# Writing Traits: Examining the Consistency of Behavioral Patterns in Writers' Composing Processes

Emily Dux Speltz[1], Jens Roeser[2], Wren Bouwman[3], Evgeny Chukharev[3], & Mark Torrance[2]

[1]Embry-Riddle Aeronautical University Worldwide | US

[2]Nottingham Trent University | UK

[3]Iowa State University | US

**Abstract:** Individual differences in writing processes have been well documented, yet the stability of a writer's behaviors across writing tasks is less studied. Behavioral stability serves as a foundational assumption in writing research, crucial for describing individual differences and essential for effective writing instruction that seeks to modify behavioral patterns across different writing tasks. Based on data from keystroke logs and eye movement, we determined the writing process stability of 30 writers across three argumentative writing tasks. This yielded 17 behavioral measures from which we identified four factors: pausing, revising, reading own text (lookback), and linearity. Hierarchical cluster analysis across scores from these factors then grouped the 90 writing processes into three clusters based on behavioral similarity. Sixteen participants remained in the same cluster (adopted similar writing processes) across all tasks; the remainder switched clusters once, and tended to be those writers whose behavior was already closer to the centroid of the cluster to which they switched. Overall, stability was much higher than predicted by chance. Results are consistent with behavioral stability across writing tasks, supporting the notion that writers have habitual cognitive processes that they carry with them between tasks.

**Keywords:** writing processes, revising behavior, pausing behavior, eye tracking, cluster analysis

journal of WRITING RESEARCH

Dux Speltz, E., Roeser, J., Bouwman, W., Chukharev, E., & Torrance, M. (Accepted for publication, 2026). Writing traits: Examining the consistency of behavioral patterns in writers' composing processes. *Journal of Writing Research, volume(issue),* ##-##. DOI: xx

Contact: Emily Dux Speltz, Embry-Riddle Aeronautical University Worldwide, 1 Aerospace Boulevard, Daytona Beach, FL 32114 | US – duxspele@erau.edu

## 1. Writing traits: Examining the consistency of behavioral patterns in writers' composing processes

Interviews with established authors suggest substantial differences in how they approach composing text. Contrast the following from the Paris Review Art of Fiction interview series: John Irving reported planning meticulously, and therefore always knowing what he was going to say before starting to write (Hansen, 1986). Joan Didion reported her fiction writing growing from a first sentence, with structure and order emerging at later stages in the process (Kuehl, 1978). Margaret Attwood reported an approach that lay between these two extremes (Morris, 1990).

There is a now quite dated body of commentary and research that aims to capture these kinds of differences in the ways in which similarly competent writers behave when composing text. Writers have variously been described as Motzartians and Beethovians (Spender, 1952), the mantic and the manic (Shapiro, 1948), makers and the possessed (Wellek & Warren, 1963), planners and discoverers (Chandler, 1992), doers and thinkers (Hartley & Branthwaite, 1989), planners and revisers (Torrance et al., 1994), water colorists, architects, bricklayers, sketchers, and oil painters (Wyllie, 1993), and elaborationist, low self-efficacy, reflective-revisionist, spontaneous-impulsive, and procedural (Lavelle, 1993).

More recently, researchers have captured writing profiles for digital writing tasks using screen recording and/or keystroke logging software. For example, Kim (2020) identified four distinct profiles among undergraduate writers based on their planning, revision, and internet search behaviors during digital writing tasks, labeling the profiles as revision-based, plan-based, search-based, and correction-based writers. Vandermeulen et al. (2024) classified upper-secondary students' writing process profiles based on their production measures captured in keystroke logs and their use of sources, resulting in profiles they called the notetaker, the source reader, and the thinker/focused reader.

Our aim in this paper is not to add our own categorization to this already long list of possible ways of describing individual differences among writers. Instead, we test the more general claim, underlying all such classifications, that how writers approach a writing task is to some important extent stable across writing tasks: that is, how writers behave when they write varies systematically across writers, and these systematic differences represent a stable trait that is evident whenever they perform tasks that require the production of texts with a broadly similar genre. By task we mean a distinct act of text production prompted by a new topic, even when genre and overall demands are held constant, a usage that overlaps with what Delgado-Osorio et al. (2025) termed distinct task performances. (This similar-genre caveat is clearly necessary. We would not necessarily expect writers to engage in similar processes when writing, for example, a research paper, a short story, or an email. The extent to which writing profiles remain stable across different genres is an interesting question that we return to in our discussion. However, it is beyond the present scope.)

## 1.1 Exisiting evidence

The theory that there are stable individual differences in writing behavior is plausible and resonates with writer experience. It has, however, received relatively little research attention. With a few exceptions that we discuss next, the studies that exist have tended to rely on questionnaires or interviews in which participants are asked about their typical writing practices. Occasionally, writing profiles have been identified from more direct observation of writing behavior or real-time self-report (e.g., think-aloud). In both cases, studies have tended to focus on single tasks, and therefore have not addressed whether the patterns that the researchers observe are a stable feature of the writer (a trait), or just a feature of the writer's response on that day to the particular topic that they have been set. To our knowledge, Torrance et al. (2000) is the only study to have used questionnaire-based methods with the same writers across multiple tasks. Forty-eight students completed questionnaires about the actions they took when writing specific coursework essays shortly after they had finished writing, for a median of six essays per student across three years of their undergraduate studies. Cluster analysis of these data identified four general patterns of self-reported writing behavior and found that 81% of students fell into one cluster more than any other, and that this cluster represented, on average, their reported behavior for 69% of their essays.

However, retrospective reporting, particularly by questionnaire, may well exaggerate cross-task behavioral stability. Students' recall and reporting of how they performed a recent task will be shaped in part by what they themselves believe is their normal writing strategy. We know of two other longitudinal studies that used more direct self-report measures. Levy and Ransdell (1995) recorded writers thinking aloud while composing ten texts spread across a university semester, and then categorized their activity into episodes of planning, composing text, reading, and revising. Their analysis focused not on the total time spent in these activities but on their sequencing, measured by counting switches between the current activity and one of the other three activities (counts of moves between planning to composing text, reviewing to revising, and so forth). Within-writer correlation among these scores across consecutive tasks gave a measure of consistency of behavior. The mean correlation across all writer and task pairs was 0.76. Torrance et al. (1999) studied the behavior of 17 undergraduate students writing two extended coursework essays. Students heard a tone at regular random intervals while writing that prompted them to report their current activity. These data were then cluster-analyzed to identify three "writing procedures" which were captured as variation across the writing timecourse in time spent in planning, drafting, revising, reading sources, and writing a polished text. They found that 76% of students (n = 13) remained in the same procedure cluster across the two tasks.

A handful of studies have used direct observation to identify writing profiles. Kim (2020) used screen recordings to determine behavioral differences among undergraduates and identified four different profiles: revision-based, plan-based, search-based, and correction-based. These profiles were linked to differences in text quality, with revision-based writers producing the strongest texts. Vandermeulen et al. (2024; an analysis of a subset of the data reported by Van Steendam et al., 2022, discussed below) used keystroke data and time spent engaging with supporting materials to identify patterns in secondary school students writing from sources. Neither study established whether the patterns that they identified remained stable within writers across tasks.

Few behavioral studies have investigated profile consistency across tasks. Van Waes and Schellens (2003) cluster-analyzed behavioral measures to identify different writing-behavior profiles. They directly observed behavior using keystroke logging and video recording to give a large number of measures relating to pausing, revision, and output rate. Cluster analysis of these identified five distinct profiles: initial planners (few revisions, lots of pausing during initial planning), fragmentary stage I writers (many revisions in the early stages of writing), stage II writers (many revisions after the completion of the first draft), non-stop writers (few revisions, less pausing, below-average writing time), and average writers (average values across measures). They conducted two studies, each with writers performing two tasks. In the first, writers composed by keyboard but with substantial variation in the number of lines of text displayed on the computer monitor. Here, 11 of 20 writers remained in the same cluster across both tasks. In the second experiment, writers composed by keyboard and with pen and paper with all but two writers changing profiles across modes.

In by far the largest existing study of this kind, Van Steendam et al. (2022) used keystrokes to identify patterns of behavior in 658 Dutch upper-secondary students who each performed four writing-from-sources tasks: two informative and two argumentative. Data comprised typing speed, total writing time, and time spent attending to source texts. They identified four behavioral patterns focused specifically on source text use. Consistency within individuals was lower than in the studies cited above, with 21% of students falling within the same pattern category for all four tasks, and a further 26% fell in the same category for three out of four tasks. Comparing these findings with those of the direct-observation studies discussed above is tricky, however, both because tasks varied in genre, and because of the focus on writing from sources.

## 1.2 The present study

Our aim, therefore, was to establish robust evidence for writing profiles—the idea that writers carry with them across tasks particular ways of writing that are different in important ways from the approach taken by other writers. We tested the following hypotheses:

1. Different writers produce texts in different but similarly successful ways: There are non-trivial differences in the behavioral patterns of experienced writers completing a

specific writing task, and these differences are independent of how well they perform (i.e., of the quality of the completed text).

2. These behavioral patterns represent a trait that is stable across same-genre tasks: Writers have habitual ways of writing that they adopt whenever they produce text of a particular type, independent of the topic. There is some existing evidence to support this claim (Levy & Ransdell, 1995; Torrance et al., 1999; 2000). The present study builds on this by presenting evidence from a larger sample studied using a broader range of behavioral measures than has been the case in previous research.

To achieve this aim, we took direct behavioral measures of the writing process by logging writers' keystrokes, following Van Waes and Schellens (2003) and Van Steendam et al. (2022), and also tracked their eye movements to capture behavior associated with writers reading their developing text. Our focus was on composition without the support of external sources. From these data, we extracted 17 measures, adapted from Dux Speltz et al. (2022), aimed at capturing four categories of writing behavior: pausing, reviewing, revising, and non-linearity. Our choice of these as descriptors of activity that happens during writing draws on a traditional understanding of written composition processes that distinguishes between planning—determining what to say and how to say it; translation—converting this plan into language; and reviewing and revising—reading and where necessary changing the text that has been produced (Hayes, 1996; Hayes & Flower, 1980). Several researchers have argued that good quality written composition requires that writers adopt explicit (deliberate), specific planning and revision strategies that modify the activities they engage in when composing text (Berninger & Winn, 2006; Harris & Graham, 1996; but see Torrance et al., 2015).

Pausing refers to extended intervals of inactivity during text production during which keys are not pressed or the pen is not moving across the page. Exactly how these pauses map onto underlying cognitive processes is complex, and dependent on pause threshold (e.g., Torrance & Conijn, 2023). Generally, though, greater pausing across the production of a text is typically assumed to be associated with the writer spending more time planning what to say next (e.g., Flower & Hayes, 1981; van Hell et al., 2008). (Pausing is also, of course, associated with reading back over text already written. However, this was accounted for separately in our study.) Total time spent pausing will increase if writers engage in explicit planning strategies. However, pausing will also increase due to more frequent lower-level planning—making deliberate word-choice decisions, for example (e.g., Olive et al., 2009).

Reviewing—looking back over text already written—can also be associated with both low-level text features—checking over the previous few words to find a spelling error, for example—and more global reading of the full text. A small number of papers have studied reviewing processes directly by tracking writers' eye movements (Alamargot et al., 2010; Beers et al., 2010; Chukharev-Hudilainen et al., 2019; Torrance et al., 2016; Wengelin et al., 2009). Torrance and co-workers found that university students composing argumentative texts—the same population and type of task as in the present study—spent a mean of around 13% of total writing time looking back into their text. Of that overall time spent looking back, 36% of occasions were associated with sustained reading. In the remainder, writers' gaze tended to hop around between different locations in the text; thus we use "lookback" rather than "reading" to describe gaze within existing text in the present study. Lookback was roughly equally divided between looks within the sentence that the writer was currently composing and looks further back into the text (Torrance et al., 2016). These findings suggest that reviewing is sometimes, but not always, a strategic activity. Reviewing sometimes results in revision, although to our knowledge, no published research reports in what proportion of lookback events this occurs. As might be expected, the vast majority of revisions, even in university student writers, involve minor, word-level corrections, with far fewer changes made at a level that affects the overall message of the text (Conijn et al, 2021; Faigley & Witte, 1981). As with planning, some researchers have argued for the value of writers strategically modifying their behavior to include dedicated time for reading and evaluating their text (e.g., De La Paz & Sherman, 2013; Graham, 1997).

Finally, non-linearity represents the extent to which the production of a text is associated with production at locations other than the front edge of the text (e.g., Buschenhenke et al., 2023; Lo Sardo et al., 2023). A maximally linear writing process might be associated with a writer carefully planning content in advance and then sticking rigidly to their plan, or with a "knowledge telling" process (Scardamalia et al, 1984) in which writers uncritically write down content as it occurs to them. Writers who either strategically or reactively shape their message and structure while they are writing are likely to write less linearly.

Participants in our study completed short argumentative writing tasks. While the prompts varied in topic, they were similar in structure and genre. We used custom software that captured and integrated keystroke and eye-movement data to extract moment-by-moment information about what text the writer was typing and on what words within their developing text the writer was fixating. From these data, we extracted summary statistics for describing the behavior of each participant for each task. We then used a confirmatory factor analysis and then cluster analysis to classify behavior into discrete patterns, again following previous studies (Crossley et al., 2014; Torrance et al., 2000; Van Waes & Schellens, 2003). A tendency for writers to remain in the same process cluster across tasks would represent evidence that writers show cross-task behavioral stability, and therefore provide evidence for the psychological validity of "writing profiles."

## 2. Methods

### 2.1 Design and participants

University students (N = 30) each composed three texts in response to similar writing prompts over the span of three tasks separated by at least one week. With three exceptions (maximum 28 years), all participants were aged between 18 and 24 years.

Participants were recruited through both word-of-mouth and advertisements in general education classes at the university. Participation was voluntary and not for course credit. They received a 90 USD gift card as compensation for their involvement in the study. Participants had normal or corrected-to-normal vision, English as a primary native language spoken during upbringing, and were QWERTY keyboard users.

Participants wrote their essays using CyWrite (Chukharev-Hudilainen et al., 2019). CyWrite is an open-source web-based text editor that provides a user interface with a word-processing experience similar to commonly used text editors like Microsoft Word and Google Docs. It captures log files containing the timings and the results of each keystroke and cursor movement. CyWrite, combined with eye-tracking hardware, also captures writers' eye movements within their emerging texts. Eye movements are captured in such a way as to permit identification of which word within the emerging text is being fixated. This makes it possible to determine, for example, the number of words that separate the word that is currently being fixated from the front edge of the text. Participants were not given access to spell-checking or grammar support and were not able to access online external resources.

### 2.2 Writing tasks

Participants were instructed to write a five-paragraph essay, performing to the best of their ability, on separate occasions, in response to each of the following prompts: (A) "Some people have argued that animals should be given similar rights to humans. To what extent do you agree or disagree with this statement?" (B) "Some people have said that finding green technologies (e.g., solar power, wind power, etc.) that do not require the burning of fossil fuels should be the focus of our efforts to avert the climate crisis. To what extent do you agree or disagree with this statement?" (C) "Science should aim to discover truth about the world, without concern for practical application or wealth creation. To what extent do you agree or disagree with this statement?". Topics were chosen to be accessible and familiar to college students in the United States. Prompts were counterbalanced within participants across tasks.

### 2.3 Equipment and procedure

Participants composed text on desktop computers with Dell U2415 monitors. The monitors had 1920x1200 display resolution with overall visible screen dimensions of 518.4 millimeters wide by 324 millimeters tall. Computers were equipped with

GazePoint GP3 eye trackers, used without head restraint. Participants sat around 580 millimeters away from the eye-tracking device, and they were positioned so that horizontal gaze was center-screen.

At the start of each task, participants completed nine-point calibration, ensuring calibration error of less than 75 pixels (2.0 degrees of visual angle). Participants then completed the writing task, taking as much time as they required. The mean time-on-task across participants was 25 minutes (SD = 10).

## 2.4 Behavioral measures

Participants' writing processes in each task were captured in terms of 17 different real-time behavioral measures. Appendix A presents descriptive statistics and correlations across these measures. We theorized that these measures captured four higher-order process dimensions as follows.

### Pausing

We used finite (two-component) mixture models to empirically distinguish between rapid inter-keystroke intervals associated with keystroke motor planning (with means typically in the range of 150 to 170 ms for competent, adult typists), and inter-keystroke intervals with substantially longer durations. We term these "hesitations" to distinguish them from more traditional measures based on a fixed threshold. We refer readers to Roeser et al. (2024) and (2025) for a full description of this approach. In brief summary, mixture models are arguably necessary in this context because the various processes necessary to generate text typically run in parallel with motor output. Interkey interval durations therefore result from two distinct data-generating processes, one associated with the case when planning of what to write next is completed entirely in parallel with previous output, and the other when planning is not completed in parallel and therefore spills over into the interval before the next keystroke. These separate data-generating processes result in different, separate probability distributions. IKI data represent a mixture of these, and a two-component model is necessary to separate them out. Whether an interkey interval is hesitant (planning is not entirely completed in parallel with previous output) is therefore derived from the data itself rather than determined a priori.

For each text we calculated the following three measures:

*Hesitations before sentences*. Percentage of pre-sentence interkey intervals (time between the keypress generating the first character in a sentence and the preceding character—either a space or a paragraph marker) that were hesitant.

*Hesitations before words*. Percentage of pre-word interkey intervals (time between the keypress generating the first character in a mid-sentence word and the preceding) that were hesitant.

*Hesitations within words*. Percentage of intervals between keypresses within words that were hesitant.

## Reviewing

These measures were extracted from eye movement data and describe cases where writers looked back into the text that they had already written. We defined lookback as a series of two or more eye movements that took place between two keystrokes. To qualify as a lookback event, eye movement had to meet the following criteria: (a) mean gaze location focused at a minimum of one word and five characters behind the last typed letter (to avoid cases where the writers was predominantly fixating the word that they were currently typing), (b) a minimum of 80% of fixations that were on text behind the current cursor position, and not forward to text in front of the cursor position or on parts of the screen that did not display text, (c) the series of eye movements did not start with fixation on the writing prompt text (displayed at the top of the screen), and (d) fewer than 50% of the eye movements were directed toward reading the writing prompt.

*Lookbacks before sentence*. Lookbacks that occurred when the cursor was in the sentence-initial position, expressed as a percentage of the total number of sentence-initial keystrokes.

*Lookbacks before word.* Defined in the same way as the previous measure but for cases where the cursor was at the start of mid-sentence words.

*Lookbacks within word*. Defined in the same way as the previous measures but for cases where the cursor was mid-word.

*Length of non-lookback sequences*. The mean number of keystrokes in sequences not broken by the writer looking back into their already-written text, with lower values indicating greater tendency to review.

*Lookback duration*. Mean lookback sequence duration, across all text locations.

*Long lookbacks*. Number of lookback sequences that were of longer duration, as defined by a two-distribution finite mixture model (described above) and expressed as a percentage of total number of lookback sequences.

## Revising

These measures captured behaviors associated with writers making any changes to the text that they had already written.

*Deletions before sentences.* The number of cases where the cursor was at the start of a sentence (i.e. after sentence-delimiting punctuation and the following space) but instead of the next keystroke forming that first character of the new sentence, it involved deletion of previous text, expressed as a percentage of all keystrokes that occurred at this sentence-initial position.

*Deletions before word.* Defined similarly to deletions before sentences, but for cases where the cursor was positioned after a mid-sentence space character.

*Deletions within word.* Defined similarly to deletions before sentences and words, but for cases where the cursor was positioned before a mid-word character.

*Length of production sequence*. Mean number of keystrokes entered in an unbroken sequence between deletion or cursor-move actions.

### Non-linearity

Non-linearity refers to the extent to which participants jumped to different insertion points within their text. A writer who composed entirely at the front edge of their text with no jumps back to insert text at an earlier point would have adopted an entirely linear process. We measured deviation from this (i.e. the extent to which the composing process was non-linear) in terms of the following:

*Number of jumps.* A count of the times that the writer's insertion point shifted to a new location within the text.

*Number of edges.* A count of new transition points or "edges" introduced into the writing sequence as a result of the jumps made when the writer's inscription point changes locations within the text. Each time the writer inserted text at a new (not previously visited) text location was recorded as a new edge.

*Number of major blocks.* A count of the number of blocks of text that the writer produced without jumping. To qualify as a major block, the writer needed to produce at least 615 characters of text (following Dux Speltz et al., 2022). A writer who composed entirely at the front edge of their text would score just one block. Values above one therefore represent an index of non-linearity.

*Number of sustained reading events.* How frequently the writer engaged in extended linear reading of at least three consecutive fixations along a horizontal line of text, as detected by the eye-tracking data.[1]

### Text-based measures

We analyzed final texts to extract several measures capturing different aspects of writing proficiency and quality. The total *number of sentences* and total *number of words* provided measures of overall productivity. The *open-to-closed class word ratio*, defined as the ratio of open-class words (nouns, verbs, adjectives, and adverbs) to closed-class words (e.g., articles and pronouns), served as an indicator of information density and ideational richness in the language used. *Lexical diversity*, measured using the MTLD statistic (McCarthy, 2005) as a sensitive and text-length independent metric, quantified the sophistication of the vocabulary employed in the writing sample (Torruella & Capsada, 2013).

[1] Inspection of bivariate correlations between the *Number of sustained reading events* variable and other variables loading onto the non-linearity factor suggested a strong association, justifying its inclusion in the non-linearity factor as opposed to the reviewing factor containing other eye movement variables. Conceptually, this association makes sense, as the *Number of sustained reading events* variable may capture instances of sustained reading related to rereading previous text when making jumps within the material, reflecting the non-linear reading processes represented by the non-linearity factor.

We also scored texts for quality using a rubric that captured context, substance, organization, style, and delivery, with scores ranging from 5 = "Excellent and outstanding" to 1 = "Inadequate" (Appendix B). All texts were scored by two trained raters (undergraduate students in linguistics) with inter-rater reliability of .77 (Krippendorff's alpha; Krippendorff, 1970).

## 3. Results

### 3.1 Approach to analysis

We reduced our 17 writing behavior measures to four broader process dimensions, as detailed above, using confirmatory factor analysis to establish whether this four-factor structure did in fact provide a good fit to our data. We then used cluster analysis to categorize the 90 writing processes that we observed (30 participants each performing 3 tasks) into groups of broadly similar processes.

We then established the extent to which each writer adopted similar processes (i.e., remained within the same process cluster) across writing tasks, and we also determined whether movement between clusters tended to be associated patterns of behavior that were already less typical (i.e., close to the boundary) of the cluster from which they were moving.

Finally, we compared text quality across clusters as a check that cluster membership was not simply determined by students' writing ability and/or motivation. We describe each of these analyses in the sections that follow.

We used Bayesian mixed-effects models for statistical inference, with random intercepts for participants and random by-participant slope adjustments for prompt and for task (first, second, and third), using weakly informative priors. These were implemented in the *brms* R package (Bьrkner, 2018). We report Bayes factors (BFs), calculated by the Savage Dickey method (Wagenmakers et al., 2010) as evidence for or against null hypotheses. For readers not familiar with this approach, by convention Bayes factors of less than three show only anecdotal evidence for or against a null hypothesis. Bayes factors of greater than five provide moderate to strong evidence.

### 3.2 Confirmatory factor analysis

We compared five models of the factor structure of the 17 process measures: A single factor model (M1), a model in which revising and linearity measures loaded onto one factor and reviewing and pausing measures loaded onto a second factor (M2), a three-factor model in which reviewing and pausing loaded onto separate factors (M3), our hypothesized four-factor structure (M4; Table 1), and finally a model in which lookback frequency (percentage) measures loaded onto a different factor from lookback duration and length of non-lookback sequence (M5). Models were fitted using Bayesian confirmatory factor analysis methods with weakly informative priors using the "blavaan" R package (Merkle et al., 2021). We used leave-one-out cross-validation (Vehtari et al.,

2024) to evaluate model predictive performance, comparing model fits in terms of difference in expected log predictive density (*elpd*), expressed as z scores (Sivula et al., 2020).

M1 and M2 provided poorer fits than M3, and M3 showed a substantially poorer fit than M4 (difference in estimated *elpd* is 5.1 SEs). M5 provided a roughly equivalent fit to M4 (difference in *elpd* is < 1 SE). We therefore preferred M4 as the more parsimonious model. Factor loadings for this model can be found in Table 1, and all model comparison indices in Appendix C.

*Table 1.* Overview of the behavioral measures used in the present study. Factor loadings with 95% probability intervals

| Factor | Measure | Factor Loading |
|---|---|---|
| Pausing | Percentage of hesitations before sentence | 1.00 |
| | Percentage of hesitations before word | 0.67 [0.16, 1.25] |
| | Percentage of hesitations within word | 0.66 [0.19, 1.19] |
| Revising | Percentage of deletions before sentence | 1.00 |
| | Percentage of deletions before word | 1.43 [1.13, 1.80] |
| | Percentage of deletions within word | 1.44 [1.14, 1.81] |
| | Length of production sequence | -1.16 [-1.53, -0.86] |
| Reviewing | Percentage of lookbacks before sentence | 1.00 |
| | Percentage of lookbacks before word | 1.09 [0.89, 1.34] |
| | Percentage of lookbacks within word | 1.09 [0.88, 1.33] |
| | Lookback duration | 0.33 [0.07, 0.60] |
| | Percentage of long lookbacks | 0.16 [-0.10, 0.43] |
| | Length of non-lookback sequence | -1.19 [-1.43, -1.00] |
| Non-linearity | Total number of major blocks | 1.00 |
| | Total number of jumps | 1.40 [1.09, 1.78] |
| | Total number of edges | 1.42 [1.11, 1.81] |
| | Number of sustained reading events | 0.99 [0.68, 1.35] |

*Note.* The first measure served as the marker variable for each factor.

## 3.3 Cluster analysis

Each of the 90 writing processes (i.e. the process adopted by each writer for each task) was treated as a single case for the purposes of cluster analysis. Clustering was on the basis of just the four process measure scores, calculated as the sum of the standardized process measure scores weighted by factor loadings. The "hclust" function in R was used to perform the hierarchical cluster analysis on the cases.

Hierarchical cluster analysis was implemented with the *hclust* R function using the complete linkage method, with a Euclidean measure of the between-case distance. This suggested a three-cluster solution[2]. Table 2 summarizes the characteristics of each cluster. Appendix D presents a dendrogram visualization of the clusters. Values represent posterior means and 95% probability intervals of factor loadings separated by cluster. Posterior factor loadings were obtained from a Bayesian mixed-effects model with Gaussian probability function) with fixed effects for all combinations of clusters.

*Table 2.* Cluster characteristics. The lower panel gives mean estimates and 95% probability intervals. For clarity, the upper panel summarizes the pattern of means (++ / – –, processes in cluster are at least 2 SD above / below grand mean; + / –, one SD above or below grand mean)

| Factor | Cluster 1 (n = 43) | Cluster 2 (n = 28) | Cluster 3 (n = 19) |
|---|---|---|---|
| F1 (Non-linearity) | | + + | – – |
| F2 (Revisions) | – | + + | – – |
| F3 (Reviewing) | + | + | – – |
| F4 (Pausing) | | | |
| F1 (Non-linearity) | -0.16 [-1.02, 0.70] | 2.85 [1.79, 3.88] | -3.6 [-4.89, -2.30] |
| F2 (Revisions) | -1.69 [-2.54, -0.82] | 4.15 [3.09, 5.21] | -2.06 [-3.34, -0.80] |
| F3 (Reviewing) | 1.66 [0.79, 2.53] | 1.29 [0.24, 2.33] | -5.42 [-6.73, -4.11] |
| F4 (Pausing) | 0.08 [-0.79, 0.95] | 0.33 [-0.74, 1.38] | -0.46 [-1.73, 0.82] |

Cluster 1, the largest cluster, was characterized by less revision behavior—that is, a lower percentage of deletions before sentence, before word, and within word, and longer uninterrupted production sequences—and more lookback behavior, as measured by the percentage of lookbacks at different locations, the length of typing sequences without lookbacks, total lookback duration, and the percentage of long lookbacks. In other words, writers in this cluster re-read previously written text often and rarely revised their text. The average time-on-task for this cluster was 22.9 minutes (SD = 7.7).

Cluster 2 was characterized by a non-linear writing process with frequent revisions and frequent lookbacks. Writers in this cluster moved the cursor in the text more often than average, thereby changing the point of inscription frequently and producing more writing blocks. They also deleted text more frequently across locations (before sentence, before word, and within word) and had shorter production sequences. Furthermore, they often looked back in the produced text for longer periods of time than average. The average time-on-task for this cluster was 31.7 minutes (SD = 10).

Cluster 3 was characterized by negative loadings on all four factors, making them essentially the opposite behaviors of Cluster 2. In other words, writers in this cluster wrote overall very linearly and rarely revised or looked back in the produced text. Cluster 3 could be characterized as fluent writing processes in which the writer essentially

[2] See https://rpubs.com/endux/stability for R code.

produces the whole text by writing at the leading edge. The average time-on-task for this cluster was 17.9 minutes (SD = 7). This was the smallest of the three clusters of writing behavior, with only 19 of the 90 cases.

### 3.4 Change in cluster membership across tasks

We investigated cluster membership for each participant across the three writing tasks. Of 60 possible switching opportunities (between Task 1 and Task 2, and between Task 2 and Task 3, for each of 30 participants), process cluster remained the same in 43 (72%) of the cases. As can be seen from Table 3, the majority of participants remained in the same cluster for all three classes. Of the remaining participants, all had processes that fell within two clusters. There were no participants whose writing processes fell within a different cluster for each task. This indicates substantially more process consistency across cluster than would be the case if cluster membership was random (Table 3, first column).

***Table 3.*** Cluster switching between tasks. Number and proportion of students

| | Predicted by chance | Observed |
|---|---|---|
| Participant's writing processes all fell within a single cluster. | 4.5 (0.15) | 16 (0.53) |
| Participant's writing processes fell within two different clusters. | 19.8 (0.66) | 14 (0.47) |
| Participant's writing processes fell within a different cluster for each task. | 5.7 (0.19) | 0 (0.00) |

Next, we compared the centrality of observed behaviors within their respective clusters. The intuition is that when a participant's behavior is more prototypical and therefore located more centrally in a cluster, then the participant is more likely to stay within that cluster in subsequent writing tasks. Specifically, in order to explore whether changes between clusters were due to texts landing on the boundary between two clusters, the centroids of each cluster were calculated. Cluster centroids are the middle of each cluster, representing a multi-factor average across all texts in that cluster. Each text's distance from the corresponding cluster's centroid and the distance to other cluster centroids were calculated to determine whether participants who switched clusters were further from the cluster centroid they switched from, and closer to the cluster centroid that they switched to, than those who did not switch clusters. Because each participant completed three writing tasks, this allowed for two across-task comparisons per participant (Task 1 to Task 2, and Task 2 to Task 3), resulting in a total of 60 comparisons across the sample of 30 participants.

Table 4 shows the results of calculating the distances from cluster centroids for each text. Texts from participants who did not switch clusters were closer to the centroid of their cluster (mean = 4.96) than those who did switch (mean = 5.26). A logistic model in which whether or not a participant switched cluster was regressed on distance from cluster centroid found some evidence that switch probability was lower for participants with a larger distance from the initial cluster's centroid (est. = -0.11, PI: [-0.57, 0.29], BF = 4.6). This indicates that participants who switched clusters from one task to the next tended to be further away from the centroid of the cluster they switched from than those who did not switch clusters. As Table 4 shows, the distance to the cluster to which they switched was, on average, shorter than the distance to the other cluster. This suggests that the movement was typically a relatively small drift in the pattern of behavior rather than a participant writing in one definite way for one task and then switching to another quite different pattern for the next.

*Table 4.* Mean (standard deviation) distance of writing processes from the centroid of the cluster for the current writing task and for the next writing task for participants who did and did not change cluster, aggregating across Task 1 to 2 and Tasks 2 to 3

| | Distance to centroid of current cluster | Distance from centroid of switched-to cluster | Mean distance from centroids of other clusters | N |
|---|---|---|---|---|
| Did not switch clusters | 4.96 (1.97) | – | 10.54 (2.38) | 43 |
| Switched clusters | 5.26 (1.75) | 7.65 (3) | 10.3 (3.19) | 17 |

## 3.5 Text quality and productivity

Finally, we established whether the pattern of writing behavior (cluster membership) predicted characteristics of participants' final texts and their total writing time. Model comparison adopted the same methods as for our confirmatory factor analysis described above. A model that included cluster membership as a predictor provided a poorer fit to our data than a baseline in which this predictor was absent (z = 3). We therefore found no evidence that the quality or nature of a text was affected by the cluster membership of the process by which the text was produced (see Appendix E).

There were, however, differences among clusters in total time to complete the writing task, calculated as the time from the first view of the writing prompt to the final keystroke or cursor movement. We modeled these data using the same strategy as for text characteristics. Including cluster membership as a fixed effect gave a better fit to our data than omitting this factor. Estimated mean duration from this model indicated the longest time-on-task for Cluster 2 and shortest for Cluster 3 (Cluster 1, 23.6 minutes, 95% PI [20.7, 26.9]; Cluster 2, 29.5 minutes, 95% PI [25.32, 34.05]; Cluster 3 20.0 minutes, 95% PI [16.8, 23.8]). We found evidence for differences between Cluster 1 and Cluster 2, and between Cluster 2 and Cluster 3, but not between Clusters 1 and 3 [BFs 9.8, > 100, and 1.4 respectively].

## 4. Discussion

This study aimed to establish evidence for "writing profiles"—the claim that writers carry with them a habitual approach to writing that differs from that of some other writers and is to some important extent stable across writing tasks, or at least tasks that require the production of similar kinds of texts. With some caveats, our findings are consistent with this claim.

The writing behaviors that we observed in our study fell into three groups: A pattern characterized by frequent lookbacks but infrequent revisions (Cluster 1); a particularly non-linear pattern, characterized by writers' tendency to compost away from the front edge of their text, and to both review and revise their text regularly (Cluster 2); and a pattern that was characterized by minimal hesitation, lookback, and revision (Cluster 3). These clusters map roughly onto, respectively, the *mixed-strategy writers*, *revisers*, and *planners* identified by Torrance et al. (1994) on the basis of student self-report data, and more clearly onto *average*, *fragmentary*, and *non-stop* ways of writing identified from behavioral data by Van Waes and Schellens (2003).

Our focus was on the extent to which these clusters of writing processes can be claimed to represent behavioral patterns that are stable within writer, across tasks. We found, across three same-genre writing tasks, that 53% of writers showed the same broad pattern of behavior throughout. All of the remaining writers behaved similarly on for two of the three tasks. No writers wrote differently for all three tasks. This consistency across tasks is substantially above chance. It is somewhat lower than the stability found in studies by Torrance et al. (1999; 2000), using self-report methods, and shows more consistency than Van Waes and Schellens (2003) based on direct behavioral measures. Comparison across studies is made difficult, however, by differences in the nature and the number of the process clusters into which writer behavior was classified. Van Waes and Schellens identified five different clusters, for example, which increases the probability of movement between clusters by chance. Conversely, in our study, participants completed three rather than just two tasks, substantially decreasing the by-chance probability of a participant's processes falling into the same cluster for all tasks.

Our approach contrasts with some of the early research on writing profiles cited in our introduction in which the focus was on writers' general attitudes toward and beliefs about their text production (Hartley & Branthwaite, 1989; Lavelle & Bushrow, 2007; Lavelle et al., 2002). Writing profiles based on writers' metacognition and metalanguage—how they explicitly think and talk about the strategies for producing text and about the text that they produce—are likely to be much more stable across tasks. How these meta-representations map onto a writer's moment-by-moment actions is, however, not straightforward.

Additional evidence for the stability of these process clusters within writers comes from our analysis of the cases where writers switched processes. Writers whose processes were less typical of the cluster to which they were allocated on a particular task (i.e., were more distant from the prototypical member of that cluster) were more likely to move to a different cluster when completing the next task. This suggests that, in at least some cases of apparent process change, movement to another cluster was more about reclassification of boundary cases than a clear switch to a substantially different process.

Our findings therefore point toward a fair degree of stability across same-genre tasks in the extent to which a writer pauses, reads or looks back at what they have written, makes changes to what they have already written, and composes away from the front edge of their text. We anticipate two possible challenges to this conclusion. First, the tasks that participants completed in our study were different in the sense that writers composed texts on different topics—with therefore some degree of individual differences across tasks in the writer's topic knowledge—but otherwise task demands were identical: In all three cases, participants wrote three-paragraph essays on general-knowledge topics. It is unclear to us whether previous authors who have described "writing profiles" would anticipate process stability across substantially divergent tasks. The existing literature does not present a strong theory about why writers carry similar behavioral patterns with them from task to task. An implicit underlying assumption is that writers over time find patterns of behaving when writing that are effective for them (or, more accurately, that the writer perceives as effective), given the kinds of writing tasks that they tend to perform. This adaptive behavior will necessarily be relatively task-specific. We would, therefore, expect whole-group, cross-task differences had we varied genre (Van Steendam et al., 2022). It also may be the case that variation across writer is greater for some text types than for others. It may or may not be that writers behave in similar ways across different text types, or that a writer who tends to compose in a particular way in one genre changes their behavior in predictable ways when writing in another genre (see, for example, Mohsen, 2021).

Note, however, that in the context of our present study, keeping genre constant across tasks was a necessary part of testing the hypothesis that writers have their own, stable writing profiles. To demonstrate this, it is necessary to show first that writers differ in important ways in how they perform the same task, and then that these within-writer differences are stable across similar tasks (i.e., this was not just down to the vagaries of how they performed on the day of the first task). Varying task type would have made it impossible to separate out variation in writing process due to the individual writer and variation in writing process resulting from varying type of task. Testing the additional effects of task-type would be possible by, for example, including an additional three tasks in a different genre. This might be a useful focus for future research.

Additionally, we note that our study used college-level writers. While this aligns with several related previous studies (e.g., Levy & Ransdell, 1995; Torrance et al., 1999, 2000), we acknowledge that our findings may not generalize to more experienced writers (e.g., professional writers).

These caveats aside, our findings align with some consistency in behavior within writer and across tasks. One interpretation of this consistency is that it is driven by a writer having developed over time a particular approach to writing that they then apply to each new task, either as an explicit strategy that they believe works for them, or just because through multiple uses their approach has become habitual. There are alternative explanations, however.

One possibility is that instead of reflecting habitual, adaptive writing processes, behavioral consistency results from how motivated the writer is to perform the task well. In the context of the present study, given the nature and similarity of our tasks, this seems quite plausible. Cluster 3 processes, for example, were very linear, with much less reviewing or revising than in other clusters. This resulted in total writing times that were only two-thirds of the duration of those in Cluster 2. It is possible that, in the context of the present study, writers whose behavior followed this pattern were less motivated. Writers in our study were paid for their participation and were free to spend as much time on the task as they wished. It may have been that writers in Cluster 3, and to a lesser extent in Cluster 1, just wanted to get the task over with as quickly as possible and were less concerned with the quality of the text that they produced.

We do not believe that this explanation is consistent with our data. We found no evidence that the differences in writing processes that were captured by our three clusters were associated with differences in the quality of the resulting text: We found no reliable differences in holistic rating of text quality ratings among clusters, nor in text length and ideational richness (as indexed by open-class to closed-class ratio). Therefore, although we would certainly expect writing behavior to be affected by writer motivation, this does not appear to provide an explanation for the behavioral stability that we observed in the present study.

The finding that writers, at least at the college level, carry with them behavioral patterns that show a fair degree of stability across tasks (or at least with tasks of a similar genre) has implications for writing instruction and support. Most fundamentally it means that interventions that target writing behavior have potential for making lasting change to the way that the writer behaves when performing similar tasks in the future (see, for example, Dux Speltz et al., 2022; Vandermeulen et al., 2023, 2024, for successful behavior-changing interventions). If it was the case that patterns in writing behavior, described at the level of this and previous studies (Kim, 2020; Levy & Ransdell, 1995, Van Waes & Schellens, 2003), were never stable across tasks, then interventions focused on changing these behaviors could not succeed.

## 5. Conclusion

The idea of writing "profiles"—that there are different types of writers, characterized by the actions they perform when producing text—has intuitive appeal and some history in research on written production. Our study tested this idea directly, evaluating two core hypotheses: (a) that writers differ meaningfully in their composing behavior even when completing the same task to a similar standard, and (b) that these behaviors are, to some extent, habitual and carried with the writer across multiple same-genre writing tasks. With some caveats, our findings support both hypotheses.

We observed robust individual differences in moment-by-moment writing behaviors, including the frequency and timing of pausing, revision, reviewing, and non-linear composition. These differences grouped into three process clusters that described quite different approaches to completing a composition task. We found no evidence that these differed in effectiveness. Over half of participants remained in the same cluster across all tasks, and the rest remained within two clusters, with transitions primarily occurring near cluster boundaries.

We conclude, therefore, that many writers do tend to indeed carry a consistent pattern of writing behavior across similar tasks. Further work is needed to understand how stable these patterns remain across task types and are affected by individual factors such as motivation or topic knowledge. Our present findings do, however, provide evidence that writers typically develop habitual approaches to composition that can be observed and have the potential to be supported or refined.

## Author Note

This material is based upon work supported by the National Science Foundation under Grant Nos. 2016868 and 2302644. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. This study was reviewed and considered exempt from regulatory oversight by the university's institutional review board. See https://rpubs.com/endux/stability for R code. Additional anonymized data files are available at https://osf.io/2z3t6/.

## References

Alamargot, D., Plane, S., Lambert, E., & Chesnet, D. (2010). Using eye and pen movements to trace the development of writing expertise: Case studies of a 7th, 9th and 12th grader, graduate student, and professional writer. *Reading and Writing, 23*, 853–888. https://doi.org/10.1007/s11145-009-9191-9

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

Berninger, V. W., & Winn, W. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (pp. 96–114). Guildford Press.

Вьrkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395. https://doi.org/10.32614/RJ-2018-017

Buschenhenke, F., Conijn, R., & Van Waes, L. (2023). Measuring non-linearity of multi-session writing processes. *Reading and Writing*. https://doi.org/10.1007/s11145-023-10449-9

Chandler, D. (1992). The phenomenology of writing by hand. *Intelligent Tutoring Media*, *3*(2-3), 65–74. https://doi.org/10.1080/14626269209408310

Chukharev-Hudilainen, E., Saricaoglu, A., Torrance, M., & Feng, H.-H. (2019). Combined deployable keystroke logging and eyetracking for investigating L2 writing fluency. *Studies in Second Language Acquisition*, *41*(3), 583–604. https://doi.org/10.1017/S027226311900007X

Conijn, R., Dux Speltz, E., & Chukharev-Hudilainen, E. (2021). Automated extraction of revision events from keystroke data. *Reading and Writing, 37*, 483–508. https://doi.org/10.1007/s11145-021-10222-w

Conijn, R., Cook, C., van Zaanen, M., & Van Waes, L. (2022). Early prediction of writing quality using keystroke logging. *International Journal of Artificial Intelligence in Education*, *32*(4), 835–866. https://doi.org/10.1007/s40593-021-00268-w

Crossley, S. A., Roscoe, R., & McNamara, D. S. (2014). What is successful writing? An investigation into the multiple ways writers can write successful essays. *Written Communication*, *31*(2), 184–214. https://doi.org/10.1177/0741088314526354

De La Paz, S., & Sherman, C. K. (2013). Revising strategy instruction in inclusive settings: Effects for English learners and novice writers. *Learning Disabilities Research and Practice, 28*(3), 129–141. https://doi.org/10.1111/ldrp.12011

Delgado-Osorio, X., Hartig, J., Harsch, C., & Koval, V. (2025). Students' behavioral patterns in integrated writing tasks: A sequence analysis approach. Journal of Educational Psychology, 117(6), 898–917. https://doi.org/10.1037/edu0000957

Ding, C. (2009). Dimension reduction techniques for clustering. In *Encyclopedia of database systems* (pp. 846-846). Springer, Boston, MA. https://doi.org/10.1007/978-0-387-39940-9_612

Dux Speltz, E., Roeser, J., & Chukharev-Hudilainen, E. (2022). Automating individualized, process-focused writing instruction: A design-based research study. *Frontiers in Communication*, *7*, 933878. https://doi.org/10.3389/fcomm.2022.933878

Faigley, L., & Witte, S. (1981). Analyzing Revision. *College Composition and Communication, 32*(4), 400–414. https://doi.org/10.2307/356602

Flower, L. S., & Hayes, J. R. (1981). The pregnant pause: An enquiry into the nature of planning. *Research in the Teaching of English, 15*(3), 229–243. https://doi.org/10.58680/rte198115763

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Chapman; Hall/CRC. https://doi.org/10.1201/b16018

Graham, S. (1997). Executive control in the revising of students with learning and writing difficulties. *Journal of Educational Psychology, 89*(2), 223–234. https://doi.org/10.1037/0022-0663.89.2.223

Hansen, R. (1986). John Irving, The Art of Fiction No. 93. *The Paris Review, 100*. https://www.theparisreview.org/interviews/2757/the-art-of-fiction-no-93-john-irving

Hartley, J., & Branthwaite, A. (1989). The psychologist as wordsmith: a questionnaire study of the writing strategies of productive British psychologists. *Higher Education*, *18*(4), 423–452. https://doi.org/10.1007/BF00140748

Harris, K.R., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation*. Cambridge, MA: Brookline Books.

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–26). Erlbaum.

Hayes, J., & Flower, L. (1980). Identifying the organisation of writing processes. In L. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Erlbaum.

Kim, H. (2020). Profiles of undergraduate student writers: Differences in writing strategy and impacts on text quality. *Learning and Individual Differences, 78*, 101823. https://doi.org/10.1016/j.lindif.2020.101823

Krippendorff, K. (1970). Bivariate agreement coefficients for reliability data. In E. R. Borgatta & G. W. Bohrnstedt (Eds.), *Sociological methodology* (pp. 139–150). Jossey Bass.

Kuehl, L. (1978). Joan Didion, The Art of Fiction No. 71. *The Paris Review, 74*. https://www.theparisreview.org/interviews/3439/the-art-of-fiction-no-71-joan-didion

Lavelle, E. (1993). Development and validation of an inventory to assess processes in college composition. *The British Journal of Educational Psychology*, *63*(3), 489–499. https://doi.org/10.1111/j.2044-8279.1993.tb01073.x

Lavelle, E., & Bushrow, K. (2007). Writing approaches of graduate students. *Educational Psychology, 27*(6), 807–822. https://doi.org/10.1080/01443410701366001

Lavelle, E., Smith, J., & O'Ryan, L. (2002). The writing approaches of secondary students. *The British Journal of Educational Psychology, 72*(3), 399–418. https://doi.org/10.1348/000709902320634564

Levy, C. M., & Ransdell, S. (1995). Is writing as difficult as it seems? *Memory & Cognition*, *23*(6), 767–779. https://doi.org/10.3758/BF03200928

Levy, C. M., & Ransdell, S. E. (1996). Writing signatures. In C. M. Levy & S. E. Ransdell (Eds.), *The Science of Writing* (1st Edition, pp. 149–162). Routledge. https://doi.org/10.4324/9780203811122

Lo Sardo, D. R., Gravino, P., Cuskley, C., & Loreto, V. (2023). Exploitation and exploration in text evolution. Quantifying planning and translation flows during writing. *PLOS ONE, 18*(3), e0283628. https://doi.org/10.1371/journal.pone.0283628

McCarthy, P. M. (2005). *An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity (MTLD)* [Doctoral dissertation, The University of Memphis].

Merkle, E. C., Fitzsimmons, E., Uanhoro, J., & Goodrich, B. (2021). Efficient Bayesian structural equation modeling in Stan. In *Journal of Statistical Software* (Vol. 100, Issue 6, pp. 1–22). https://doi.org/10.18637/jss.v100.i06

Mohsen, M. A. (2021). Second Language Learners' Pauses over Different Times Intervals in L2 Writing Essays: Evidence from a Keystroke Logging Program. *Psycholinguistics, 30*(1), 180–202. https://doi.org/10.31470/2309-1797-2021-30-1-180-202

Morris, M. (1990). Margaret Atwood, The Art of Fiction No. 121. *The Paris Review, 117*. https://www.theparisreview.org/interviews/2262/the-art-of-fiction-no-121-margaret-atwood

Olive, T., Alves, R. A., & Castro, S. L. (2009). Cognitive processes in writing during pause and execution periods. *European Journal of Cognitive Psychology, 21*(5), 758–785. https://doi.org/10.1080/09541440802079850

Roeser, J., De Maeyer, S., Leijten, M., & Van Waes, L. (2024). Modelling typing disfluencies as finite mixture process. *Reading and Writing*, *37*, 359–384. https://doi.org/10.1007/s11145-021-10203-z

Roeser, J., Conijn, R., Chukharev, E., Ofstad, G. H., & Torrance, M. (2025). Typing in tandem: Language planning in multisentence text production is fundamentally parallel. *Journal of Experimental Psychology: General, 154*(7), 1824–1854. https://doi.org/10.1037/xge0001759

Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflective processes in written composition. *Cognitive Science, 8*, 173–190. https://doi.org/10.1207/s15516709cog0802_4

Shapiro, K. (1948). The Meaning of the Discarded Poem. In Arnheim, R Auden, WH Shapiro, K Stauffer, DA (Ed.), *Poets at work*. Oxford University Press.

Sivula, T., Magnusson, M., Matamoros, A. A., & Vehtari, A. (2020). Uncertainty in Bayesian leave-one-out cross-validation based model comparison. In *arXiv [stat.ME]*. arXiv. http://arxiv.org/abs/2008.10296

Spender, S. (1952). The making of a poem. In B. Ghiselin (Ed.), *The Creative Process*. Greenwood Publishing Group.

Torrance, M., & Conijn, R. (2023). Methods for studying the writing time-course. *Reading and Writing*. https://doi.org/10.1007/s11145-023-10490-8

Torrance, M., Fidalgo, R., & Robledo, P. (2015). Do sixth-grade writers need process strategies? *British Journal of Educational Psychology, 85*(1), 91–112. https://doi.org/10.1111/bjep.12065

Torrance, M., Johansson, R., Johansson, V., & Wengelin, E. (2016). Reading during the composition of multi-sentence texts: an eye-movement study. *Psychological Research, 80*(5), 729–743. https://doi.org/10.1007/s00426-015-0683-8

Torrance, M., Thomas, G. V., & Robinson, E. J. (1994). The writing strategies of graduate research students in the social sciences. *Higher Education*, *27*(3), 379–392. https://doi.org/10.1007/BF03179901

Torrance, M., Thomas, G. V., & Robinson, E. J. (1999). Individual differences in the writing behavior of undergraduate students. *The British Journal of Educational Psychology*, *69*(2), 189–199. https://doi.org/10.1348/000709999157662

Torrance, M., Thomas, G. V., & Robinson, E. J. (2000). Individual differences in undergraduate essay-writing strategies: A longitudinal study. *Higher Education*, *39*(2), 181–200. https://doi.org/10.1023/A:1003990432398

Torruella, J., & Capsada, R. (2013). Lexical statistics and tipological structures: A measure of lexical richness. *Procedia-Social and Behavioral Sciences*, *95*, 447–454. https://doi.org/10.1016/j.sbspro.2013.10.668

Vandermeulen, N., Van Steendam, E., De Maeyer, S., Lesterhuis, M., & Rijlaarsdam, G. (2024). Learning to write syntheses: the effect of process feedback and of observing models on performance and process behaviors. *Reading and Writing, 37*(6), 1375-1405. https://doi.org/10.1007/s11145-023-10483-7

Vandermeulen, N., Van Steendam, E., De Maeyer, S., & Rijlaarsdam, G. (2023). Writing process feedback based on keystroke logging and comparison with exemplars: Effects on the quality and process of synthesis texts. *Written Communication*, *40*(1), 90–144. https://doi.org/10.1177/07410883221127998

van Hell, J. G., Verhoeven, L., & van Beijsterveldt, L. M. (2008). Pause Time Patterns in Writing Narrative and Expository Texts by Children and Adults. *Discourse Processes, 45*(4–5), 406–427. https://doi.org/10.1080/01638530802070080

Van Steendam, E., Vandermeulen, N., De Maeyer, S., Lesterhuis, M., Van den Bergh, H., & Rijlaarsdam, G. (2022). How students perform synthesis tasks: An empirical study into dynamic process configurations. *Journal of Educational Psychology, 114*(8), 1773–1800. https://doi.org/10.1037/edu0000755

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*(6), 829–853. https://doi.org/10.1016/S0378-2166(02)00121-2

Vehtari, A., Simpson, D., Gelman, A., Yao, Y., & Gabry, J. (2024). Pareto smoothed importance sampling. *Journal of Machine Learning Research, 25*(72), 1–58. http://jmlr.org/papers/v25/19-556.html

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing, 27*, 1413-1432. https://doi.org/10.1007/s11222-016-9696-4

Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage--Dickey method. *Cognitive Psychology*, *60*(3), 158–189. https://doi.org/10.1016/j.cogpsych.2009.12.001

Wellek, R., & Warren, A. (1963). *Theory of Literature*. Penguin.

Wengelin, E., 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, 41*(2), 337–351. https://doi.org/10.3758/BRM.41.2.337

*Wyllie, A. (1993). On the Road to Discovery: A Study of the Composing Strategies of Academic Writers on the Word Processor [Doctoral dissertation, University of Lancaster].*

## Appendix A

*Table A1.* Descriptive statistics for behavioral measures

| Measure | By-cluster means (with SD) | | |
|---|---|---|---|
| | Cluster 1 | Cluster 2 | Cluster 3 |
| 1. Total number of major blocks | 1.5 (1.2) | 1.9 (1.5) | 0.79 (0.54) |
| 2. Total number of jumps | 13 (14) | 26 (15) | 4.9 (3.6) |
| 3. Total number of edges | 10 (8.4) | 18 (10) | 4.6 (3) |
| 4. Number of sustained reading events | 35 (15) | 54 (21) | 10 (7.8) |
| 5. Percentage of hesitations before sentence | 12 (6.3) | 16 (6) | 9.9 (5) |
| 6. Percentage of hesitations before word | 12 (3.8) | 21 (4.9) | 11 (3.1) |
| 7. Percentage of hesitations within word | 11 (3) | 19 (4.4) | 10 (2.3) |
| 8. Percentage of lookbacks before sentence | 14 (6.6) | 12 (6.5) | 4.1 (2.8) |
| 9. Percentage of lookbacks before word | 3.7 (2.4) | 3.9 (2.7) | 1.1 (0.82) |
| 10. Percentage of lookbacks within word | 0.56 (0.26) | 0.68 (0.42) | 0.16 (0.086) |
| 11. Length of non-lookback sequence (keystrokes) | 78 (32) | 89 (86) | 320 (180) |
| 12. Length of production sequence (keystrokes) | 29 (9.2) | 17 (4.1) | 27 (10) |
| 13. Lookback duration (seconds) | 6.7 (0.18) | 6.7 (0.12) | 6.6 (0.19) |
| 14. Percentage of long lookbacks | 54 (18) | 52 (17) | 39 (16) |
| 15. Percentage of hesitations before sentence | 65 (13) | 66 (11) | 57 (18) |
| 16. Percentage of hesitations before word | 30 (13) | 34 (11) | 31 (26) |
| 17. Percentage of hesitations within word | 8.7 (4.8) | 10 (4.1) | 7.4 (2.5) |

*Table A2.* Correlations for behavioral measures

| Measure | Correlations | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 1. Total number of major blocks | | | | | | | | | | | | | | | | |
| 2. Total number of jumps | 0.19 | | | | | | | | | | | | | | | |
| 3. Total number of edges | 0.35 | 0.94 | | | | | | | | | | | | | | |
| 4. Number of sustained reading events | 0.49 | 0.62 | 0.61 | | | | | | | | | | | | | |
| 5. Percentage of hesitations before sentence | 0 | 0.47 | 0.44 | 0.45 | | | | | | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 6. Percentage of hesitations before word | 0.2 | 0.5 | 0.47 | 0.58 | 0.51 | | | |
| 7. Percentage of hesitations within word | 0.1 | 0.5 | 0.44 | 0.54 | 0.43 | 0.93 | | |
| 8. Percentage of lookbacks before sentence | 0.28 | -0.05 | -0.02 | 0.3 | -0.08 | -0.12 | -0.08 | |
| 9. Percentage of lookbacks before word | 0.18 | -0.12 | -0.1 | 0.28 | 0.02 | 0.01 | 0.04 | 0.69 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10. Percentage of lookbacks within word | 0.21 | 0.04 | 0.04 | 0.3 | 0.02 | 0.06 | 0.11 | 0.55 | 0.72 | | | | |
| 11. Length of non-lookback sequence (keystrokes) | -0.24 | -0.18 | -0.18 | -0.46 | -0.18 | -0.12 | -0.14 | -0.58 | -0.58 | -0.59 | | | |
| 12. Length of production sequence (keystrokes) | 0.13 | -0.36 | -0.34 | -0.21 | -0.34 | -0.65 | -0.68 | 0.13 | 0.01 | -0.01 | 0.02 | | |
| 13. Lookback duration (seconds) | 0.16 | 0.04 | 0.07 | 0.17 | 0.12 | 0.13 | 0.08 | 0.3 | 0.16 | 0.16 | -0.24 | -0.11 | |
| 14. Percentage of long lookbacks | 0.14 | 0.44 | 0.41 | 0.43 | 0.09 | 0.03 | 0 | 0.14 | -0.1 | 0.04 | -0.18 | 0.04 | 0.17 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15. Percentage of hesitations before sentence | 0.29 | 0 | 0.01 | 0.25 | 0.08 | 0.17 | 0.08 | 0.27 | 0.31 | 0.24 | -0.29 | 0.13 | 0 | -0.12 | | |
| 16. Percentage of hesitations before word | -0.13 | 0.02 | -0.02 | 0.1 | 0.16 | 0.2 | 0.25 | 0.01 | 0.1 | 0.23 | -0.01 | -0.3 | -0.12 | 0.01 | 0.21 | |
| 17. Percentage of hesitations within word | 0.08 | 0.16 | 0.2 | 0.08 | 0.04 | 0.17 | 0.16 | 0.11 | 0.21 | 0.43 | -0.21 | -0.22 | 0.1 | 0.05 | 0.16 | 0.28 |