (0.20)	−1.71 (0.18)	−1.71 (0.18)
3rd quartile	−1.06 (0.35)	−1.43 (0.24)	−.67 (0.51)

Between-sentence pauses. Figures 12–14 show the distribution of between-sentence-initial keystrokes in copy typing, drafting, and editing/proofreading across all students. The distributions for copy typing and editing/proofreading suggest a mixture of two distributions: a small but significant group of very fast pauses and a broader distribution of longer pauses. The distribution for drafting has a single peak, though there may be a small second peak at about the pause length where we see the peak for short responses for between-sentence pauses for the other two tasks.

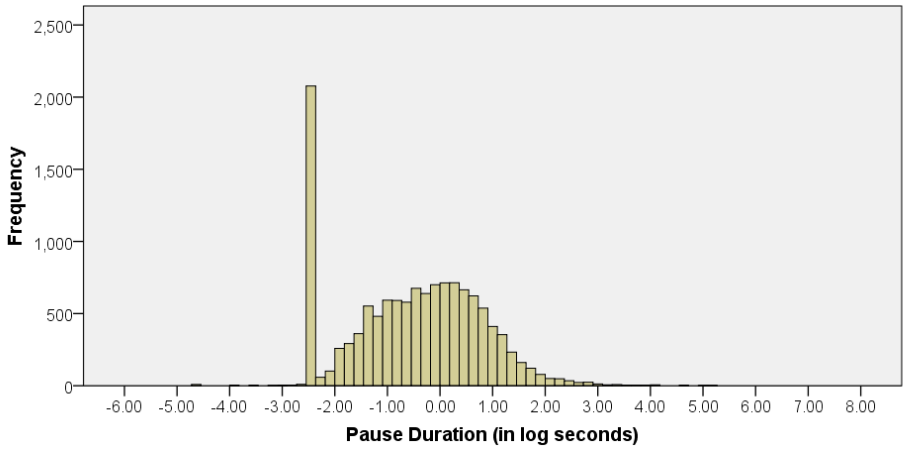

Figure 12. Between-sentence pauses during copy typing.

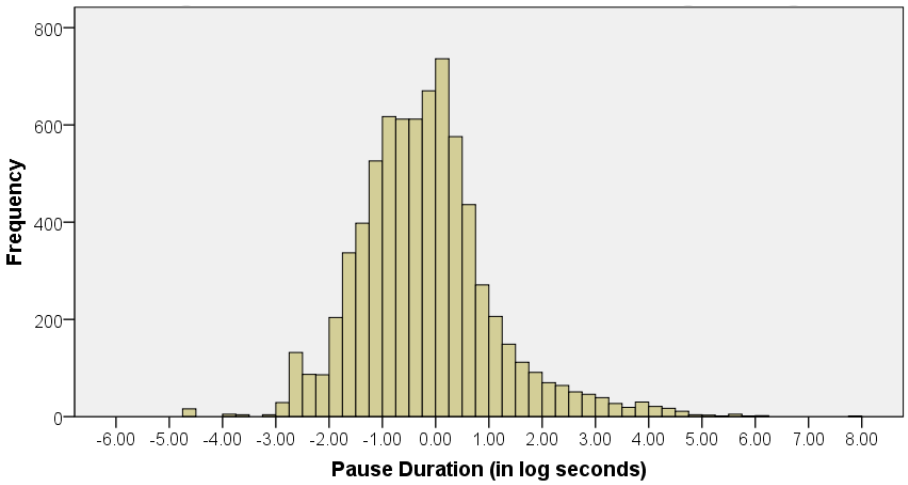

Figure 13. Between-sentence pauses during drafting.

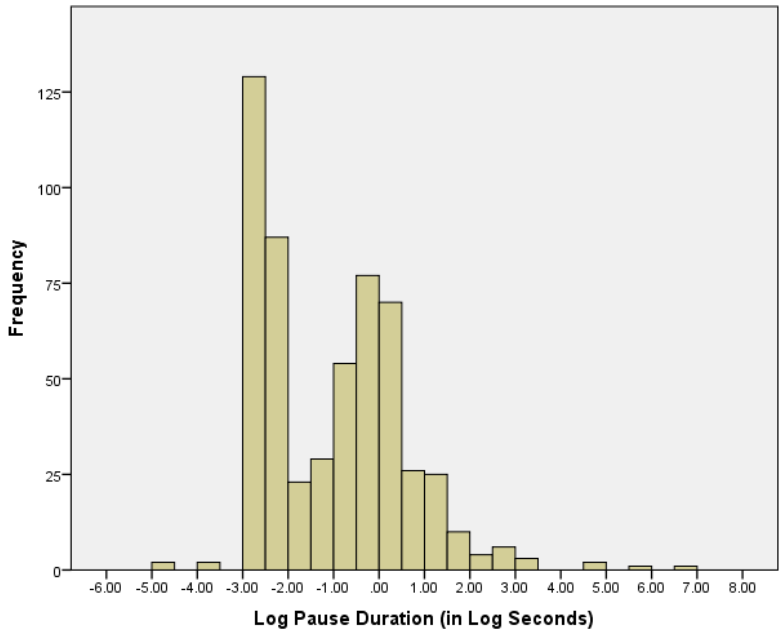

Figure 14. Between-sentence pauses in editing/proofreading.

However, we conducted three Wilcoxon signed-ranks tests examining the grand medians of log between-sentence pause durations during copy typing, drafting, and editing/proofreading. After Bonferroni corrections, none of these comparisons were significantly different (copy typing vs. drafting, $p = .02$; copy typing vs. editing/proofreading, $p = .90$; drafting vs. editing/proofreading, $p = .31$).

Backspace pauses. Figures 15–17 show the distribution of backspace keystrokes in copy typing, drafting, and editing/proofreading across all students. There are actually three peaks in the distribution of backspaces for all three tasks, suggesting some kind of mixture distribution (copy typing median = .43 seconds or −.84 log seconds; drafting median = .19 seconds or −1.66 log seconds; editing/proofreading median = .23 seconds or −1.49 log seconds). Examining random individual student samples also indicated more than one peak to the distribution, though many students only showed two peaks.

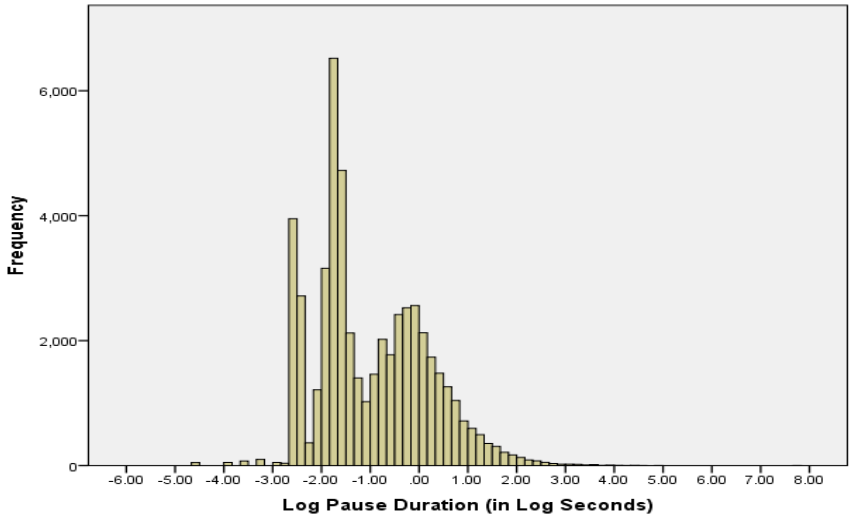

Figure 15. Backspace pauses during copy typing.

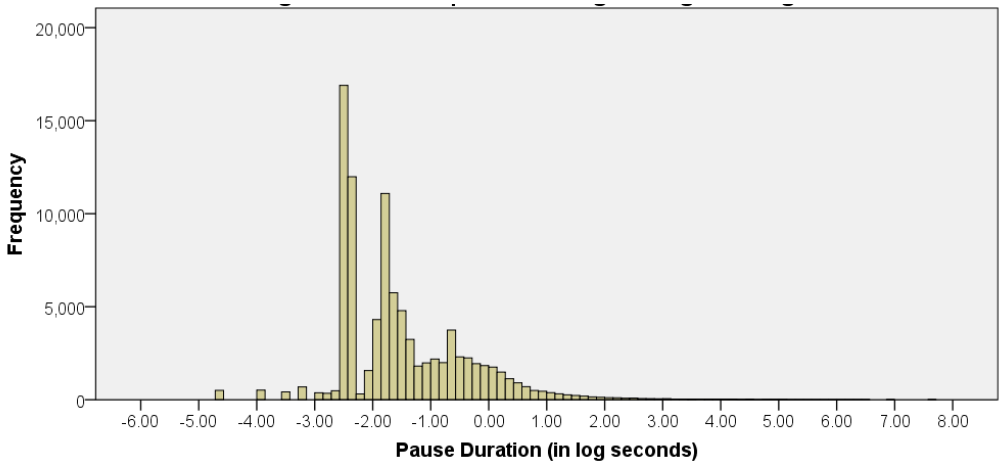

Figure 16. Backspace pausing during drafting.

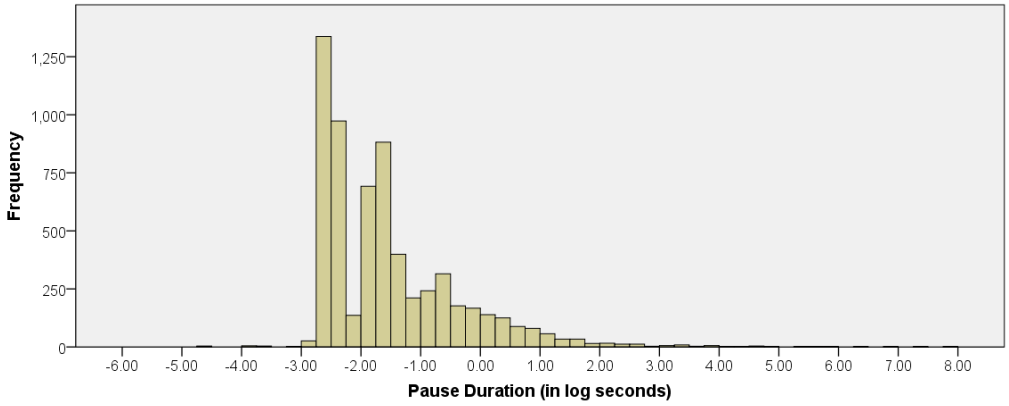

Figure 17. Backspace pausing during editing/proofreading.

A Wilcoxon signed-ranks test indicated that the grand median of log backspace pause durations was greater during copy typing than during drafting, $Z = -12.06$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the grand median of log backspace pause durations was smaller during drafting than during editing/proofreading, $Z = 4.71$, $p < .001$. After the Bonferroni correction was applied, a third Wilcoxon signed-ranks test indicated that the median values of log backspace pause durations during copy typing and editing/proofreading were not significantly different, though there was a marginal effect indicating that the grand median of backspace pauses during copy typing might be longer than during editing/proofreading, $p = .013$. Table 6 shows the median and first and third quartile values for backspace pauses in each of the three tasks.

Table 6. Medians and First and Third Quartiles of Backspace Pause Times for Copy Typing, Drafting, and Editing/Proofreading, Reported in Log Seconds

Period	Copy typing	Drafting	Editing/proofreading
Median	−.84 (0.43)	−1.66 (0.19)	−1.49 (0.23)
1st quartile	−1.40 (0.25)	−1.83 (0.16)	−1.77 (0.17)
3rd quartile	−.46 (0.63)	−1.43 (0.24)	−.58 (0.56)

Local edit pauses. Figures 18–20 show the distribution of local edit keystrokes in copy typing, drafting, and editing/proofreading across all students. The number of local edits for any one student is quite small.

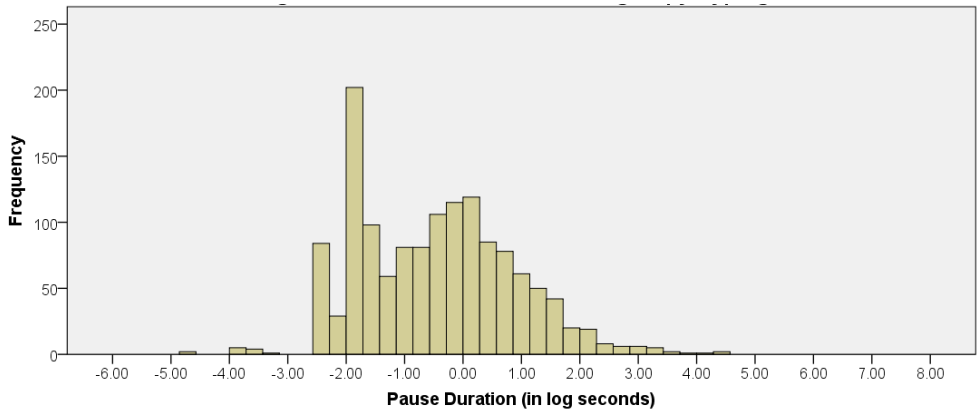

Figure 18. Local edit pauses during copy typing.

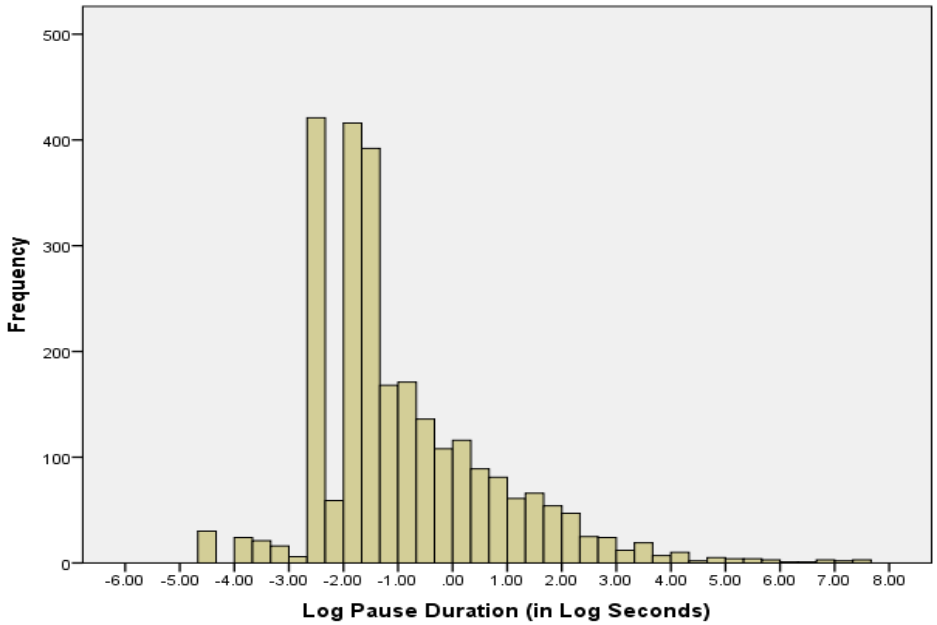

Figure 19. Local edit pauses during drafting.

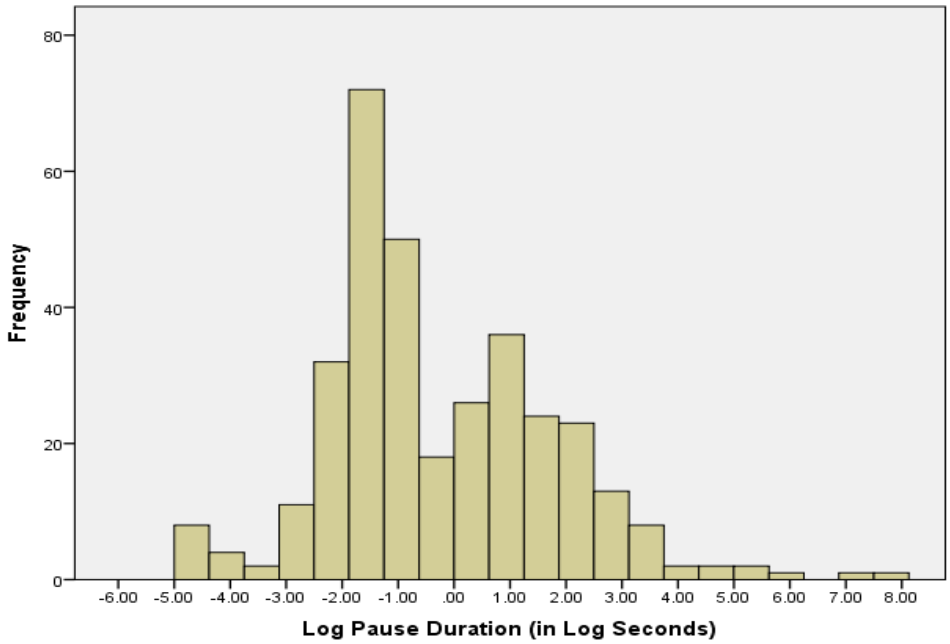

Figure 20. Local edit pauses during editing/proofreading.

We conducted three Wilcoxon signed-ranks tests examining the relation between the median values of log local edit durations during copy typing, drafting, and editing/proofreading. After Bonferroni corrections, none of these comparisons were significantly different (copy

typing vs. drafting, $p = .92$; copy typing vs. editing/proofreading, $p = .99$; drafting vs. editing/proofreading, $p = .32$).

Jump-to-edit pauses. Figures 21–23 show the distribution of jump-to-edit keystrokes in copy typing, drafting, and editing/proofreading across all students, which looks slightly bimodal, with a primary peak at a relatively long pause duration and a possible smaller second peak involving very short pauses (copy typing grand median = 3.94 seconds or 1.37 log seconds; drafting grand median = 11.66 seconds or 2.46 log seconds). The number of individual jump-to-edit pauses can be quite low.

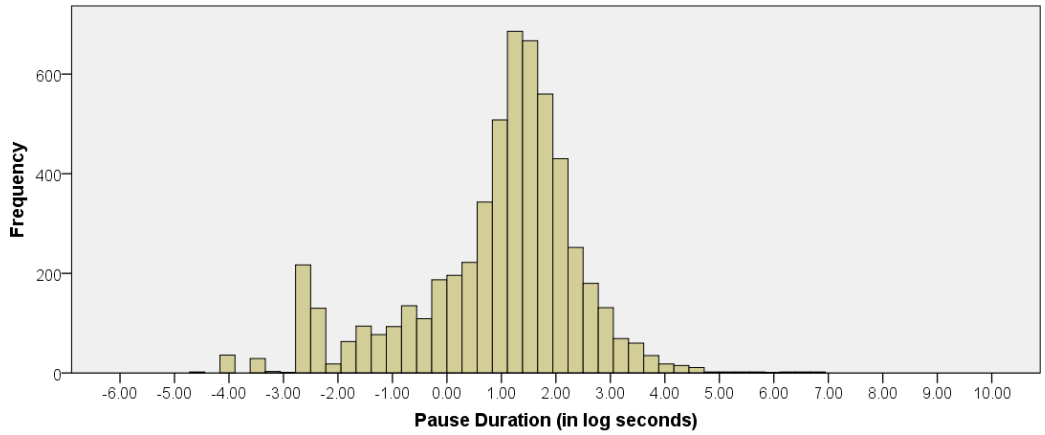

Figure 21. Jump-to-edit pauses during copy typing.

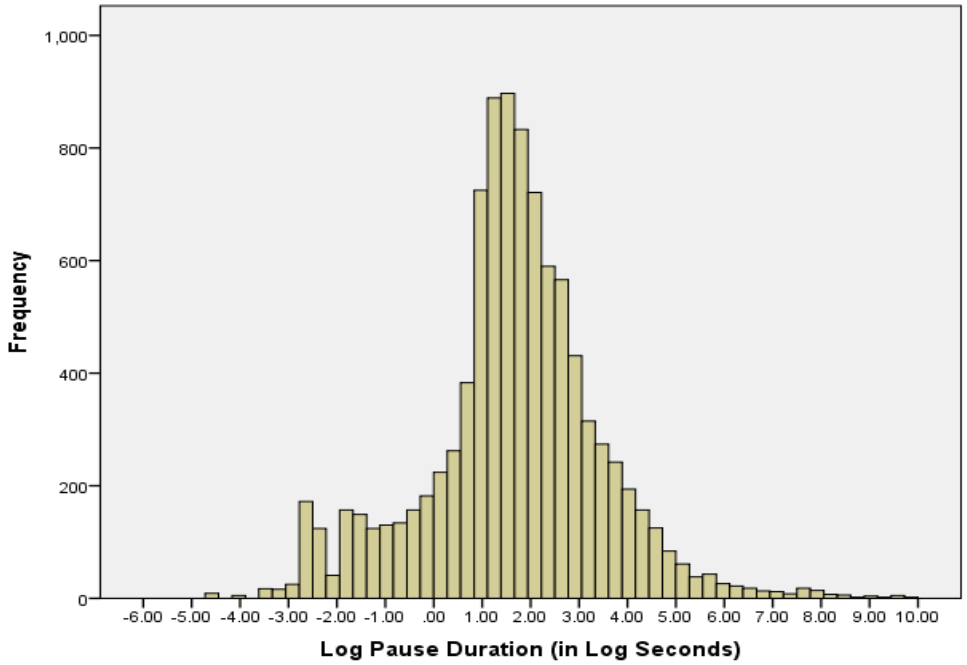

Figure 22. Jump-to-edit pauses during drafting.

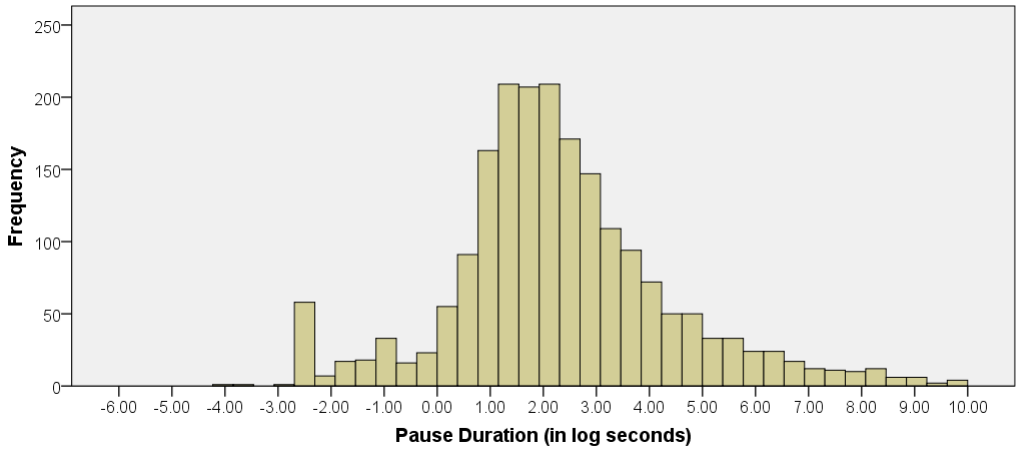

Figure 23. Jump-to-edit pauses during editing/proofreading.

A Wilcoxon signed-ranks test indicated that the grand median log jump-to-edit pause durations was smaller during copy typing than during drafting, $Z = 8.29$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the grand median of log jump-to-edit pause durations was smaller during drafting than during editing/proofreading, $Z = 5.99$ $p < .001$. A third Wilcoxon signed-ranks test indicated that the grand median of log jump-to-edit pause durations was smaller during copy typing than during editing/proofreading, $Z = 8.13$, $p < .001$ (see Table 7).

Table 7. Medians and First and Third Quartiles of Jump-to-Edit Pause Times for Copy Typing, Drafting, and Editing/Proofreading, Reported in Log Seconds

Period	Copy typing	Drafting	Editing/proofreading
Median	1.37 (3.94)	1.69 (5.41)	2.46 (11.66)
1st quartile	0.90 (2.46)	1.33 (3.79)	1.76 (5.80)
3rd quartile	1.59 (4.89)	2.13 (8.42)	2.93 (18.71)

Differences in Percentage of Total Time

We examined percentage of total time for the following event types: start time, long pause time, word-internal pause time, word-initial pause time, between-word pause time, backspace pause time, between-sentence pause time, local edit time, and jump-to-edit time.

To gain a clear sense of how time is distributed across tasks, we examined histograms for the distribution of percentage of total time for each event type, and then we applied the Wilcoxon sign-ranks test to determine whether the median value for each event's percentage of total time was different across tasks. Figure 24 shows the general pattern that we observed.

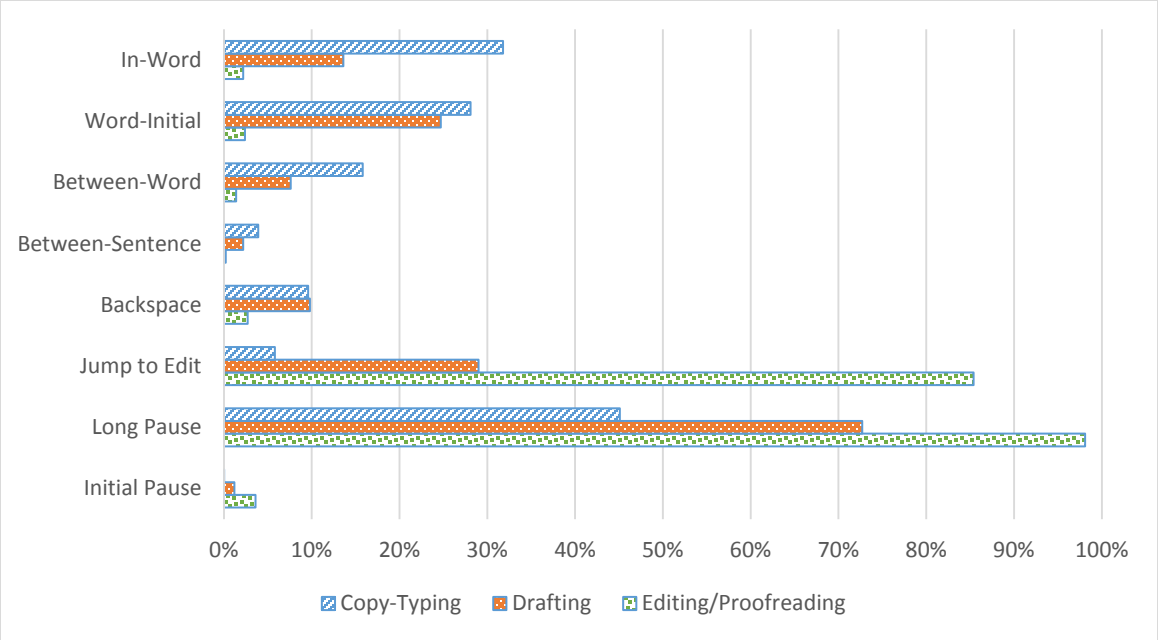

Figure 24. Percentage of total time by keystroke event category.

Note that start time and long pause time overlap with the other event categories, which are otherwise mutually exclusive. The general pattern we observe is that copy typing involves a larger proportion of time devoted to text-production events (word internal, word initial, and between word), significantly less time devoted to jump-to-edit events, and less time in long pauses or pausing initially than the other two task types. Editing/proofreading, in contrast, involves significantly less percentage of total time dedicated to text-production events, significantly more time on jump-to-edit events, significantly more time accounted for by long pauses, and significantly longer initial pauses. Drafting is distinguished by nearly twice as much time being devoted to word-initial pauses than to word-internal pauses and to increased time for jump-to-edit events, to long pauses generally, and to initial pauses, though these increases are not as large as for the editing/proofreading task. Details of the individual analyses are outlined in the following sections.

Percentage word-internal pause time. Figures 25–27 show the distributions of total word-internal pause time as a percentage of total time on task. The distributions differ in obvious ways. As Figure 25 illustrates, the percentage of word-internal pause times for copy typing peaks strongly around 32% of total time (median = 31.8%). By contrast, the distribution of word-internal pause time for drafting has a broad peak running from about 5% to 20% of total time (median = 13.6%), though it appears to be bimodal, with one peak around 8% of total time and a

second around 16% of total time. Word-internal pauses account for only a small proportion of total time in editing/proofreading, typically less than 1% (median = .5%).

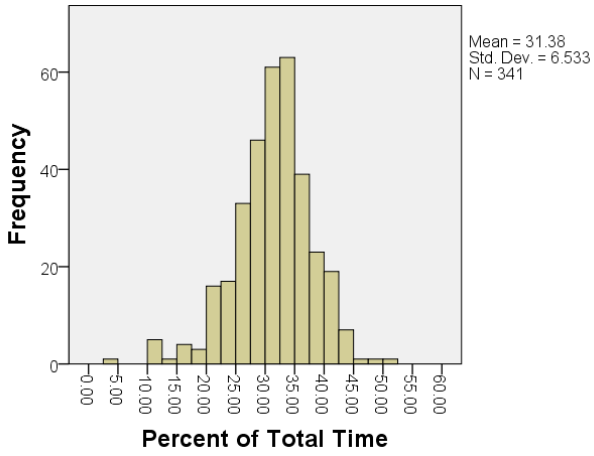

Figure 25. Distribution of in-word pause time during copy typing.

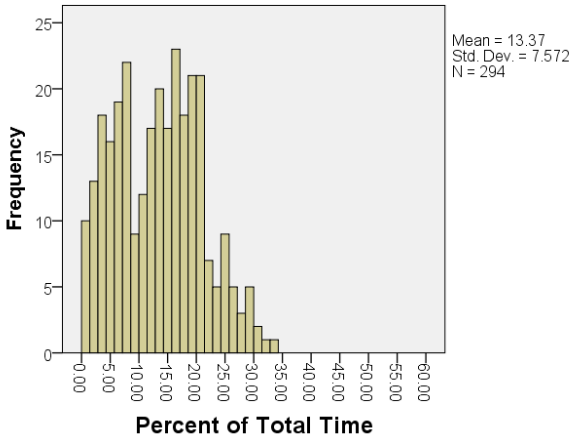

Figure 26. Distribution of in-word pause time during drafting.

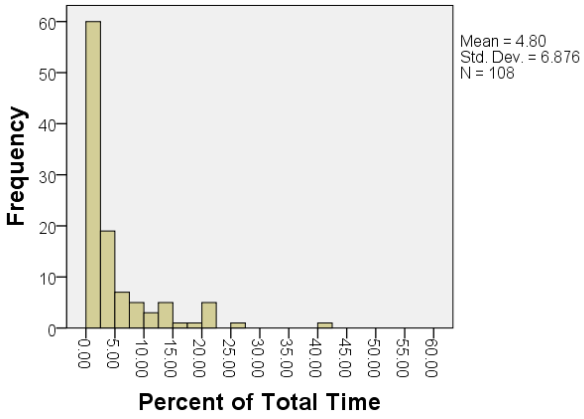

Figure 27. Distribution of in-word pause time during editing/proofreading.

A Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by word-internal pauses during copy typing was significantly greater than the median percentage of total time occupied by word-internal pauses during drafting, $Z = -14.06$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by word-internal pauses during drafting was significantly greater than the median percentage of total time occupied by word-internal pauses during editing/proofreading, $Z = -6.74$, $p < .001$. A third Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by word-internal pauses during copy typing was significantly greater than the median percentage of total time occupied by word-internal pauses during editing/proofreading, $Z = -8.45$, $p < .001$. Table 8 shows the median and first and third quartiles for word-internal pauses across the three tasks.

Table 8. Medians and First and Third Quartiles of Percentage of Total Time Occupied by Word-Internal Pauses for Copy Typing, Drafting, and Editing/Proofreading

Period	Copy typing (%)	Drafting (%)	Editing/proofreading (%)
Median	31.8	13.6	2.2
1st quartile	27.7	6.8	0.47
3rd quartile	35.3	19.0	5.4

Percentage word-initial pause time. A similar pattern can be observed with word-initial pauses, as shown by Figures 28–30. For copy typing, there is a single sharp peak around 28% of total time on task (median = 28.1%). For drafting, there again appears to be two peaks, one around 5% and the other around 25% (median = 24.4%). Editing/proofreading strongly peaks at low percentages, with the vast majority of writers spending less than 5% of their editing/proofreading time on word-initial pauses (median = 2.4%).

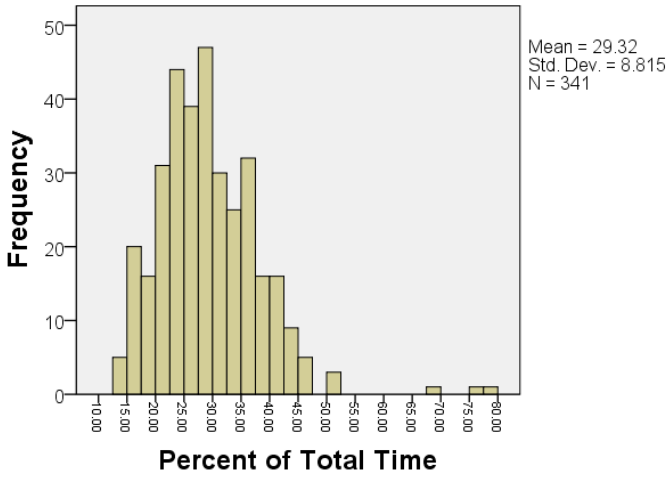

Figure 28. Distribution of word-initial pause time during copy typing.

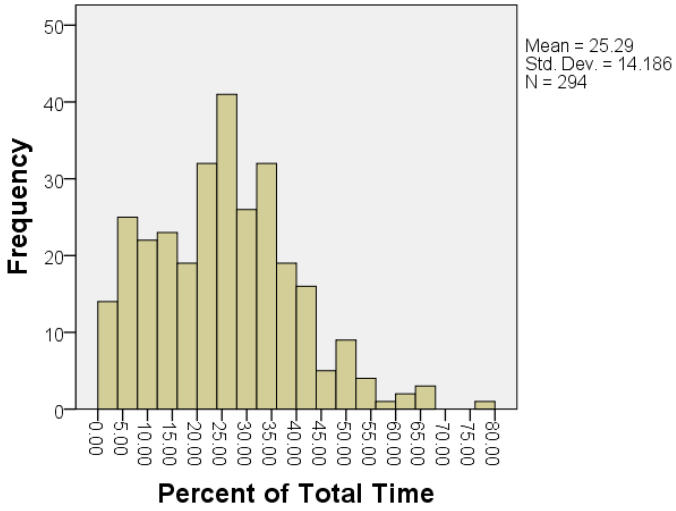

Figure 29. Distribution of word-initial pause time during drafting.

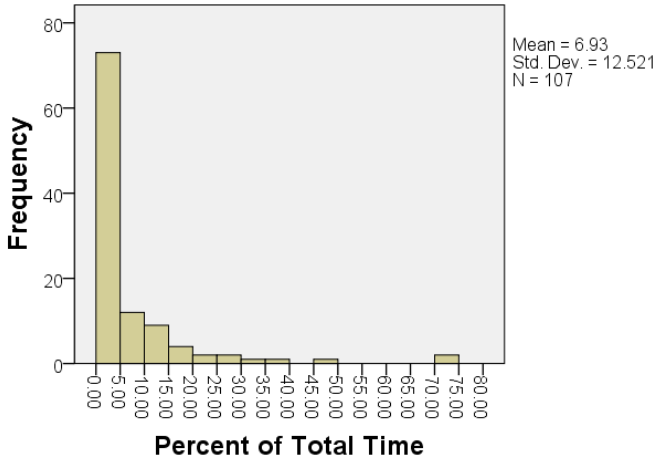

Figure 30. Distribution of word-initial pause time during editing/proofreading.

A Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by word-initial pauses during copy typing was significantly greater than the median percentage of total time occupied by word-initial pauses during drafting, $Z = -4.13$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by word-initial pauses during drafting was significantly greater than the median percentage of total time occupied by word-initial pauses during editing/proofreading, $Z = -7.03$, $p < .001$. A third Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by word-initial pauses during copy typing was significantly greater than the median percentage of total

time occupied by word-initial pauses during editing/proofreading, $Z = -7.89$, $p < .001$. Table 9 shows the median and first and third quartile values for word-initial pauses across the three tasks.

Table 9. Medians and First and Third Quartiles of Percentage of Total Time Occupied by Word-Initial Pauses for Copy Typing, Drafting, and Editing/Proofreading

Period	Copy typing (%)	Drafting (%)	Editing/proofreading (%)
Median	28.1	24.7	2.4
1st quartile	23.3	14.4	0.7
3rd quartile	35.0	34.4	7.5

Percentage between-word pause time. The distributions of between-word pauses across the three tasks are shown in Figures 31–33. Like word-internal and word-initial pauses, between-word pauses also display a sharply peaked distribution for copy typing, with a peak around 16% of total time (median = 15.8%). For drafting, there is also a sharply peaked distribution with a peak of around 8% of total time (median = 7.6%). For editing/proofreading, the vast majority of writers spend less than 2% of total time on between-word pauses (median = 1.4%).

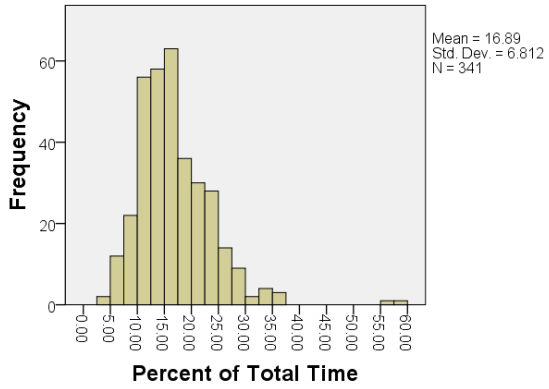

Figure 31. Distribution of between-word pause time during copy typing.

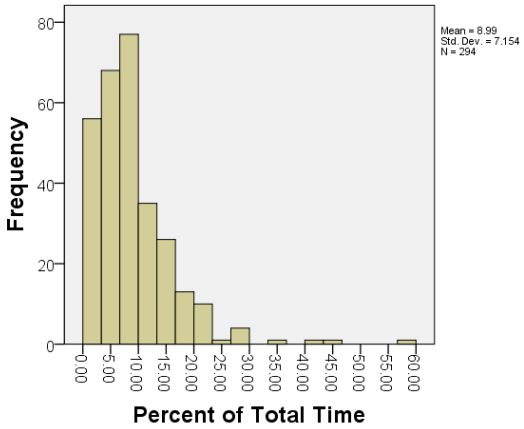

Figure 32. Distribution of between-word pause time during drafting.

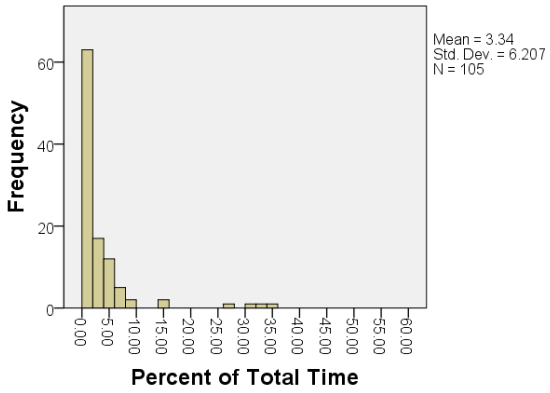

Figure 33. Distribution of between-word pause time during editing/proofreading.

A Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by between-word pauses during copy typing was significantly greater than the median percentage of total time occupied by between-word pauses during drafting, $Z = -12.15$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by between-word pauses during drafting was significantly greater than the median percentage of total time occupied by between-word pauses during editing/proofreading, $Z = -5.78$, $p < .001$. A third Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by between-word pauses during copy typing was significantly greater than the median percentage of total time occupied by between-word pauses during editing/proofreading, $Z = -7.54$, $p < .001$. Table 10 shows the medians and first and third quartile values for between-word pauses for each of the three tasks.

Table 10. Medians and First and Third Quartiles of Percentage of Total Time Occupied by Between-Word Pauses for Copy Typing, Drafting, and Editing/Proofreading

Period	Copy typing (%)	Drafting (%)	Editing/proofreading (%)
Median	15.8	7.6	1.4
1st quartile	12.3	4.5	0.3
3rd quartile	20.5	11.4	3.7

Percentage between-sentence pause time. The distributions of time spent on between-sentence pauses are shown in Figures 34–36. We again see a sharply peaked distribution, with most writers in the copy typing condition using 2%–6% of total time on between-sentence pauses (median = 3.9%). In the drafting condition, the most common responses involve very little time on between-sentence pauses (less than 1%), but there was a very long tail of writers who used much more time on between-sentence pauses, up to 15% of total time (median = 2.2%). In the

retyping condition, the vast majority of writers used less than 1% of total time on between-sentence pauses (median = 0.2%).

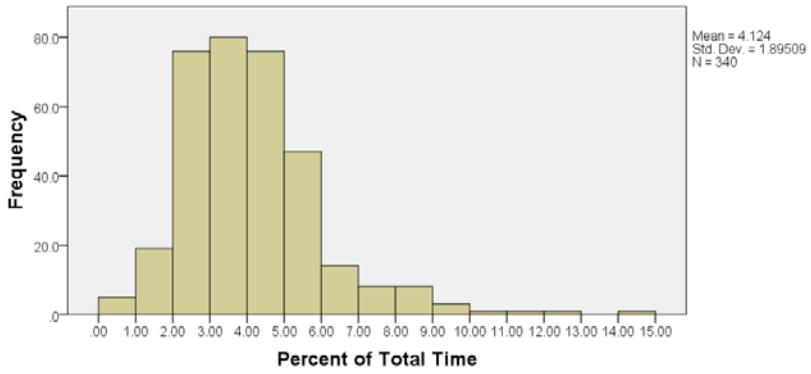

Figure 34. Distribution of between-sentence pause time during copy typing.

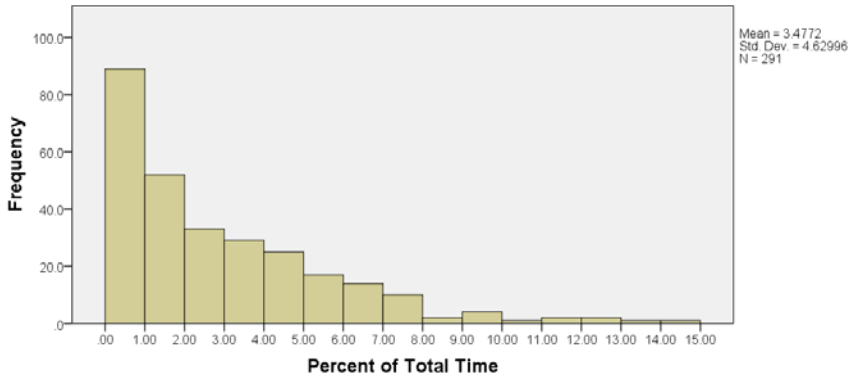

Figure 35. Distribution of between-sentence pause time during drafting.

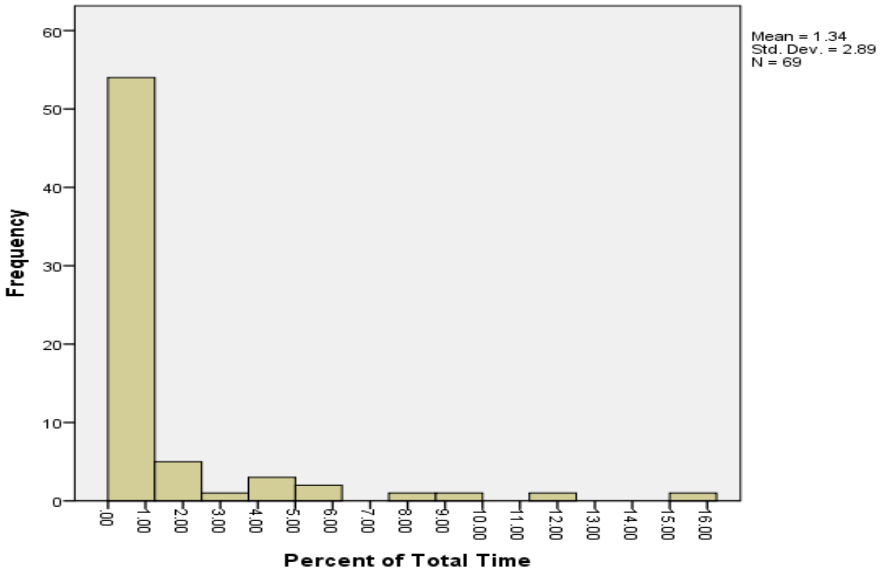

Figure 36. Distribution of between-sentence pause time during editing/proofreading.

A Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by between-sentence pauses during copy typing was significantly greater than the median percentage of total time occupied by between-sentence pauses during drafting, $Z = -5.34$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by between-sentence pauses during drafting was significantly greater than the median percent of total time occupied by between-sentence pauses during editing/proofreading, $Z = -3.60$, $p < .001$. A third Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by between-sentence pauses during copy typing was significantly greater than the median percentage of total time occupied by between-sentence pauses during editing/proofreading, $Z = -5.84$, $p < .001$. Table 11 shows the medians and first and third quartile values for between-sentence pause time in the three conditions.

Table 11. Medians and First and Third Quartiles of Percentage of Total Time Occupied by Between-Sentence Pauses for Copy Typing, Drafting, and Editing/Proofreading

Period	Copy typing (%)	Drafting (%)	Editing/proofreading (%)
Median	3.9	2.2	0.2
1st quartile	2.8	0.8	0.1
3rd quartile	5.0	4.6	0.6

Percentage backspace pause time. Figures 37–39 show the distributions of total time on backspace events as a proportion of total time for each of the three tasks. Once again, copy typing is sharply peaked around 10% (median = 9.6%). Drafting is less sharply peaked and has a long tail of writers who pause for a relatively large percentage of total time, between 20% and 30% (median = 9.8%). The distribution of editing/proofreading backspaces has a modal group who spend less than 1% of total time on backspacing, but there is a long tail of writers who paused for larger percentages of the time on backspace events (median = 2.7%).

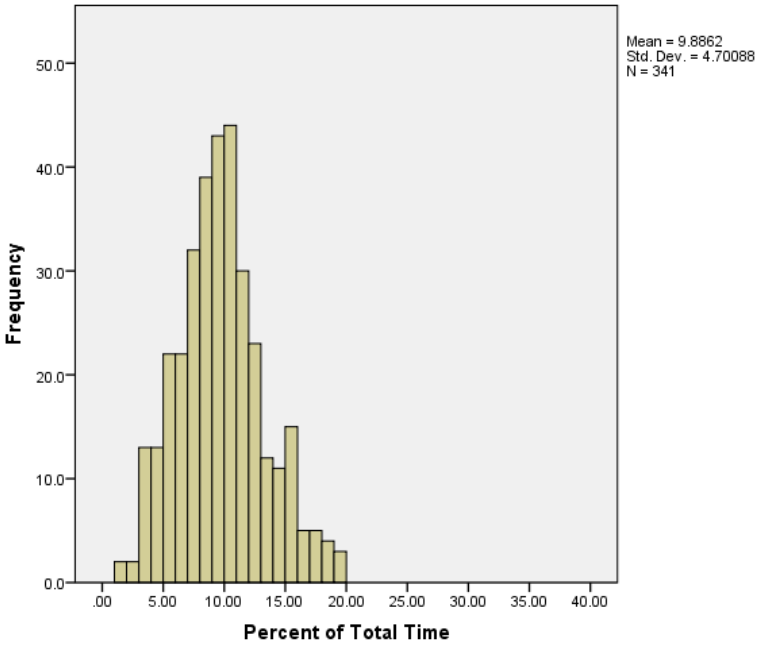

Figure 37. Distribution of backspace time during copy typing.

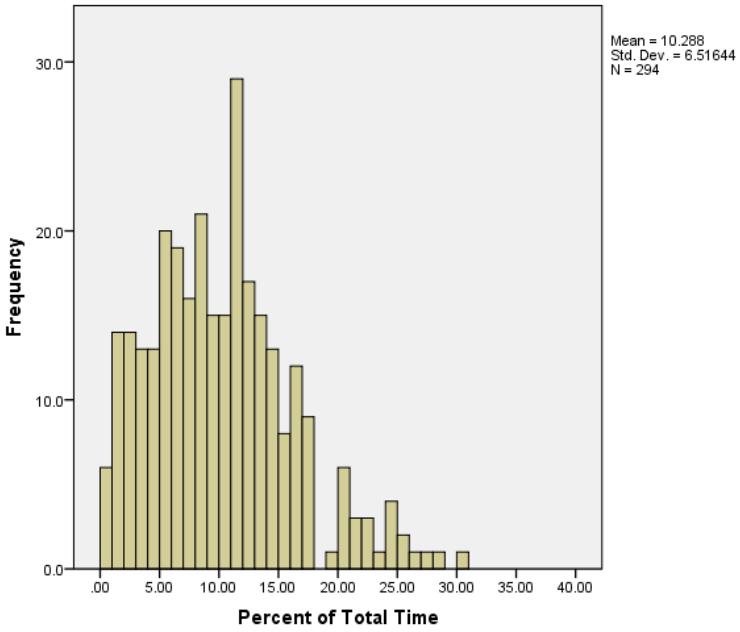

Figure 38. Distribution of backspace time during drafting.

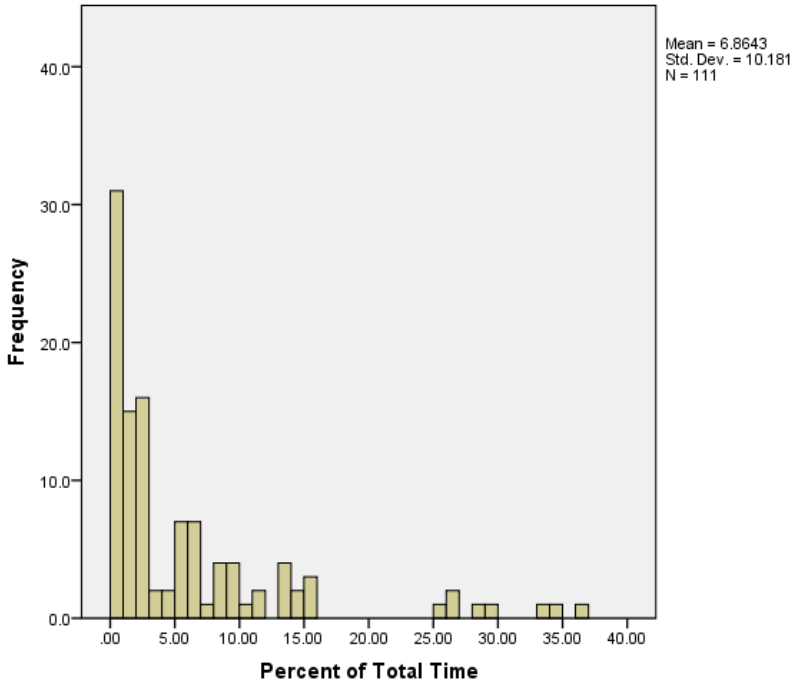

Figure 39. Distribution of backspace time during editing/proofreading.

A Wilcoxon signed-ranks test indicated no significant difference between the median percentage of total time occupied by backspace pauses between copy typing and drafting, $p = .75$. A second Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by backspace pauses during drafting was significantly greater than the median percentage of total time occupied by backspace pauses during editing/proofreading, $Z = -4.91$, $p < .001$. A third Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by backspace pauses during copy typing was significantly greater than the median percentage of total time occupied by backspace pauses during editing/proofreading, $Z = -4.91$, $p < .001$. Table 12 shows the medians and first and third quartile values for backspace pause time for each of the three tasks.

Table 12. Medians and First and Third Quartiles of Percentage of Total Time Occupied by Backspace Pauses for Copy Typing, Drafting, and Editing/Proofreading

Period	Copy typing (%)	Drafting (%)	Editing/proofreading (%)
Median	9.6	9.8	2.7
1st quartile	7.5	5.7	0.8
3rd quartile	11.8	13.7	8.3

Percentage local edit time. A Wilcoxon signed-ranks test indicated that there was no significant difference in percentage of total time occupied by local edit pauses between copy typing and drafting, $p = .94$. A second Wilcoxon signed-ranks test indicated that there was no significant difference in percentage of total time occupied by local edit pauses between drafting and editing/proofreading, $p = .51$. A third Wilcoxon signed-ranks test indicated that there is no significant difference in the percentage of total time occupied by local edit pauses between copy typing and editing/proofreading, $p = .86$.

Percentage jump-to-edit time. Jump-to-edit time displayed rather distinct patterns across the three tasks, as Figures 40–42 illustrate. During copy typing, the modal percentage of time used for jump-to-edit keystrokes was very small, less than 5% of total time, but there was a long tail reflecting writers who spent significant time making jump edits (median = 5.8%). During drafting, the modal percentage of time used for jump-to-edit keystrokes was larger, around 15% of total time, but writers varied widely in the likelihood that they would perform jump-to-edit actions, with some writers spending 80% or 90% of their time pausing before jump edits (median = 29.0%). By contrast, the modal pattern for editing/proofreading was for writers to spend almost all their time pausing before a jump, though there was a long tail of writers who used a much smaller percentage of total time (median = 85.4%).

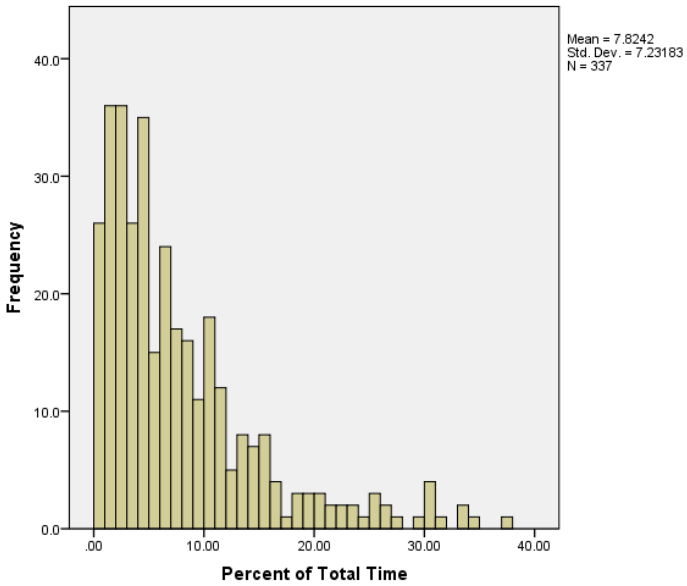

Figure 40. Distribution of jump to edit during copy typing.

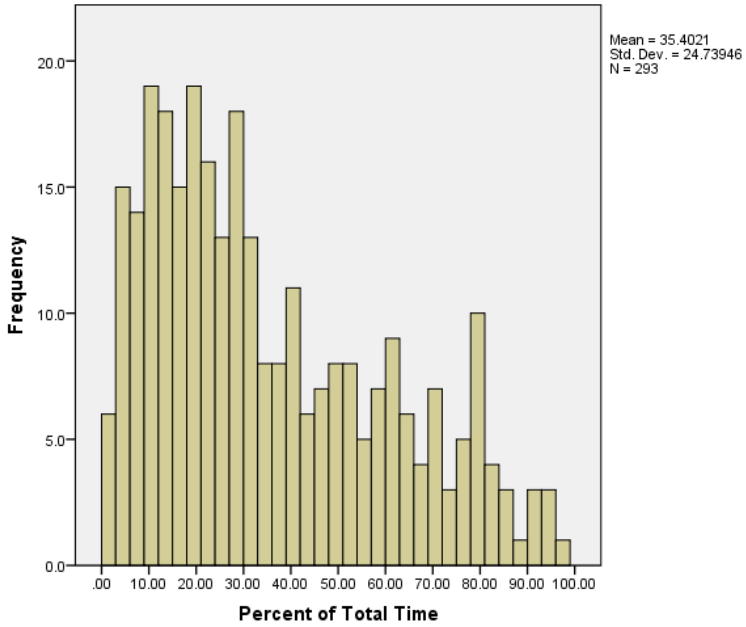

Figure 41. Distribution of jump to edit during drafting.

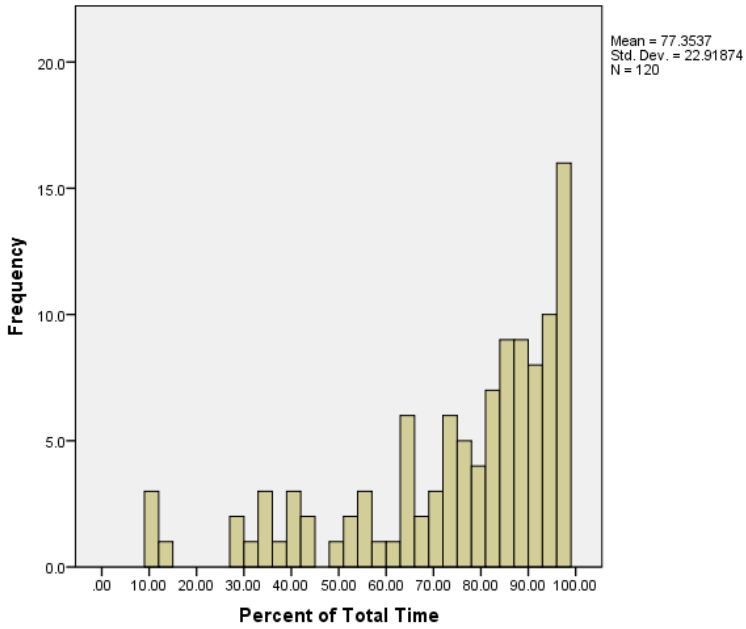

Figure 42. Distribution of jump to edit during editing/proofreading.

A Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by jump-to-edit pauses during copy typing was significantly less than the median percentage of total time occupied by jump-to-edit pauses during editing/proofreading, $Z = 13.30$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by jump-to-edit pauses during drafting was significantly less than the median percentage of total

time occupied by jump-to-edit pauses during editing/proofreading, $Z = 7.83$, $p < .001$. A third Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by jump-to-edit pauses during copy typing was significantly less than the median percentage of total time occupied by jump-to-edit pauses during editing/proofreading, $Z = 8.75$, $p < .001$. Table 13 shows the medians and first and third quartile values for the percentage of time writers spent on jump-to-edit pauses under the three conditions.

Table 13. Medians and First and Third Quartiles of Percentage of Total Time Occupied by Jump-to-Edit Pauses for Copy Typing, Drafting, and Editing/Proofreading

Period	Copy typing (%)	Drafting (%)	Editing/proofreading (%)
Median	5.8	29.0	85.4
1st quartile	2.5	15.1	65.8
3rd quartile	10.6	53.2	95.6

Percentage start time. The distributions of time spent on the initial pause before typing begins varied considerably across the three tasks, though most students started writing quite quickly in all three tasks, as Figures 43–45 demonstrate. However, in both drafting and editing/proofreading, there was a long tail of writers who spent relatively large proportions of total time pausing before they began to write (copy typing median = 0.02%, drafting median = 1.2%, editing/proofreading median = 3.6%).

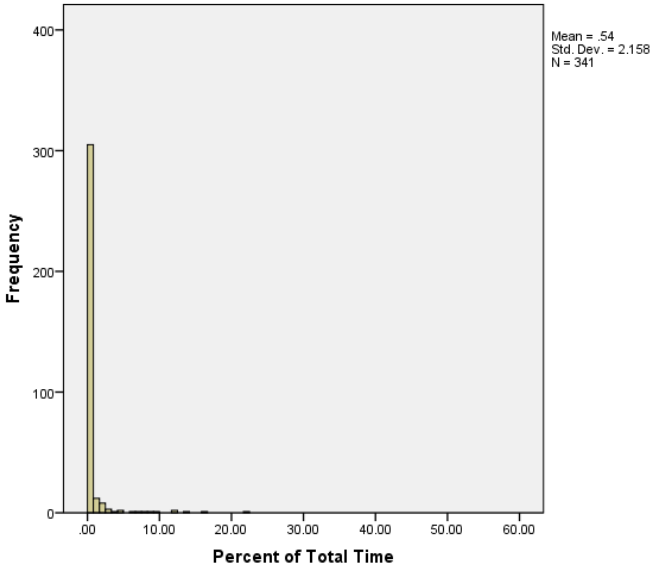

Figure 43. Distribution of initial pause time during copy typing.

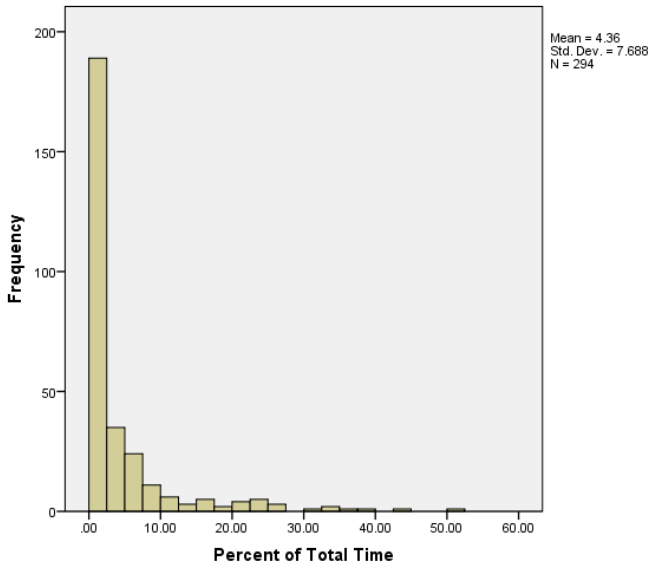

Figure 44. Distribution of initial pause time during drafting.

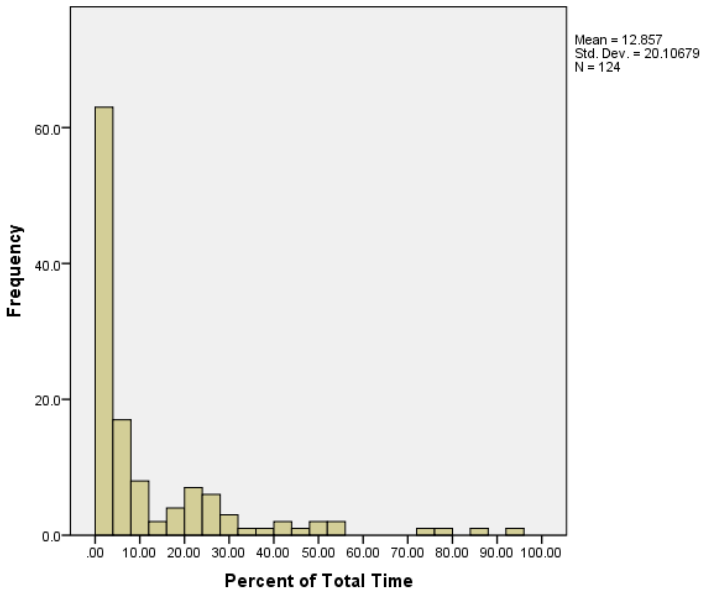

Figure 45. Distribution of initial pause time during editing/proofreading.

A Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by the initial pause during copy typing was significantly less than the median percentage of total time occupied by the initial pause during drafting, $Z = 10.77$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by the initial pause during drafting was significantly less than the median percentage of total time occupied by the initial pause during editing/proofreading, $Z = 4.02$, $p < .001$. A third Wilcoxon signed-ranks test

indicated that the median percentage of total time occupied by the initial pause during copy typing was significantly less than the median percentage of total time occupied by the initial pause during editing/proofreading, $Z = 8.17$, $p < .001$. Table 14 displays the medians and first and third quartile values for initial pauses for each of the three tasks.

Table 14. Medians and First and Third Quartiles of Percentage of Total Time Occupied by the Initial Pause for Copy Typing, Drafting, and Editing/Proofreading

Period	Copy typing (%)	Drafting (%)	Editing/proofreading (%)
Median	0.02	1.2	3.6
1st quartile	0.01	0.3	0.7
3rd quartile	0.06	4.7	19.2

Percentage long pause time. If we aggregate long pauses (>2 s) regardless of the specific keystroke action involved (see Figures 46–48), we find that long pauses constituted a large part of writers' behaviors even during copy typing. That distribution peaks around 49% of total time (median = 45.1%), whereas the distribution of long pause time for drafting spreads out across a broad range, with most writers spending at least 50% of their time in long pauses (median = 72.7%). During editing/proofreading, the tendency to spend time in long pauses is even more marked, with most writers spending more than 80%, and often nearly 100%, of their time in long pauses (median = 98.1%).

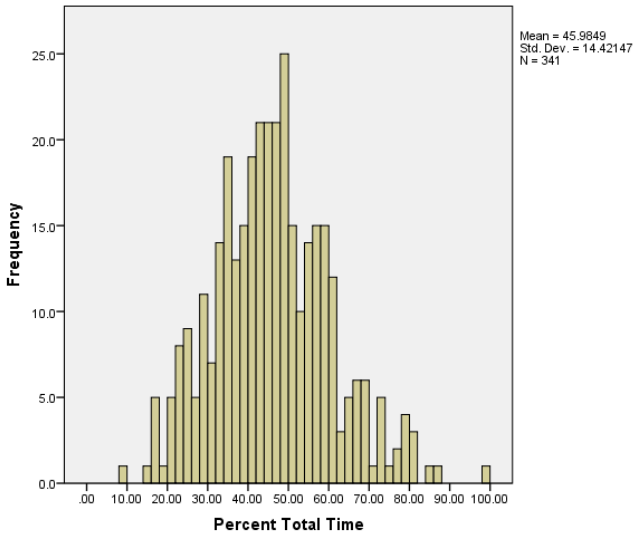

Figure 46. Distribution of long pause time during copy typing.

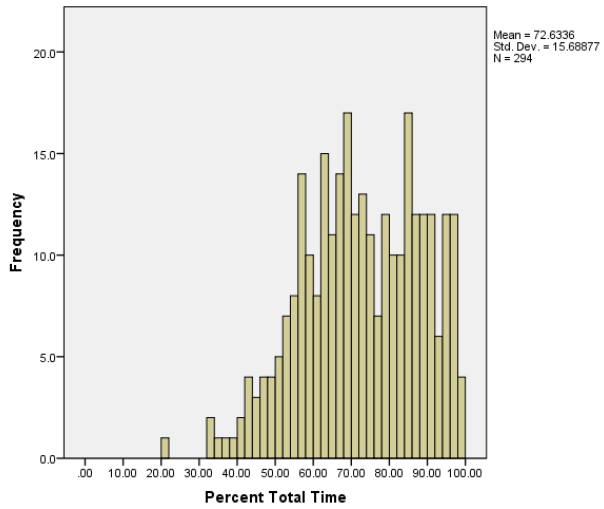

Figure 47. Distribution of long pause time during drafting.

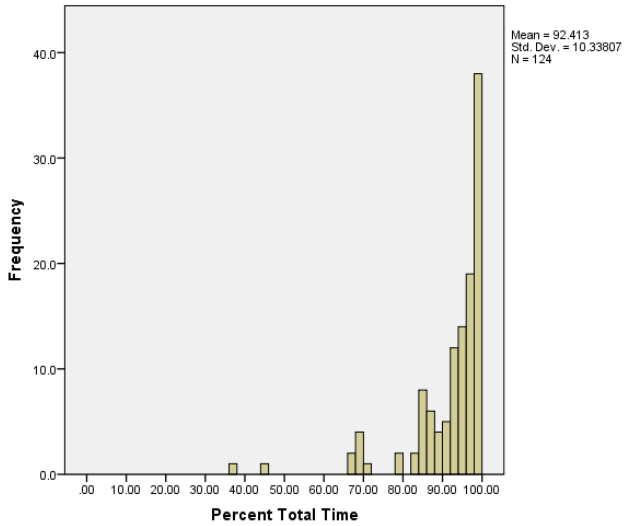

Figure 48. Distribution of long pause time during editing/proofreading.

A Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by the initial pause during copy typing was significantly less than the median percentage of total time occupied by the initial pause during drafting, $Z = 13.54$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by long pauses during drafting was significantly less than the median percentage of total time occupied by long pauses during editing/proofreading, $Z = 7.95$, $p < .001$. A third Wilcoxon signed-ranks test indicated that the median percentage of total time occupied by long pauses during copy typing was significantly less than the median percentage of total time occupied by long pauses during

editing/proofreading, $Z = 9.06$, $p < .001$. Table 15 shows the medians and first and third quartile values for long pauses for each of the four tasks.

Table 15. Medians and First and Third Quartiles of Percentage of Total Time Occupied by Long Pauses (>2 s) for Copy Typing, Drafting, and Editing/Proofreading

Period	Copy typing (%)	Drafting (%)	Editing/proofreading (%)
Median	45.1	72.7	98.1
1st quartile	35.7	61.1	89.5
3rd quartile	55.5	85.7	98.7

Differences in Other Features

The remaining features we examined help to characterize the kinds of editing that took place during each task type. They indicate the following:

- Length of jump edits is a major differentiator between the three tasks, with copy typing involving shorter jump edits than drafting and drafting involving shorter jump edits than editing/proofreading. There is also a moderate association between the log length of jump edits and their log length in seconds.
- Drafting involves a significantly larger proportion of multiword deletions than the other tasks.
- Editing/proofreading is differentiated by the fact that a significantly larger number of individual words are edited than in the other tasks.

The details of these analyses are presented in the following sections.

Median jump length. As Figures 49–51 demonstrate, jump-to-edit events are distributed very differently across the three tasks. The modal length of jumps to edit in copy typing is less than six characters, though there was a long tail of more extended jump edits (median = 1.79 log characters). The modal length of jumps to edit in drafting is also less than six characters, though there the shape of the distribution suggests a mixture of two different processes, one with very short jumps and another with relatively long jumps (median = 3.26 log characters). The distribution of jumps to edit in editing/proofreading is clearly bimodal, with one peak corresponding to the short jumps that are modal for the other two tasks and the second peak corresponding to very long pauses (median = 5.09 log characters). Moreover, log pause time during jumps to edit appears to be moderately associated with log jump length ($R = .50$, $R^2 = .25$), as Figure 52 attests.

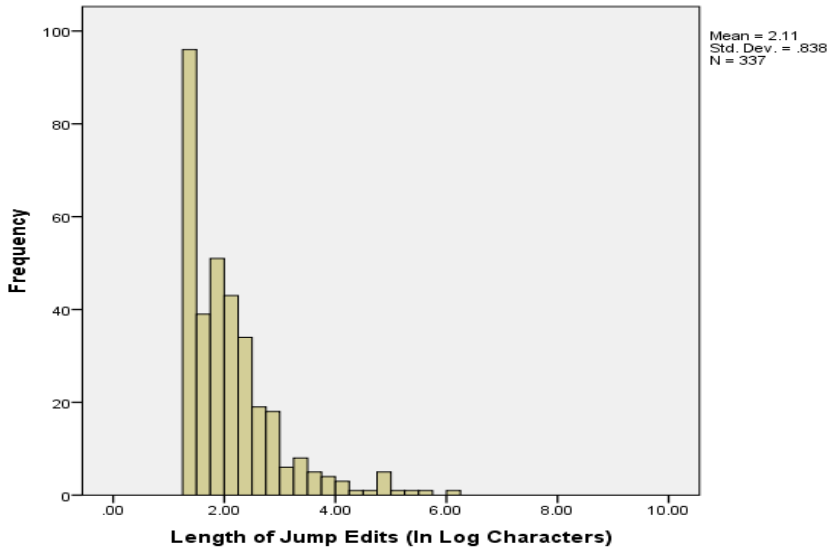

Figure 49. Distribution of jump edits by length during copy typing.

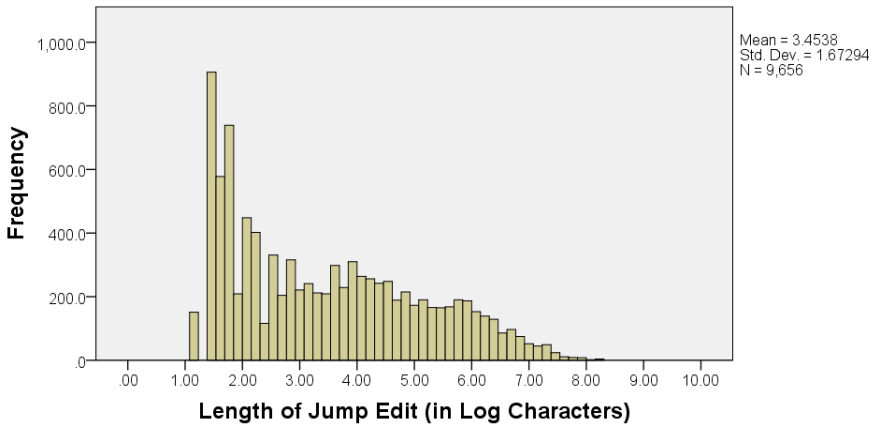

Figure 50. Distribution of jump edits by length during drafting.

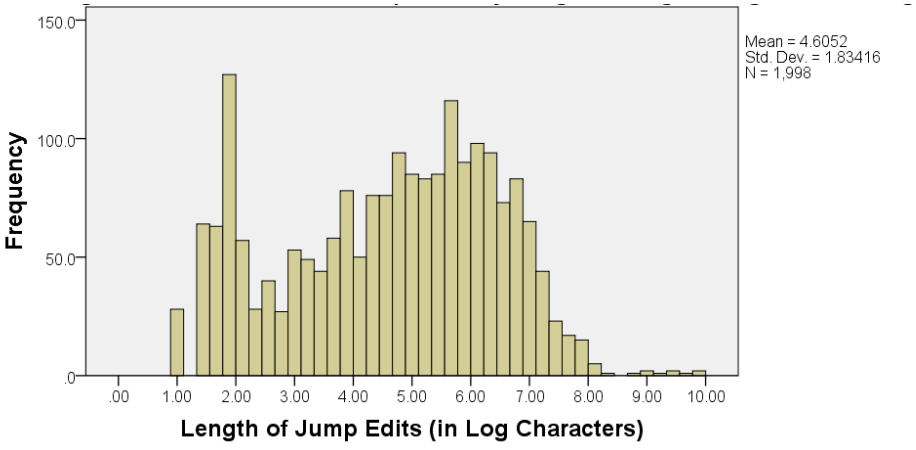

Figure 51. Distribution of jump edits by length during editing/proofreading.

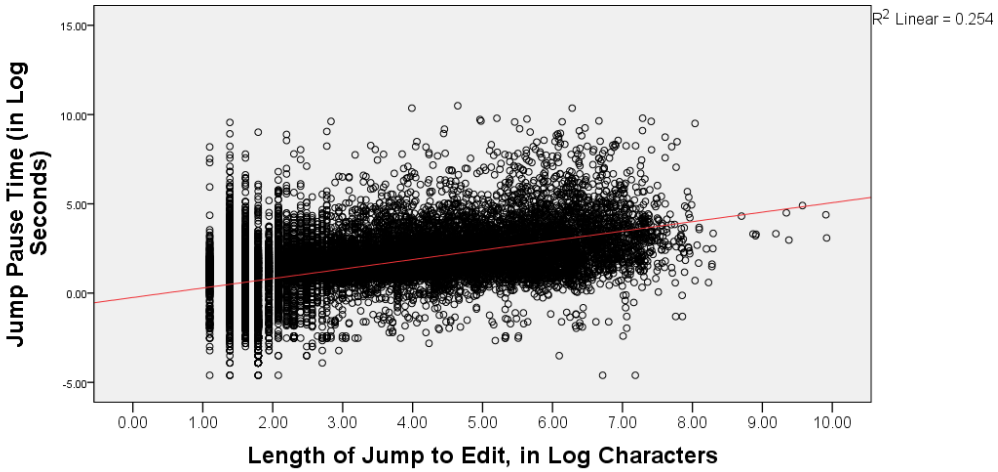

Figure 52. Relation between jump length and jump pause time.

When we calculate individual medians by task and examine how the resulting measure differs by task, we observe significant differences. A Wilcoxon signed-ranks test indicated that the median of per-individual median log jump lengths during copy typing was significantly less than the median of per-individual median log jump lengths during drafting, $Z = 12.06$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the median of per-individual median log jump lengths during drafting was significantly less than the median of per-individual median log jump lengths during editing/proofreading, $Z = 7.81$, $p < .001$. A third Wilcoxon signed-ranks test indicated that the median of per-individual median log jump lengths during copy typing was significantly less than the median of per-individual median log jump lengths during editing/proofreading, $Z = 8.75$, $p < .001$. Table 16 shows the medians and first and third quartile values for individual median jump lengths for each of the four tasks.

Table 16. Medians and First and Third Quartiles of Individual Median Jump Lengths for Copy Typing, Drafting, and Editing/Proofreading, in Log Characters

Period	Copy typing	Drafting	Editing/proofreading
Median	1.79 (6)	3.26 (26)	5.09 (162.5)
1st quartile	1.39 (4)	2.46 (12)	4.49 (89.5)
3rd quartile	2.44 (12)	3.97 (53)	5.78 (325)

Proportion of multiword deletion events. When we compare proportions of multiword deletion events by task, we observe significant differences (copy typing median = 3.5%, drafting median = 17.1%, editing/proofreading median = 0%). A Wilcoxon signed-ranks test indicated that the median proportion of multiword deletions is significantly smaller in copy typing than in

drafting, $Z = 12.69$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the median proportion of multiword deletions is significantly larger in drafting than in editing/proofreading, $Z = -7.76$, $p < .001$. A third Wilcoxon signed-ranks test indicated no significant difference in the proportion of multiword deletions in copy typing and editing/proofreading. Table 17 provides the medians and first and third quartiles for each task.

Table 17. Medians and First and Third Quartiles of Proportion of Multiword Deletions for Copy Typing, Drafting, and Editing/Proofreading

Period	Copy typing (%)	Drafting (%)	Editing/proofreading (%)
Median	3.5	17.1	0.0
1st quartile	0.0	3.6	0.0
3rd quartile	1.2	60.5	2.0

Proportion of edited words. When we compare proportions of individually edited words (i.e., words that had at least one character changed while the word was being typed), we observe significant differences across tasks (copy typing median = 1.3%, drafting median = 1.5%, editing/proofreading median = 2.3%). A Wilcoxon signed-ranks test indicated that the median proportion of edited words is significantly smaller in copy typing than in drafting, $Z = 4.99$, $p < .001$. A second Wilcoxon signed-ranks test indicated that the median proportion of edited words is significantly smaller in drafting than in editing/proofreading, $Z = 5.03$, $p < .001$. A third Wilcoxon signed-ranks test indicated that the median proportion of edited words in copy typing is significantly smaller than in editing/proofreading, $Z = 5.65$, $p < .001$. Table 18 provides the medians and first and third quartiles for each task.

Table 18. Medians and First and Third Quartiles of Proportion of Individually Edited Words for Copy Typing, Drafting, and Editing/Proofreading

Period	Copy typing (%)	Drafting (%)	Editing/proofreading (%)
Median	1.3	1.5	2.3
1st quartile	1.0	1.2	0.9
3rd quartile	1.7	2.0	4.6

Correlations Across Tasks

Table 19 displays Pearson correlations between the median pause times for each keystroke event type across tasks. As this table demonstrates, word-internal pauses are strongly correlated between the copy typing and drafting tasks (.81). Certain other production events (between-word and between-sentence pauses) have moderate correlations between copy typing

and drafting, but otherwise, correlations are weak to nonexistent. In fact, the editing/proofreading task has weak correlations with other tasks even with respect to word-internal and between-word pauses and no other significant correlations with either of the other tasks.

Table 19. Correlations Across Tasks for Median Pause Duration Features

Feature	Copy typing/drafting	Copy typing/editing	Drafting/editing
Word-internal pauses	.81**	.08	.24*
Word-initial pauses	.21**	.16	.14
Between-word pauses	.56**	.22*	.32**
Between-sentence pauses	.39**	.21	.14
Backspace pauses	.12*	−.08	.04
Local edit pauses	−.07	.20	.01
Jump-to-edit pauses	−.06	.03	−.04

*$p < .05$. **$p < .01$.

Discussion

The results we have just reviewed indicate that writing process logs have very different characteristics, depending on the task. These differences are consistent with the predictions we made based on our cognitive theory.

Characteristics of the Copy Typing Task

We expected, based on the underlying cognitive processes involved, that the copy typing task would reflect transcription skills very directly and would have little involvement of the evaluator (except for checking accuracy of spelling/typing), translator, or proposer. However, we also expected copy typing to impose a significant executive processing load, because the writer must visually scan the text to be copied, remember its content, and actively switch attention back and forth between the text to be copied and the keyboard and/or the text being produced.

Action of the proposer. Because no idea generation is taking place—only scanning to find and remember content to be copied—we predicted that percentage start time and the percentage of time in long pauses would be significantly lower for copy typing than for drafting. Our results confirm this prediction.

Action of the translator. Because no translation is taking place and copy typing is primarily driven by the "outer loop," during which writers visually scan the next word or short phrase to be copied, we predicted that copy typing would typically involve very short bursts (i.e., single words or short phrases rather than clauses or sentences). Because the outer loop is also supposed to operate word by word, resulting in consistently longer word-initial pauses as each

word is successively scanned, we also predicted that word-initial pause durations would be greater in copy typing than in drafting. Our results confirm these predictions.

Action of the transcriber. Because we expected the visual scanning that takes place during copy typing to add to working memory load and therefore interfere with transcription fluency, we predicted that the duration of word-internal pauses would be greater in copy typing than in drafting. Our results confirm this prediction.

Action of the evaluator. Because copy typing should only invoke the evaluator to catch and correct orthographic errors in the most recently produced text, we predicted that copy typing would involve fewer editing actions and that these actions would tend to be shorter, take a smaller percentage of total time on task, and typically involve local jumps to edit recently produced text than in drafting or editing/proofreading. Our results confirm these predictions.

Characteristics of the Essay Drafting Task

We expected, based on the underlying cognitive processes involved, that the drafting task would primarily reflect the functions of the translator and the transcriber, with some involvement of the proposer if ideas are generated during drafting rather than during prior planning activities. We expected that the action of the evaluator during drafting would differ from its role in copy typing by monitoring text produced and evaluating whether the ideas or their expression communicated what the writer intended.

Action of the proposer. On the basis of the preceding assumptions, we predicted that initial pauses and long pauses would take up a larger percentage of total time in drafting than in copy typing. This prediction was confirmed.

Action of the translator. We also predicted that the process of generating novel text would force writers to pause for long periods more often at natural boundaries (e.g., between words and sentences), resulting in them taking a larger percentage of total time on the event types where those pauses would most likely fall (word initially or between sentences). We further predicted that writers would produce longer bursts of text in drafting, corresponding to phrase- or clause-length conceptual units. By contrast, we expected copy typing and editing/proofreading to involve shorter bursts driven by the visual inspection of the text to be copied or edited. These predictions were confirmed.

Action of the transcriber. We also predicted that word-internal pauses would be shorter in drafting than in either copy typing or editing/proofreading, because both of those tasks impose

an additional visual executive control process that can be expected to increase cognitive load and decrease the working memory available for transcription. This prediction was confirmed.

Action of the evaluator. We predicted that drafting would involve more cases where multiple-word sequences were deleted, often by backspacing over the entire text to be deleted. Because, in many cases, the evaluator might cause the writer to decide to modify rather than delete the current conceptual unit being produced, we also predicted that drafting would involve more jump-to-edit actions than copy typing, that these jumps would far more often extend to edits beyond the last word or two produced than in copy typing, and that the writer would pause longer before them, resulting in jump-to-edit actions accounting for a larger proportion of total time in drafting than in copy typing. These predictions were confirmed.

Characteristics of the Editing/Proofreading Task

We expected, based on the underlying cognitive processes involved, that the editing/proofreading task would most strongly reflect the functions of the evaluator. This entails less emphasis on idea generation, translation, and transcription events; greater frequency of editing actions; and more time spent in particular on jump-to-edit actions, in which writers scan all or large portions of the text to find locations where changes need to be made.

Action of the proposer. Very little idea generation should take place during editing/proofreading. Initial pauses and long pauses should instead represent deliberation time during which writers are scanning the text to identify issues. Because the combination of scanning, evaluation, and planning text changes is likely to impose a heavy load on executive processing and attention, we predicted that percentage start time and the percentage of time in long pauses would be significantly greater for editing/proofreading than for drafting. Our results confirm this prediction.

Action of the translator. Very little translation should take place during editing/proofreading. Most bursts of keystrokes should consist of the execution of changes (i.e., deletions and insertions) that were identified and planned between bursts. We therefore predicted that editing/proofreading, like copy typing, would typically involve very short bursts (i.e., edits to single words or modifications to short phrases rather than production of whole clauses or sentences). We expected that the increased memory load demanded to evaluate errors, plan edits, and monitor corrections for accuracy would reduce the resources available to plan each burst and

would therefore result in longer word-initial pauses in editing/proofreading than in drafting. Our results confirm these predictions.

Action of the transcriber. Because we expected that the visual scanning, evaluation, and planning of edits that take place during editing/proofreading would add to working memory load and therefore interfere with transcription fluency, we predicted that the duration of word-internal pauses would be greater in editing/proofreading than in drafting. Our results confirm this prediction.

Action of the evaluator. Because editing/proofreading is driven by the evaluator and because the time needed to scan the text, evaluate it for errors, and plan edits should be considerable, we predicted that editing/proofreading would involve more editing actions and that these actions would tend to be longer in duration, take a larger percentage of total time on task, and typically involve global jumps to almost any location in the text as compared to drafting or copy typing. Our results confirm these predictions.

General Discussion

In this analysis, we have primarily been concerned with relatively easily measured features: the number, duration, and total time spent on certain types of keystroke events; the distribution of initial pauses and long pauses; the length of jump-to-edit actions; and the proportion of time spent in multiword deletions and single-word editing actions. The differences we have observed between copy typing, drafting, and editing/proofreading are consistent with what we would expect on the basis of a cognitive account of writing processes. Our observations raise important issues that need to be taken into account when interpreting a writing process log and suggest specific hypotheses that should be explored in future work.

One point that cannot be overemphasized is that many of the features we have examined mean different things in different tasks and have very different distributions as a result. For instance, the duration of initial pauses and of long pauses more generally seems likely to reflect very different cognitive processes, depending on the task. Jump-to-edit actions reflect the action of the evaluator, but their distribution both in time and in length of jump suggests that observed jump-to-edit actions may reflect three different kinds of evaluation processes, reflecting immediate monitoring of typing, immediate monitoring of translation and/or idea generation, and more global evaluation of the text produced. The results suggest, more generally, that we need to

obtain a much better understanding of how the distribution of keystroke actions is affected by the cognitive processes that drive performance.

The kind of patterns we observe suggest that it should be possible to identify characteristic distributions for certain writing processes, for instance, a pattern of very long pauses followed by fast, short edit sequences for the evaluation process typical of proofreading or a pattern of a long pause followed by a relatively long burst, possibly followed by deletion or editing of the text produced, as characteristic of fluent drafting. To the extent that we can build up such profiles, we may be able to characterize the patterns of performance in a writing process log in a way that characterizes what mix of cognitive processes controlled performance during different portions of an individual's writing process.

Conclusions, Limitations, and Future Directions

This study indicates that when we change the task to emphasize different parts of the writing process, the resulting writing process logs can differ significantly in interpretation. These differences and the failure of most of the writing process features to correlate strongly across those tasks underlines the extent to which writing is a complex skill that requires the coordination of multiple processes. However, the logs appear to provide consistent information across tasks about one component process—transcription—and to provide significant information about the mix of processes individual writers are bringing to bear while they write.

Limitations

The results we report are in many ways preliminary. We sampled student writing from a single school and grade level with students predominantly drawn from a largely minority population with low SES and limited English proficiency. We did not analyze logs for all students in our population due to technical issues with the keystroke log collection process, and we have not yet scored the student essays or compiled scores for the work students did during the prewriting process, which was collected in paper form. Some of the details of student behavior that we have observed may have been affected by specifics of our implementation, such as the requirement that students paste their plans into the essay submission box so that they would be readily available to the students when they began drafting. Caution should therefore be exercised in generalizing these results to other tasks or other populations or in drawing inferences about how the patterns we have observed relate to writing ability or other cognitive skills.

In addition, because our analysis focused on task differences and did not examine individual differences in performance or motivation, we must also allow for the possibility that the pattern of results we observed may vary if we sample a population with a different mix of ability levels or under circumstances that affect student motivation. For example, the high proportion of long pauses in the editing task might reflect greater effort focused on evaluation, but it might also reflect a higher level of frustration with or disengagement from the editing task, particularly for students who have very little idea how to improve an existing text through the editing process. To the extent that this is the case, we might observe very different patterns of long pauses during editing, depending on skill and interest levels. It is impossible to determine the extent to which this concern is valid without conducting further studies.

It is also important to note that this report is, by intention, focused on the distribution of a small number of features we believe may provide evidence about different writing processes. Given the large differences between the tasks and the relatively low correlations between features across tasks, our results also raise important issues about the use of writing process features for assessment purposes. If student performance reflects a different mix of writing processes across tasks, it would be very easy to draw inappropriate conclusions from summary features that might in actual fact mean different things, depending on the exact nature of the writing task assigned in a particular assessment. In particular, the almost complete lack of correlation between parallel features between editing/proofreading and other tasks suggests that heavy involvement of the evaluator may have significant impact on a wide range of features, whether they would most obviously be linked to the action of the evaluator or not.

Implications for Improving the Keystroke Log Analysis Process

Our results indicate that the tempo and distribution of keystroke actions vary considerably when writers undertake tasks as different as retyping, drafting, and editing/proofreading. One possible implication of this result is that it suggests that it should be possible to classify events in a keystroke log dynamically, by taking into account the pattern of keystroke events in the immediate neighborhood. If a writer switches from direct text production (translation and transcription of existing ideas) to some other mode, whether for purposes of idea generation or editing or merely to copy a quotation from a source text, our results suggest that this switch will create markers in the keystroke log that should make it possible to identify that the writer has made such a switch between writing subtasks. Such a dynamic classification of

keystroke log events might allow us to infer what writing processes are engaged and hence what the writer is currently attempting to accomplish.

However, our results emphasize the extent to which individual writing process features may be ambiguous in their interpretation. To achieve reliable inferences about student writing processes, it will probably be necessary to aggregate over multiple indicators. In fact, the patterns we have observed in this study provide a first step in that direction. For example, if a writer shifts from typical drafting behaviors to typical editing behaviors during the same writing session, we may be well justified in inferring that the writer has transitioned from drafting to editing, even if we could not draw that conclusion by examining many of the individual features in isolation.

Making these kinds of inferences will probably require considerably more information about how writers switch between tasks and what the distribution of keystroke events looks like when different cognitive processes are in play than what we can support in the current study. However, our results do suggest specific patterns—such as the duration of and distance traversed by jump-to-edit events—that may indicate what kinds of evaluation processes a writer is performing at particular points in the writing process.

Implications for Assessment

The results we have obtained also have significant implications when we combine them with specific results obtained in prior research. In particular, they suggest that getting accurate measurement of student writing skill may require different strategies than currently applied in most direct writing assessments.

In prior studies (Deane 2014; Deane & Zhang, 2015; Zhang & Deane, 2015), we examined keystroke log patterns collected from direct writing assessments, where students were expected to draft written essays in 35–45 minutes (though they were provided with 45 minutes of preparatory work on the same topic, as part of a scenario-based assessment). In these studies, we observed a pattern in which writing fluency and time on task were positively connected with essay score and in which deletion of text already produced, long pauses at the start of writing, and longer word-initial pauses were negatively associated with score. In these tasks, very little editing, revision, or even proofreading took place. Many students stopped writing long before the task time limit was reached, and nearly all students produced text sequentially, appending new

text to the end of the existing document, and appeared to pause primarily when needed to plan the next text segment. They did very little proofreading, editing, or revision.

The literature has suggested that novice writers typically follow a knowledge-telling strategy (Bereiter & Scardamalia, 1987), in which idea generation tends to be reduced to retrieval of relevant prior knowledge and the translation and transcription processes are only minimally regulated by executive control processes that monitor the output and help students apply other, more effective writing strategies. The pattern of behavior we observed in our earlier studies is consistent with the hypothesis that writers completing these tasks follow a knowledge-telling strategy in which they retrieve relevant information from long-term memory, translate that information into text, and stop writing when they are unable to retrieve any more information relevant to the topic about which they are writing.

If this is true, the direct writing assessments arguably provided evidence about how effectively students could retrieve relevant information (idea generation) and how efficiently they could express that information in words and transcribe it digitally, but they provided almost no information about how well students can monitor and evaluate their own performance or adopt appropriate strategies to break up the writing work into manageable subtasks. In the context of a direct writing assessment, a simple knowledge-telling approach to writing may in fact be safest and most efficient, as it does not require the writer to engage in complex and time-consuming planning, evaluation, and revision processes that might interfere with completing the writing task in a limited time frame. And yet, the literature has indicated that planning, evaluation, and revision skills are among the critical skills students need to master to engage effectively in more complex writing practices. We would contend that an effective writing assessment should attempt to measure these kinds of information in addition to measuring fluency and accuracy of text production.

In the current data set, collected under classroom conditions, students spent much more time on the writing task than they did in the quite similar direct writing assessments we administered earlier. Moreover, they appear to have separated at least two parts of the writing process (planning and editing/proofreading) from drafting when the task was structured to encourage them to do so.[2] Because the population largely consisted of students from underserved populations (students from backgrounds with low socioeconomic status and from largely African American and Hispanic backgrounds with a fairly high proportion of English learners), the

results of our study suggest that we may be able to obtain information about very different aspects of writing competency even for students with potentially lower levels of achievement by requiring them to undertake multiple writing subtasks that emphasize different components of the writing process.

It is easiest to see how these goals could be accomplished under formative conditions. In the classroom, it is feasible to assign rich writing tasks that provide multiple opportunities for revision and editing. Digital writing tools could be used to capture product and process information and assess student writing in the background. The inferences suggested by this kind of background assessment system could be used to support effective formative assessment practices that would increase student engagement and motivation and the probability of seeing the more advanced forms of writing behavior even from relatively weak students.

In a summative assessment, alternatively, it may be relatively hard to elicit all of the writing behaviors we wish to measure, for all the reasons we have just surveyed. Various strategies are worth considering, such as assigning explicit peer review, revision, or editing tasks either as follow-ups to a direct writing assessment or as additional tasks. Writing process logs provide valuable new tools for analyzing students' writing processes, but only if the tasks they undertake elicit the evidence we wish to measure.

References

Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing*. New York, NY: Springer. https://doi.org/10.1007/978-94-010-0804-4

Alamargot, D., & Fayol, M. (2009). Modeling the development of written composition. In R. Beard, D. Myhill, M. Nystrand, & J. Riley (Eds.), *Handbook of writing development* (pp. 23–47). London, England: Sage. https://doi.org/10.4135/9780857021069.n3

Allen, L. K., Jacovina, M. E., Dascalu, M., Roscoe, R. D., Kent, K. M., Likens, A. D., & McNamara, D. S. (2016). *{ENTERING}ing the time series {SPACE}: Uncovering the writing process through keystroke analysis*. Paper presented at the 9th International Conference on Educational Data Mining, June 29-July 2, Raleigh, NC.

Allen, L. K., Mills, C., Jacovina, M. E., Crossley, S. A., D'Mello, S., & McNamara, D. S. (2016). *Investigating boredom and engagement during writing using multiple sources of information: The essay, the writer, and keystrokes*. Paper presented at LAK'16, Edinburgh, Scotland. https://doi.org/10.1145/2883851.2883939

Allen, L. K., Perret, C., & McNamara, D. S. (2016). *Linguistic signatures of cognitive processes during writing*. Paper presented at Cogsci 2016, the annual meeting of the Cognitive Science Society, Austin, TX.

Andersson, B., Dahl, J., Holmqvist, K., Holsanova, J., Johansson, V., Strömqvist, S., . . . Wengelin, Å. (2006). Combining keystroke logging with eye tracking. In L. van Waes, M. Leijten, & C. Neuwirth (Eds.), *Writing and digital media* (pp. 166–172). Oxford, England: Elsevier.

Baaijen, V. M., Galbraith, D., & de Glopper, K. (2012). Keystroke analysis: Reflections on procedures and measures. *Written Communication, 29*, 246–277. https://doi.org/10.1177/0741088312451108

Beaufort, A. (2008). Writing in the professions. In C. Bazerman (Ed.), *Handbook of research on writing: History, society, school, individual, text* (pp. 221–236). New York, NY: Lawrence Erlbaum.

Becker, A. (2006). A review of writing model research based on cognitive processes. In A. Horning & A. Becker (Eds.), *Revision: History, theory, and practice* (pp. 25–48). West Lafayette, IN: Parlor Press.

Beers, S., Quinlan, T., & Harbaugh, A. G. (2010). Adolescent students' reading during writing behaviors and relationships with text quality: An eyetracking study. *Reading and Writing, 23*, 743–775. https://doi.org/10.1007/s11145-009-9193-7

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

Berninger, V. W. (1999). Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes. *Learning Disability Quarterly, 22*, 99–112. https://doi.org/10.2307/1511269

Berninger, V. W. (2009). Multiple processes that matter in writing instruction and assessment. In G. Troia (Ed.), *Instruction and assessment for struggling writers: Evidence-based practices* (pp. 15–50). New York, NY: Guilford Presss.

Bixler, R., & D'Mello, S. (2013). *Detecting boredom and engagement during writing with keystroke analysis, task appraisals, and stable traits*. Paper presented at the International Conference on Intelligent User Interfaces, ACM, March 19–22, Santa Monica, CA. https://doi.org/10.1145/2449396.2449426

Caporossi, G., & Leblay, C. (2011). *Online writing data representation: A graph theory approach*. Paper presented at IDA 2011: Advances in Intelligent Data Analysis X, October 29–31, 2011, Porto, Portugal. https://doi.org/10.1007/978-3-642-24800-9_10

Carl, M. (2012, May 23–25). *Translog-II: A program for recording user activity data for empirical reading and writing research.* Paper presented at the 8th International Conference on Language Resources and Evaluation, Istanbul, Turkey.

Carl, M., Kay, M., & Jensen, K. T. H. (2010). *Long-distance revisions in drafting and post-editing.* Paper presented at the 11th International Conference on Intelligent Text Processing and Computational Linguistics, March 21–27, 2010, Iasu, Romania.

Chenu, F., Pellegrino, F., Jisa, H., & Fayol, M. (2014). Interword and intraword pause threshold in writing. *Frontiers in Psychology, 5*, 182. https://doi.org/10.3389/fpsyg.2014.00182

Cho, Y. (2003). Assessing writing: Are we bound by only one method? *Assessing Writing, 8*, 165–191. https://doi.org/10.1016/S1075-2935(03)00018-7

Chukharev-Hudilainen, E. (2014). Pauses in spontaneous written communication: A keystroke logging study. *Journal of Writing Research, 6*(1), 61–84. https://doi.org/10.17239/jowr-2014.06.01.3

Connelly, V., Dockrell, J. E., Walter, K., & Critten, S. (2012). Predicting the quality of composition and written language bursts from oral language, spelling, and handwriting skills in children with and without specific language impairment. *Written Communication, 28,* 278–302. https://doi.org/10.1177/0741088312451109

Deane, P. (2011). *Writing assessment and cognition* (Research Report No. RR-11-14). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2011.tb02250.x

Deane, P. (2013). Covering the construct: An approach to automated essay scoring motivated by a socio-cognitive framework for defining literacy skills. In M. Shermis & J. Burstein (Eds.), *Handbook of automated essay evaluation: Current applications and new directions* (pp. 298–312). New York, NY: Routledge.

Deane, P. (2014). *Using writing process and product features to assess writing quality and explore how those features relate to other literacy tasks* (Research Report No. RR-14-03). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12002

Deane, P., Odendahl, N., Quinlan, T., Fowles, M., Welsh, C., & Bivens-Tatum, J. (2008). *Cognitive models of writing: Writing proficiency as a complex integrated skill* (Research Report No. RR-08-55). Princeton, NJ: Educational Testing Service. Retrieved from https://www.ets.org/Media/Research/pdf/RR-08-55.pdf

Deane, P., & Zhang, M. (2015). *Exploring the feasibility of using writing process features to assess text production skills* (ETS RR-15–26). Retrieved from Princeton, NJ: http://onlinelibrary.wiley.com/doi/10.1002/ets2.12071

Dimitrova, B. E. (2006). Segmentation of the writing process in translation: Experts versus novices. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging and writing* (pp. 189–202). Oxford, England: Elsevier.

DiStefano, P., & Killion, J. (1984). Assessing writing skills through a process approach. *English Education, 16,* 203–207.

Englert, C. S., Mariage, T. V., & Dunsmore, K. (2006). Tenets of sociocultural theory in writing instruction research. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 208–221). New York, NY: Guilford Press.

Fayol, M. (1999). From on-line management problems to strategies in written composition. In M. Torrance & G. Jeffery (Eds.), *The cognitive demands of writing: Processing capacity and*

working memory effects in text production (pp. 15–23). Amsterdam, Netherlands: Amsterdam University Press.

Fayol, M. (2012). Cognitive processes of children and adults in translating thought into written language in real time: Perspectives from 30 years of programmatic cognitive psychology and linguistics research. In V. W. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 27–60). New York, NY: Psychology Press.

Flinn, J. Z. (1987). Case studies of revision aided by keystroke recording and replaying software. *Computers and Composition*, *5*(1), 31–44. https://doi.org/10.1016/S8755-4615(87)80013-0

Flower, L., & Hayes, J. R. (1980). A cognitive process theory of writing. *College Composition and Communication, 32*, 365–387. https://doi.org/10.2307/356600

Grabowski, J. (2008). The internal structure of university students' keyboard skills. *Journal of Writing Research, 1*(1), 27–52. https://doi.org/10.17239/jowr-2008.01.01.2

Graham, S., & Harris, K. R. (2012). The role of strategies, knowledge, will, and skills in a 30-year program of writing research (with homage to Hayes, Fayol, and Boscodo). In V. W. Berninger (Ed.), *Past, present and future contributions of cognitive writing research to cognitive psychology* (pp. 177–196). New York, NY: Psychology Press.

Harris, K. R., & Graham, S. (2009). Self-regulated strategy development in writing: Premises, evolution, and the future. *British Journal of Educational Psychology Monograph Series, II, 6*, 113–135.

Harris, K. R., Graham, S., & Mason, L. H. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. *American Educational Research Journal, 43*, 295–340. https://doi.org/10.3102/00028312043002295

Hayes, J. R. (2011). Kinds of knowledge-telling: Modeling early writing development. *Journal of Writing Research, 3*(2), 73–92. https://doi.org/10.17239/jowr-2011.03.02.1

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

Hayes, J. R., & Chenoweth, N. A. (2006). Is working memory involved in the transcribing and editing of texts? *Written Communication, 23*, 135–149. https://doi.org/10.1177/0741088306286283

Hayes, J. R., & Chenoweth, N. A. (2012). Working memory in an editing task. *Written Communication, 24*, 283–294. https://doi.org/10.1177/0741088307304826

Hayes, J. R., & Flower, L. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–29). Hillsdale, NJ: Lawrence Erlbaum.

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

Heilman, A., & Neumann, S. (2016). *Dynamic assessment of keystroke logged data for the detection of complexity in translation and monolingual text products*. Paper presented at the Workshop on Computational Linguistics for Text Complexity, Osaka, Japan.

Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. Ann Arbor, MI: University of Michigan Press.

Jakobsen, A. L. (2006a). Research methods in translation—Translog. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging and writing* (pp. 95–101). Oxford, England: Elsevier.

Jakobsen, A. L. (2006b). Translog: Research methods in translation. In K. P. H. Sullivan & E. Lindren (Eds.), *Computer keystroke logging and writing: Methods and applications* (pp. 95–105). Oxford, England: Elsevier.

Johansson, R., Johansson, V., & Wengelin, Å. (2012). Text production in handwriting versus computer typing. In *Learning to write effectively: Current trends in European research* (pp. 375–377). Bingley, UK: Emerald Group.

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

Kellogg, R. T. (1987). Writing performance: Effects of cognitive strategies. *Written Communication, 4*, 269–298. https://doi.org/10.1177/0741088387004003003

Kellogg, R. T. (1996a). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–71). Hillsdale, NJ: Lawrence Erlbaum.

Kellogg, R. T. (1996b). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*, 355–365. https://doi.org/10.1037/0278-7393.14.2.355

Kellogg, R. T. (2001). Competition for working memory among writing processes. *American Journal of Psychology, 114*, 175–191. https://doi.org/10.2307/1423513

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

Kellogg, R. T., Turner, C. E., Whiteford, A. P., & Mertens, A. (2016). The role of working memory in planning and generating written sentences. *Journal of Writing Research, 7*, 397–416. https://doi.org/10.17239/jowr-2016.07.03.04

Leblay, C., & Caporossi, G. (2015). A graph theory approach to online writing data visualization. In G. Cislaru (Ed.), *Writing(s) at the crossroads: The process-product interface* (pp. 171–184). Amsterdam, Netherlands: John Benjamins. https://doi.org/10.1075/z.194.09leb

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

Limpo, T., & Alves, R. A. (2013). Teaching planning or sentence-combining strategies: Effective SRSD interventions at different levels of written composition. *Contemporary Educational Psychology, 38*, 328–341. https://doi.org/10.1016/j.cedpsych.2013.07.004

Limpo, T., & Alves, R. A. (2014). Children's High-Level Writing Skills: Development of planning and revising and their contribution to writing quality. *British Journal of Educational Psychology, 84*, 177–193. https://doi.org/10.1111/bjep.12020

Lindgren, E., Stevenson, M., & Sullivan, K. P. H. (2008). Supporting the reflective language learner with computer keystroke logging. In F. Zhang & B. Barber (Eds.), *Handbook of research on computer-enhanced language acquisition* (pp. 189–204). Hershey, PA: Information Science Reference. https://doi.org/10.4018/978-1-59904-895-6.ch011

Macken, L., Hoste, V., Leijten, M., & Van Waes, L. (2012). *From keystrokes to annotated process data: Enriching the output of InputLog with linguistic information*. Paper presented at the 8th International Conference on Language Resources and Evaluation, Istanbul, Turkey.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review, 8*, 299–325. https://doi.org/10.1007/BF01464076

McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist, 35*, 13–23. https://doi.org/10.1207/S15326985EP3501_3

McCutchen, D. (2011). From novice to expert: Implications of language skills and writing-relevant knowledge for memory during the development of writing skill. *Journal of Writing Research, 3*(1), 51–68. https://doi.org/10.17239/jowr-2011.03.01.3

Medimorec, S., & Risko, E. F. (2017). Pauses in written composition: On the importance of where writers pause. *Reading and Writing: An Interdisciplinary Journal, 30*, 1267–1285. https://doi.org/10.1007/s11145-017-9723-7

Mikulski, A., & Elola, I. (2011). Spanish heritage language learners' allocation of time to writing processes in English and Spanish. *Hispania, 94*, 715–733.

Olive, T., Alves, R. A., & Castro, S. L. (2009). Cognitive processes in writing during pause and execution periods. *European Journal of Cognitive Psychology, 21*, 758–785. https://doi.org/10.1080/09541440802079850

Olive, T., & Kellogg, R. T. (2002). Concurrent activation of high- and low-level production processes in written composition. *Memory and Cognition, 30*, 594–600. https://doi.org/10.3758/BF03194960

Perrin, D., & Wildi, M. (2008). Cumulated deviation of a linear trend: Statistical modeling of writing processes. In C. Bazerman, R. Krut, K. Lunsford, S. McLeod, S. Null, P. Rogers, & A. Stansell (Eds.), *Traditions of writing research* (pp. 378–393). New York, NY: Routledge.

Prunty, M. M., Barnett, A. L., Wilmut, K., & Plumb, M. S. (2014). An examination of writing pauses in the handwriting of children with developmental coordination disorder. *Developmental Disabilities, 35*, 2894–2905. https://doi.org/10.1016/j.ridd.2014.07.033

Quinlan, T., Loncke, M., & Leijten, M. (2012). Coordinating the cognitive processes of writing: The role of the monitor. *Written Communication, 29*, 345–368. https://doi.org/10.1177/0741088312451112

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

Roscoe, R. D., Snow, E. L., Allen, L. K., & McNamara, D. S. (2015). Automated detection of essay revising patterns: Applications for intelligence feedback in a writing tutor. *Technology, Instruction, Cognition, and Learning, 10*, 59–79.

Rosenqvist, S. (2015). *Developing pause thresholds for keystroke logging analysis* (Unpublished master's thesis). Umea University, Umea, Sweden.

Sabbaghan, S. (2013). How noticing is affected by replay of writing process during stimulated recall. *Procedia Social and Behavioral Sciences, 83*, 629–633. https://doi.org/10.1016/j.sbspro.2013.06.119

Schilperoord, J., & Sanders, T. (1999). How hierarchical text structure affects retrieval processes: Implications of pause and text analysis. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write: Conceptual processes in text production* (pp. 13–30). Amsterdam, Netherlands: Amsterdam University Press.

Shaffer, L. H. (1975). Control processes in typing. *Quarterly Journal of Experimental Psychology, 27*, 419–432. https://doi.org/10.1080/14640747508400502

Shaffer, L. H. (1978). Timing in the motor programming of typing. *Quarterly Journal of Experimental Psychology, 30*, 333–345. https://doi.org/10.1080/14640747808400680

Shaffer, L. H., & Hardwick, J. (1968). Typing performance as a function of text. *Quarterly Journal of Experimental Psychology, 20*, 360–369. https://doi.org/10.1080/14640746808400175

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review, 78*(1), 40–59. https://doi.org/10.17763/haer.78.1.v62444321p602101

Spelman Miller, K. (2002). Units of production in writing: Evidence of topic "framing" in on-line writing research. *Reading Working Papers in Linguistics, 6*, 244–272.

Spelman Miller, K., & Lindgren, E. (2011). The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging. *TESOL Quarterly*, *42*, 433–454. https://doi.org/10.1002/j.1545-7249.2008.tb00140.x

Spelman Miller, K., & Sullivan, K. P. H. (2006). Keystroke logging: An introduction. In K. P. H. Sullivan & E. Lindren (Eds.), *Computer keystroke logging: Methods and applications* (pp. 1–9). Oxford, England: Elsevier.

Stevenson, M., Schoonen, R., & de Glopper, K. (2006). Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL. *Journal of Second Language Writing, 15*, 201–233. https://doi.org/10.1016/j.jslw.2006.06.002

Sullivan, K., Kollberg, P., & Palsson, E. (1997). L2 writing: A pilot investigation using key-stroke logging. In L. Diaz & C. Perez (Eds.), *Views on the acquisition and use of a second language* (pp. 553–566). Barcelona, Spain: Universitat Pompeu Fabra.

Sumner, E., Connelly, V., & Barnett, A. L. (2013). Children with dyslexia are slow writers because they pause more often and not because they are slow at handwriting execution. *Reading and Writing: An Interdisciplinary Journal, 26*, 991–1008. https://doi.org/10.1007/s11145-012-9403-6

Thorson, H. (2000). Using the computer to compare foreign- and native-language writing processes: A statistical and case study approach. *Modern Language Journal*, *84*, 55–70. https://doi.org/10.1111/0026-7902.00059

Torkildsen, J. V. K., Morken, F., Helland, W. A., & Helland, T. (2015). The dynamics of narrative writing production in primary grade children: Writing process factors predict story quality. *Reading and Writing: An Interdisciplinary Journal*, *29*, 529–554. https://doi.org/10.1007/s11145-015-9618-4

Torrance, M., & Jeffery, G. (1999). Writing processes and cognitive demands. In M. Torrance & G. Jeffery (Eds.), *The cognitive demands of writing: Processing capacity and working memory in text production* (pp. 1–11). Amsterdam, Netherlands: Amsterdam University Press.

van Orden, G. C., Kloos, H., & Wallot, S. (2011). Living in the pink: Intentionality, well-being, and complexity. In C. A. Hooker (Ed.), *Philosophy of complex systems: Handbook of the philosophy of science* (pp. 639–683). Amsterdam, Netherlands: Elsevier. https://doi.org/10.1016/B978-0-444-52076-0.50022-5

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

Van Waes, L., & Schellens, P. J. (2003). Writing profiles: The effect of the writing mode on pausing and revision patterns of experienced writers. *Journal of Pragmatics, 35*, 829–853. https://doi.org/10.1016/S0378-2166(02)00121-2

Wallot, S., & Grabowski, J. (2013). Typewriting dynamics: What distinguishes simple from complex writing tasks? *Ecological Psychology, 25*, 1–14. https://doi.org/10.1080/10407413.2013.810512

Wallot, S., O'Brien, B. A., & Van Orden, G. (2012). Fractal and recurrence analysis of psycholinguistic data. In G. Liben, G. Jarema, & C. Westbury (Eds.), *Methodological and analytical frontiers in lexical research* (pp. 395–430). Amsterdam, Netherlands: Bejamins. https://doi.org/10.1075/bct.47.18wal

Wengelin, Å. (2006). Examining pauses in writing: Theory, methods and empirical data. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging and writing* (pp. 107–130). Oxford, England: Elsevier.

Wengelin, Å., Torrance, M., Holmqvist, K., Simpson, S., Galbraith, D., Johansson, V., & Johansson, R. (2009). Combined eyetracking and keystroke-logging methods for studying cognitive processes in text production. *Behavior Research Methods, Instruments, and Computers, 41*, 337–351. https://doi.org/10.3758/BRM.41.2.337

Wiggins, G. (1994). The constant danger of sacrificing validity to reliability: Making writing assessment serve writers. *Assessing Writing, 1*, 129–139. https://doi.org/10.1016/1075-2935(94)90008-6

Zhang, M., & Deane, P. (2015). *Process features in writing: Internal structure and incremental value over product features* (Research Report No. RR-15-27). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12075

Appendix A. Student Packet for the Junk Food Formative Task

Name: _____

Student ID#: _____

Teacher: _____

"Should Junk Food Be Sold in Your School?"

Scenario: Everyone is discussing whether or not junk food (unhealthful food and drinks) should be sold at school. You and your classmates are trying to learn more and make up your own minds. In this project, you will research the issue, explore arguments on both sides of the issue, and write an essay for your school newspaper to explain your point of view.

© Vivian Cardentey/iStockphoto # 10891410

School officials are trying to decide whether the middle schools in your area should stop selling unhealthful junk food and drinks that contain large amounts of sugars and/or fats. At your school, some people think that selling junk food to students should be allowed, and some people think that it should not be allowed.

Writing Assignment: Now the editor of the school newspaper has asked students to write essays about their own views.

You may first need to consider questions such as:

- What are the arguments for and against selling junk food in your school?
- What is your view, and how will you present it?
- Should students make the decision about whether or not junk food is sold in their school?

Table of Contents and Calendar

The calendar below will help you understand the sequence of activities and the work you miss if you're absent.

Day	Date	Activities
1	Sept. 21	– Introduction to the Argumentation Unit – Free Write – Pro/Con Chart
2	Sept. 22	– Argumentation Vocabulary – Sorting Arguments
3	Sept. 23	– Writing a Summary – "Yes" Article and Summarization
4	Sept. 24	– Read "Might Not Solve" Article, Write a Summary – Read "No" Article, Write a Summary
5	Sept. 25	– Peer Review Summaries – Evaluating Evidence
6	Sept. 28	– Essay Writing Guidelines – Argumentation Rubric – Creating a Writing Plan (Graphic Organizer)
7	Sept. 29	– Sentence Starters for Peer Review – Essay Planning and Review
8	Sept. 30	– Submit an Essay Draft
9	Oct. 1	– Peer Review, Self– Assess
10	Oct. 2	– Submit Final Draft

Day 1: Free Write

In the space below, write down what **you** think about junk food being sold in schools.

- What is junk food?
- Do you think it should be sold in schools?
- Why or why not?

What your partner thinks about junk food:

Day 1: Pro and Con

Directions: Use the chart below to take notes from your class discussion.

Topic: Should junk food be sold in schools?

Pro (reasons ***for*** it)	Con (reasons ***against*** it)

Day 2: Argumentation Vocabulary

Directions: Write the definition below for each vocabulary word. Be sure to write the definition in a way that makes sense to you and helps you understand what it means.

Ban:

Allow:

Summary:

Source:

Reason:

Warrant:

Evidence:

Claim:

Counterclaim:

Rebuttal:

Day 2: Analyze Arguments

Consider Arguments for and Against the Issue

Are there good reasons to *ban* the sale of junk food at your school? Are there good reasons to *allow* it?

Your job is to help organize the list below of some reasons people have relating to this issue. Divide the statements into two groups: reasons to ban the sale of junk food and reasons to allow it.

Directions: Read each statement and decide whether it gives a reason to ban or to allow the sale of junk food at school. Put a check in the appropriate box.

Some Reasons People Give to Ban or Allow the Sale of Junk Food at School	Ban	Allow
1. Candy bars, sodas, and other junk foods give students the quick energy they sometimes need.		
2. Kids can get a balanced diet, even if they eat some junk food.		
3. Parents are very concerned about the impact of junk food on their children's health.		
4. If kids can't buy junk food at school, they'll just get it somewhere else.		
5. By selling junk food, schools can earn a lot of money for important school activities.		
6. Beginning in middle school, kids have to learn to exercise self-control.		
7. The sale of junk food undermines a school's responsibility to provide students with nutritious food.		
8. The more junk food kids eat, the more they become addicted to it.		
9. Middle school students have the right to make up their own minds about what to eat.		
10. Too much sugar makes kids jittery, and then they can't focus on their studies.		
11. Many kids argue that it's okay to eat junk food, but very few doctors agree.		

Day 3: Consider Arguments FOR Selling Junk Food in School

Some people think that all junk food is harmful and should not be sold at school. Others think that your school should sell junk food. Before developing your own position, you will need to understand the arguments on both sides*: ban* the sale of junk food at your school or a*llow* it.

To research the issue, you and your classmates will read several articles and write a brief *summary* of each one.

Before beginning your research, read the summary guidelines below. The guidelines will help you to write your summaries.

Summary Guidelines

1. State the central idea or position that the source discusses. (What is its main point?)
2. Report only the most important supporting ideas. (Leave out minor ideas and details.)
3. Report only what the author wrote. (Summaries do not include your own opinions on the subject.)
4. Be accurate. (Be careful not to distort information from the source.)
5. Use mostly your own words. (But if you do quote from the source, use quotation marks.)

distort = twist, misrepresent
source = the article, interview, or other text that you need to summarize

Directions: Read the article below. Then use the article to answer the questions on the following pages that are related to the Summary Guidelines on the previous page. Later, you will use the Guidelines to write your own summaries.

Should Schools Sell Junk Food? (YES!)

By M. Carter, Principal of a Large High School in Florida

Junk food should be sold in schools—along with other food.

It would be great if all of our students ate the nutritious meals served in the cafeteria. Our cafeteria offers a wide variety of meals, but some students are allergic to tomatoes, certain spices, or other ingredients. Others cannot deal with the cafeteria's noise and long lines and find it easier to simply grab a snack from the vending machine during our short lunch periods. Also, some students just do not recognize cafeteria meals as food. Even our best meals have no nutritional value when they end up in the garbage can.

Another reason for selling junk food is that it provides quick energy. It can take the place of a meal or help fill out an inadequate meal. Without snacks, students may fall asleep or have trouble paying attention in class.

On the economic side, junk-food sales and the recycling of aluminum cans provide a steady income for school organizations. A school in the Washington, DC, area earned over $98,000 (about a quarter of the school's budget) from vending machines that sell junk food.

It is true that we should teach and model good eating habits, and it is true that we have health problems brought on by poor nutrition. But many schools now offer milk, fruit, and other nutritious alternatives in their vending machines, along with the junk food. Students like and need freedom of choice—including their choice of foods.

Source: Mark Carter, "Perspectives of a Principal" 2008

Directions: Read the following discussion between two pretend students. Then answer the questions that follow.

BRIAN: Let's start with Summary Guideline 1: State the central idea or position that the source discusses. (What is its main point?)

EVA: How about this statement for the central idea? *Some parts of the principal's argument make sense, but it's not good for kids to eat junk food instead of a nutritional lunch.*

BRIAN: Eva, you can't use that as the central idea.

Why shouldn't Eva use that sentence as the central idea?

It's just a supporting detail.

It gives her own opinion about the topic.

It distorts information in the article.

It just copies words exactly as they appear in the article.

Below are statements from other students. Circle which one Brian and Eva should use at the central idea of the "Yes!" article.

The main point of the article is that school principals and students do not always agree on the benefits of eating nutritional food.

The author, a school principal, actually admits that his school's cafeteria does not serve food that appeals to most students.

The article was written by the principal of a school in Florida; it focuses on the topic of junk food.

In this article, a school principal explains why he thinks schools should sell junk food.

Directions: Other students wrote summaries that do not meet the Summary Guidelines. Read Summary #1. Then answer the questions that follow.

Summary #1

(1) The main argument in the article, "Should Schools Sell Junk Food? (YES!)," is that it's a good idea for schools to sell junk food. (2) The author gives several reasons to support this position. (3) One reason is that students sometimes can't eat cafeteria food or don't want to go to the cafeteria or just need some quick energy, so junk food should be available in vending machines, along with other snacks. (4) Another reason is that the fastest and easiest way for any school to get money is to sell junk food. (5) Although the author acknowledges that good eating habits are important, he claims that students should nevertheless be free to choose what they eat.

What is the main problem with Summary #1?

It includes minor details.

It gives the student's own opinion on the subject.

It distorts information in the article.

It copies too much from the article.

In which of the following sentences from the student's summary does the problem you identified appear?

(1) The main argument in the article, "Should Schools Sell Junk Food? (YES!)," is that it's a good idea for schools to sell junk food. The author gives several reasons to support this position.

(3) One reason is that students sometimes can't eat cafeteria food or don't want to go to the cafeteria or just need some quick energy, so junk food should be available in vending machines, along with other snacks

(4) Another reason is that the fastest and easiest way for any school to get money is to sell junk food

(5) Although the author acknowledges that good eating habits are important, he claims that students should nevertheless be free to choose what they eat.

Directions: Other students wrote summaries that do not meet the Summary Guidelines. Read Summary #2. Then answer the questions below.

Summary #2

(1) In this article, the principal of a Florida high school explains why both students and schools benefit from the sale of junk food. (2) Students benefit because if they are in a hurry or are allergic to food in the cafeteria, they can get junk food and other snacks from a vending machine. (3) Schools benefit because the sale of junk food can bring in money for school organizations. (4) Fruit is an example of a nutritious food that might be sold in a vending machine.

What is the main problem with Summary #2?

It includes minor details.

It gives the student's own opinion on the subject.

It distorts information in the article.

It copies too much from the article.

In which of the following sentences from the student's summary does the problem you identified appear?

(1) In this article, the principal of a Florida high school explains why both students and schools benefit from the sale of junk food.

(2) Students benefit because if they are in a hurry or are allergic to food in the cafeteria, they can get junk food and other snacks from a vending machine.

(3) Schools benefit because the sale of junk food can bring in money for school organizations.

(4) Fruit is an example of a nutritious food that might be sold in a vending machine.

Directions: Other students wrote summaries that do not meet the Summary Guidelines. Read Summary #3. Then answer the questions below.

Summary #3

(1) The article "Should Schools Sell Junk Food? (YES!)" is basically a principal's argument in favor of selling junk food at school. (2) According to the author, Mark Carter, one reason is that students with allergies might need to buy junk food and other snacks from a vending machine. (3) However, I know lots of students with dietary restrictions, and most of them prefer to pack a lunch with their special foods from home. (4) Mr. Carter presents other reasons, too, including economic advantages and the need for

What is the main problem with Summary #3?

It includes minor details.

It gives the student's own opinion on the subject.

It distorts information in the article.

It copies too much from the article.

In which of the following sentences from the student's summary does the problem you identified appear?

(1) The article "Should Schools Sell Junk Food? (YES!)" is basically a principal's argument in favor of selling junk food at school.

(2) According to the author, Mark Carter, one reason is that students with allergies might need to buy junk food and other snacks from a vending machine.

(3) However, I know lots of students with dietary restrictions, and most of them prefer to pack a lunch with their special foods from home.

(4) Mr. Carter presents other reasons, too, including economic advantages and the need for students to be free to choose what they eat.

Directions: First, read the draft summary below. Pay special attention to the bolded and underlined words. Is each word the best and most accurate one to use, given the context, or should another word be substituted?

After you have read the summary, answer the questions that follow, which ask you to decide whether Eva should use each highlighted word or change it to one of the other choices.

Draft Summary

In the article "Should Schools Sell Junk Food? (YES!)," a school principal from Florida explains why he thinks schools should sell junk food. He makes an economic argument, **and** schools can earn money by selling junk food and recycling soda cans. He also argues that vending machines offer an alternative to cafeteria food, **where** some students cannot eat cafeteria food for health reasons and others cannot tolerate cafeteria conditions. He concludes his argument with the **claim** that students should have the freedom to decide what foods to eat.

He makes an economic argument, _____ schools can earn money by selling junk food and recycling soda cans.

and
and yet
pointing out that
if we know that

He also argues that vending machines offer an alternative to cafeteria food, _____ some students cannot eat cafeteria food for health reasons and others cannot tolerate cafeteria conditions.

where
although
because
and, as a result,

He concludes his argument with the _____ that students should have the freedom to decide what foods to eat.

claim
prediction
example
evidence

Day 4: Read and Summarize another Article

Directions: Read "Why Banning Junk Food Sales in School Might Not Solve the Problem." Then use the space below the article to write a brief summary (2–3 sentences) of the article's main points. Make sure to refer to the Summary Guidelines to help you write your summary. When you finish, continue on to the "Should Schools Allow the Sale of Junk Food? (No!)" article, and write a summary for it in the space provided.

Why Banning Junk Food Sales in School Might Not Solve the Problem

Some schools are learning that banning junk food from school doesn't always keep students from eating junk food.

That is the experience in Oakland, California, which **banished** soda, candy, caffeinated drinks, and other products from schools in the 52,000-student district. The school district's food-service manager, Amy Lins, says students are now eating things like soy-based burgers, salads, and grilled chicken and are now drinking fruit juices.

''But more kids may be sneaking off to get their junk food off campus,'' Ms. Lins said.

At a school in Fremont, California, students run a thriving business selling nutritious food made by local merchants. But some students go to the local fast-food restaurant for fries and the corner market for cookies, sodas, and chips. About an hour after the students return, teachers say, they will crash in a daze as the sugar high wears off.

''You can chart it,'' Michael Moore, a deputy principal, said. ''The students say these foods keep them awake. But it ends up putting them to sleep in the afternoon.''

Adapted from Timothy Egan, "In Bid to Improve Nutrition, Schools Expel Soda and Chips," *The New York Times,* May 20, 2002.

Banished: got rid of

Summary

_____

_____

_____

Should Schools Allow the Sale of Junk Food? (NO!)

By J. Warner, School Nurse in St. Louis, Missouri

Why do so many Americans eat junk food? According to a national survey conducted in 2007, one of the main reasons is that junk food is heavily advertised. It's everywhere, even at school. We teach students good nutrition and then sell them junk food--what a double message! When students refuse to eat cafeteria food, it's because they are used to junk food and are perhaps even **addicted** to sugar. Research shows that, given a choice, most students prefer to eat junk food even though they know that is terrible for their teeth, bones, and weight.

Dentists are seeing massive tooth decay in the mouths of young adults, which is caused in part by drinking soda rather than milk. Also, when students drink soda instead of milk, their bones do not develop properly. After about the age of 18, our bodies stop building bone mass. If kids haven't developed strong bones by then, they will face serious problems when they are older: their fragile bones will be easily broken.

Some people argue that junk food sales support school activities. Well, student health is more important than making money. Had any dental work done lately? Costly! Bone problems in old age are even more expensive.

Do teens need a sugar pick-me-up? Not if they eat according to the Food Pyramid. Let's guide them to life-long healthy eating habits!

Source: Warner, Jane. "School Health Journal," April, 2002.]

Addicted: needing, dependent on

Summary

Day 5: Consider Evidence From the Articles

Now you and your class will evaluate arguments by judging whether claims from the letters to the editor can be supported using information from articles that the class has gathered. This will prepare you to use claims and evidence in your own essay.

Directions: For each question below, decide whether the evidence:

- supports the claim
- weakens the claim
- neither supports nor weakens the claim

CLAIM: Because sugary and fatty foods are bad for our health, schools should *ban* the sale of junk food.
EVIDENCE: "Dentists are seeing massive tooth decay in the mouths of young adults, which is caused in part by drinking soda rather than milk."
The evidence

A. supports the claim
B. weakens the claim
C. neither supports nor weakens the claim

CLAIM: Because some students don't have extra money to spend on snacks, schools should *ban* the sale of junk food.
EVIDENCE: According to the results of a 2007 survey, many people eat junk food because "it is heavily advertised."
The evidence

A. supports the claim
B. weakens the claim
C. neither supports nor weakens the claim

CLAIM: Because there are other ways to raise money for school projects besides selling candy and soda, schools should *ban* the sale of junk food.

EVIDENCE: "At a school in Fremont, California, students run a thriving business selling nutritious food made by local merchants."

The evidence:

A. supports the claim
B. weakens the claim
C. neither supports nor weakens the claim

CLAIM: Because most teens are mature enough to make good nutritional choices, schools should *allow* the sale of junk food.

EVIDENCE: "Research shows that, given a choice, most students prefer to eat junk food, even though it is bad for their teeth, bones, and weight."

The evidence

A. supports the claim
B. weakens the claim
C. neither supports nor weakens the claim

CLAIM: Because schools need more money for computers, band uniforms, and special projects, we should *allow* the sale of junk food.

EVIDENCE: "A school in the Washington, D.C., area earned over $98,000 (about a quarter of the school's budget) from vending machines that sell junk food."

The evidence

A. supports the claim
B. weakens the claim
C. neither supports nor weakens the claim

CLAIM: Because kids who can't buy junk food at school will just get it somewhere else, schools should *allow* the sale of junk food.

EVIDENCE: An Oakland, California, middle school that banned the sale of junk food reported that, as a result, "more kids started sneaking off to buy junk food off campus."

The evidence

A. supports the claim
B. weakens the claim
C. neither supports nor weakens the claim

Days 6–10: Present Your View in an Essay

Graphic Organizer

Now you are ready to explain your own position on the issue.

Directions: Write an essay for your school newspaper. Explain your view on the question: Should your school sell junk food to students?

You can, but need not, use information from the articles. You can also use your own arguments and examples.

In developing your essay, use logical reasons and/or relevant examples to support your position and to argue against the opposite position. If you use information from the documents, be sure to use it accurately.

You should use a planning tool to help you organize your essay before you start writing.

Now you are ready to explain what <u>you</u> think about the issue.

Many people use a planning tool or graphic organizer to organize information (main points, supporting details, examples, etc.) for essays, lectures, articles, or other types of text. In this task, you use the **Idea Web** in CRITERION to organize and plan for your essay.

Writer's Checklist: Argumentative Essay

Make sure that your essay has good:

Content

- *A clear thesis statement* presenting your position on the issue
- *Relevant reasons and/or examples to explain your view, with supporting details*

  These may come from the documents, from your own experience, or from both. If you use information from the documents, be sure to use it accurately.

Note: You might also want to mention any important *counterarguments* and say why they are not convincing.

Organization

- *An introduction that draws in the reader and sets up the topic*
- *A topic sentence* for each paragraph in the body of the essay
- *A clear, strong conclusion*

Use of Language

- Check your wording, grammar, spelling, and punctuation.
- Use your own words.

If you quote any statements, use quotation marks.

Thesis statement: sometimes called a *theme statement.*

Topic sentence: sentence that states the main idea of a paragraph.

Counterarguments: arguments that can be used to "counter," or oppose, arguments made by the other side

Argumentation Essay Rubric

EXEMPLARY (5)

An EXEMPLARY response meets all of the requirements for a score of 4 and distinguishes itself with such qualities as *insightful analysis (recognizing the limits of an argument, identifying possible assumptions and implications of a particular position); intelligent use of claims and evidence to develop a strong argument (including particularly well-chosen examples or a careful rebuttal of opposing points of view); or skillful use of rhetorical devices, phrasing, voice and tone to engage the reader and thus make the argument more persuasive or compelling.*

CLEARLY COMPETENT (4)

The response demonstrates a competent grasp of argument construction and the rhetorical demands of the task, by displaying all or most of the following characteristics:

Command of Argument Structure

- *States a clear position on the issue of whether schools should sell junk food to students*
- *Uses claims and evidence to build a case in support of that position*
- *May also consider and address obvious counterarguments*

Quality and Development of Argument

- *Makes reasonable claims about the issue*
- *Supports claims by citing and explaining relevant reasons and/or examples*
- *Is generally accurate in its use of evidence*

Awareness of Audience

- *Focuses primarily on content that is appropriate for the target audience (students, teachers, administrators)*
- *Expresses ideas in a tone that is appropriate for the audience and purpose for writing*

DEVELOPING HIGH (3)

While a response in this category displays considerable competence, it differs from Clearly Competent responses in at least one important way, such as a

- *Vague claim*
- *Somewhat unclear, limited, or inaccurate use of evidence*
- *Simplistic reasoning*
- *Occasionally inappropriate content or tone for the audience*

DEVELOPING LOW (2)

A response in this category differs from Developing High responses because it displays problems that seriously undermine the writer's argument, such as

- *A confusing claim*
- *A seriously underdeveloped or unfocused argument*
- *Irrelevant or seriously misused evidence*
- *An emphasis on opinions or unsupported generalizations rather than reasons and examples*
- *Inappropriate content or tone throughout much of the response*

MINIMAL (1)

A response in this category differs from Developing Low responses in that it displays little or no ability to construct an argument. For example, there may be *no claim, no relevant reasons and examples, no development of an argument, or little logical coherence throughout the response.*

Peer Review Sentence Starters

Directions: Use these sentence starters as a springboard for your Peer Review comments in the Dialogue box in CRITERION. Fill in the blanks with specific details from your classmate's writing so that you give them useful feedback (both positive and constructive) that is connected to the scoring rubric and will help them write a stronger essay.

Argument/ Claim and Thesis Sentence:

It sounds like you think junk food [should/should not] be allowed in schools. But I'm having trouble finding your thesis statement.

You helped me really understand _____.

I think you could add more evidence from the _____ article into your essay. This might make your point about _____ clearer.

You did a good job explaining _____, but I think your [explanation/evidence/claim] could be stronger
about _____.

Evidence From the Text:

I like how you used textual evidence from the _____ article to show _____.
Maybe you could use textual evidence from the _____article add to your point about _____.
I think you could explain more about how _____and _____are similar/ different.
Could you add more details or some evidence from the _____ article to support your claim?
It was/wasn't easy to follow your ideas about_____. (Maybe you could add more details.)

Writing: Language and Conventions:

Go back to your essay and take a look at your [capitalization/punctuation/grammar/ spelling].
You have [few/some/many] typos in your essay. Don't forget to fix these before you submit your final draft!

Appendix B. Teacher's Guide

Learning Opportunity Teacher's Guide

September 21–October 2, 2015

Day 1: Monday, September 21

Tasks & Activities	Rationale/Goal & Notes to Teachers
Quick introduction to Junk Food topic: Review table of contents (p. 2) Free Write (p. 3) Pair & Share: Partner's thoughts (p. 3) Whole Class: Pro/Con t-chart (p. 4)	To introduce the topic and generate thinking, students will begin with a free write exercise. Students will then share their writing with a partner, and as a class they will generate a pro/con chart on the topic. The goal is to establish students' prior knowledge on this topic, and to illuminate their existing opinions and reasoning/evidence. Over the next few days, they will read source material and gather additional information that will further inform their position or change their position altogether.

Teacher Notes:

Day 2: Tuesday, September 22

Tasks & Activities	Rationale/Goal & Notes to Teachers
Do now: Stand up in the room: one side Yes JF, one side No JF in schools. Have a brief discussion about the topic by asking students to share their reasons.	At first, students will likely be engaged by having to take stance on the issue and having an opportunity to explain their positions. However, they might quickly realize how difficult it can be to articulate their positions with evidence.
Argumentation vocabulary words—building a common ground for work in argumentation (p. 5)	Students will learn vocabulary words related to argumentation and ideally the teacher will post the words and definitions around the room and make multiple references to these words during the course of this unit of study.
Sorting arguments (p. 6)	The process of forming a strong argument is also important for the students to understand, so that they have an explanation for the activities that will come in the following days. Research skills depend critically on being able to understand and **summarize** individual sources, on **analyzing** how different sources develops their ideas, and using **textual evidence** to support that analysis. Argument skills require students to understand how a source text makes a **claim**, presents **reasons** to support that claim, and backs up those reasons with **factual evidence**.

Teacher Notes:

Day 3: Wednesday, September 23

Tasks & Activities	Rationale/Goal & Notes to Teachers
How to write a summary—teacher will review the Summary Guidelines. (p. 7) Read "Yes!" article, encouraging students to highlight/write in margins (p. 8) Go through Sample Summaries (p. 9–13)	Students will discuss how to write a summary so that they understand what's expected of them as they read some source articles and write their own summaries over the next few class periods. By reading sample summaries and identifying the issues in them, they will become more attuned to what to include and what to discard from their own summaries. These activities will help students learn how to analyze texts and identify main and supporting ideas. If this is the first time students have been exposed to criteria for writing summaries, make sure they understand the purpose of a summary: to convey, in their own words, the main points and most important supporting details in a source, not to give their own ideas or opinions about the topic. The *Guidelines for Summarizing* is a useful precursor for the later tasks, when once again, students will use a new set of guidelines to help them plan and write their essays

Teacher Notes:

Day 4: Thursday, September 24

Tasks & Activities	Rationale/Goal & Notes to Teachers
Read "Might Not Solve" and "No" articles. Use Summarization Guidelines to write a summary for each article. (p. 14–15)	Students should use their understanding of how to write a strong summary from the previous day's work along with the Summarization Guidelines in order to draft their own summaries of the articles. This independent reading and writing time will give the teacher an opportunity to conference with those students who are struggling, catch students up who may have been absent, and also help students gather information on the subject that they will ultimately use in their essays, all while practicing the skills of summarization.

Teacher Notes:

Day 5: Friday, September 25

Tasks & Activities	Rationale/Goal & Notes to Teachers
Peer review summaries from yesterday using the Guidelines as a checklist. Evaluate evidence for claims (p. 16–17)	Preview next week: review essay assignment, rubric, expectations. This activity will warm students up to peer review that will continue during the essay writing next week. Helping students think through the criteria for a "good" summary will help them connect the purpose of peer review focused on a rubric upon which their final essays will be graded. Reviewing the essay assignment now will help

Teacher Notes:

Day 6: Monday, September 28 (Anna Litz on-site afternoon only)

Tasks & Activities	Rationale/Goal & Notes to Teachers
Move into essay	Students need to know what's expected of them before they attempt a task, so review of the task and rubric will be beneficial to start the week.
Review essay writing guidelines, (p. 18) Review argumentation rubric (p. 20)	Having worked through the preceding activities, students should be in a strong position to write a thoughtful essay supporting their position on the issue.
Students log into CRITERION and review the CRITERION Assignment *Junk Food* by clicking on the assignment's name.	Working through the writing process and beginning with a writing plan in class will help students internalize the steps to the writing process.
Students start their CRITERION writing plan, if time allows to plan their essay using planning tool and evidence from texts	By completing a graphic organizer in CRITERION, students can organize their thinking and ensure they're on the right track for completing the assignment appropriately. Also, at this stage they can gather and organize their evidence both from the 3 articles presented last week and their own thinking. Teachers should encourage students to use a combination of reasons and evidence that were generated last week from other students during class discussions, from their own thinking, and from the source material.
	To help students think critically, it is important for them to be able to identify counterclaims and draft strong rebuttals to help bolster their own position statements.

Critical CRITERION Notes:

1. Students access CRITERION via http://criterion-pilot.ets.org using **Firefox** with their Student IDs as their Username & Password
2. If time allows, CRITERION Writing Plans that are started, **should NOT be submitted to CRITERION, but SAVED to finish in class the next day**.

Teacher Notes:

Day 7: Tuesday, September 29 (Anna Litz on-site)

Tasks & Activities	Rationale/Goal & Notes to Teachers
Mini Lesson: How do we peer review each other's writing plans? Using the Writer's Checklist as a guide (p. 19) Finish writing plan (Attempt 1) and submit it. Peer review another student's plan (if time, review more than one other student in your PR group).	Providing students with sentence starters for peer review will give them the tools necessary to not only make strong comments to each other and help normalize critique, but will also help them internalize what's required of a strong essay that will score well on the rubric.

Critical CRITERION Notes:

1. The Writing Plan and Peer Review must be **completed by the end of this class period.**

2. After clicking SAVE on CRITERION's Plan tab, students will then be able to copy their Writing Plans and paste it into CRITERION's Response tab to SUBMIT it to CRITERION. They will not be able to complete Peer Review until this is done.

3. In the Response tab, students click on a hyperlinked name of a writer appearing in the Peer Review column to give him/her Feedback via the Dialogue button.

   Note. When a peer reviewer is giving feedback to a student writer's essay via Dialogue, **the reviewer should note the writer's Name and his/her Attempt number followed by their feedback**. (e.g., *Bill Smith-Attempt 1, I like how you used evidence from the ____ source material to show ____.*)

   Note. Students should review feedback on their own writing plans by clicking on Attempt 1 in the Response tab and clicking on the Dialogue box.

Teacher Notes:

Day 8: Wednesday, September 30 (Anna Litz on-site)

Tasks & Activities	Rationale/Goal & Notes to Teachers
Write during class. Submit a draft by end of class (Attempt 2) Peer review if time allows, using Sentence Starters (p. 21)	It's valuable for students to complete their writing in class, where the teacher can circulate the room and conference with struggling students and help those who might be stuck for any reason. Encouraging additional peer review is also beneficial to students who both give and receive the feedback. The Sentence Starters were written to mirror the expectations in the argumentation rubric. This will help students focus on critical and meaningful feedback to each other and normalize the critique process.

Critical CRITERION Notes:

1. Students click on Attempt 1 in the *Response* tab and then click REVISE to begin their essay drafts (Attempt 2).
2. Students should click SAVE frequently and **do NOT click SUBMIT until they have completed their essay draft by the end of class.**
3. If time allows for Peer Review of **Attempt 2**, follow previous steps: In the *Response* tab, students click on a hyperlinked name of a writer appearing in the *Peer Review* column to give him/her Feedback via the *Dialogue* button.
    a. When a peer reviewer is giving feedback to a student writer's essay via *Dialogue*, **the reviewer should note the writer's Name and his/her Attempt number followed by their feedback**. (e.g., *Bill Smith-Attempt 2, I like how you used evidence from the ____ source material to show ____.*)
    b. Students should review feedback on their own writing plans by clicking on **Attempt 2** in the *Response* tab and clicking on the *Dialogue* box.

Teacher Notes:

Day 9: Thursday, October 1 (Anna Litz on-site)

Tasks & Activities	Rationale/Goal & Notes to Teachers
Finish peer review on Attempt 2. Review CRITERION automated feedback and peer feedback, and students should also compare their essay to the argumentation rubric. Start final draft (Attempt 3).	It is critical to get students actively engaged in applying the rubric to their own and other's writing and to provide useful feedback to one another based on the rubric. It is important to engage students with the content and to force them to confront different perspectives.

Critical CRITERION Notes:

1. Complete Peer Review for **Attempt 2** by following previous steps: In the *Response* tab, students click on a hyperlinked name of a writer appearing in the Peer Review column to give him/her Feedback via the *Dialogue* button.

   Note. When a peer reviewer is giving feedback to a student writer's essay via *Dialogue*, **the reviewer should note the writer's Name and his/her Attempt number followed by their feedback**. (e.g. Bill Smith-Attempt 2, I like how you used evidence from the ____ source material to show ____.)
   Note. Students should review feedback on their own writing plans by clicking on **Attempt 2** in the *Response* tab and clicking on the *Dialogue* box.
2. Students click on **Attempt 2** in the *Response* tab and then click REVISE to begin their final essays (**Attempt 3**).
3. Students review CRITERION's diagnostic feedback via the trait category drop down menus appearing above their **Attempt 2** in the Revise screen. Remind students that feedback changes as each category and sub-category is clicked and that students should read roll-over messages and click on WRITER'S HANDBOOK, if they need further explanation, examples, commentary about the error detected.
4. Students should click SAVE frequently and **NOT click SUBMIT so that they can finish their final essays in class the next day.**

Teacher Notes:

Day 10: Friday, October 2

10 10/2	Submit final draft of essay in class (Attempt 3). Submit to teacher a self-score on argumentation rubric.	At this point, students will have completed a graphic organizer writing plan, at least two iterations of the essay, and at least two rounds of peer review. Asking them to provide a self-assessment on the argumentation rubric will further encourage the students to be critical of their own work.

Critical CRITERION Notes:

1. Students should click SAVE frequently and **NOT click SUBMIT until they have completed their final essays by the end of class.**

Teacher Notes:

*** All packets should be collected from students and sent back to ETS via the prepaid shipping labels provided to the school once the project is complete and teachers have had sufficient time to review the student responses in the packets.*

Notes

[1] http://www.americaslibrary.gov/aa/twain/aa_twain_huckfinn_1.html; http://www.americaslibrary.gov/aa/twain/aa_twain_huckfinn_2.html

[2] We structured the peer review prompts to focus students on the content and the argument, hoping for more substantive revisions, and not just proofreading, in students' final sessions, but in this implementation, we were not yet successful in getting students to make larger revisions, consistent with information in the literature that suggests student writers tend to conflate revision with editing and proofreading and to undertake mostly proofreading actions when asked to revise.