EasyChair Preprint

№ 8453

Access to Real-Time Typing Shapes Perception of a Collaborator's Work Quality

Arielle Elliott and William Horton

July 11, 2022

PREPRINT

Access to Real-time Typing Shapes Perception of a Collaborator's Work Quality

Arielle V. Elliott and William S. Horton

Department of Psychology

Northwestern University

Author Note

The authors declare that there are no conflicts of interest with respect to this preprint.

Correspondence should be addressed to Arielle Elliott: a.v.elliott@u.northwestern.edu

Abstract

Online shared workspaces provide real-time access to others' work as it is being written. Given that typing patterns are shaped by stress and cognitive load, access to an online partner's real-time typing behaviors may also inform perceptions of another's epistemic state more broadly, influencing judgments of their task contributions. Participants completed a joint editing task with a "partner" whose edits were pre-recorded to be delivered fluently or disfluently. Participants then rated the partner's edited sentences. Our results show that visible typing dynamics in an online workspace can influence perception of writing quality.

Access to Real-time Typing Shapes Perception of Collaborator's Work Quality

In face-to-face interactions, people routinely draw epistemic inferences about an interlocutor's competence or certainty about what is being said. Such inferences may be influenced not only by *what* that person says (content), but also *how* they say it (delivery). Although face-to-face contexts routinely provide access to paralinguistic cues like facial expressions or speech disfluencies, text-based contexts typically allow only limited delivery-based cues, such as the relative timing of production (Kalman, Scissors, & Gergle, 2013). However, in shared, fully synchronous workspaces like Google Docs, users have real-time access to the dynamics of one others' written contributions, permitting access to other kinds of potentially informative cues to epistemic states.

We are interested in how visible typing behaviors in Google Doc-like environments facilitate certain kinds of epistemic inferences. Shifts in typing patterns have been reliably shown to be indicators of different mental states, such as cognitive load or mood changes (Epp, Liphold, & Mandryk, 2011). As such, typing dynamics could potentially function as a rich paralinguistic cue about the partner that is particularly available in synchronous collaborative writing. In previous work, we found that viewers can make explicit judgments about a user's certainty and task familiarity from video clips of that person's typing (Elliott & Horton, in prep). In the present study, we investigated how having access to a partner's real-time typing behavior in a collaborative environment would shape judgments about the quality of that partner's contributions to the task. Participants completed an online collaborative editing task with a "partner" whose task contributions were pre-scripted to be delivered fluently or disfluently. The participant then independently judged the quality of their partner's edits. Our results show that

observable differences in typing behavior can contribute to perceptions of how well a partner is managing the task, which has implications for interfaces that enable real-time text-based collaboration.

Method

Participants. We recruited 38 college students for a collaborative editing task that was carried out online. All participants were native speakers of English and completed the study for partial course credit.

Materials. For the editing task, we constructed four paragraphs, each ten sentences long, based on items taken from an online SAT preparation website. To motivate the need for editing, the sentences in each original paragraph were rewritten to contain errors that were sometimes grammatical, but more often stylistic and structural in nature. Some example sentences are shown in Table 1.

Table 1. *Sample sentences assigned to the "partner" for potential revision.*

(Line 1) Globalization, or the integration of cultures across nations, can be to some, a prominent yet controversial topic.
(Line 2) Some defend globalization for its benefits greater creativity and appreciation of heritage.
(Line 5) Now, many nations gets to enjoy booming domestic film markets that can even compare with or surpass the United States in production and popularity.
(Line 10) Certainly, many individuals prefer films that reflect their own cultural identities, but trends indicate rising popular interest, even in Hollywood, in multicultural and cross-cultural movies.

From these items, we created a second set of revised sentences designed to be the partner's real-time contributions to the editing task. To examine the impact of visible typing behavior over a range of contribution "quality," we constructed these revisions to be either *sufficient* or *insufficient*. Sufficient revisions were intended to be generally satisfactory edits of the target sentences, although the occasional flaw was allowed to remain for verisimilitude. For the insufficient revisions, we corrected at least one error in the original but often inappropriately modified one or more remaining errors (e.g., by shifting word order in a way that did not improve the sentence structure). These insufficient revisions were intended to represent stylistically or grammatically unacceptable changes to the target sentences. The number of sufficient/insufficient partner revisions alternated between 2 or 3 sentences of each type across paragraphs, ultimately totaling ten sufficient revisions and ten insufficient revisions across all four paragraphs. Examples of the final scripted revisions are provided in Table 2.

Table 2. *Sample scripted revisions produced by the "partner."*

Sufficient	Insufficient
Globalization, or the integration of cultures across nations, has become an increasingly prominent yet controversial topic. (Line 1)	Due to it's benefits of greater creativity and appreciation of heritage, some defend globalization. (Line 2)
Despite many individual's preference in films that reflect their own cultural identities, trends indicate rising popular interest in cross-cultural movies. (Line 10)	The domestic film markets in many nations are now booming and can even compare with or surpass the United States in production and popularity. (Line 5)

For each scripted revision sentence, we created two prerecorded typing "clips" to mimic the partner's contributions during the editing task: a fluent version and a disfluent version. These typing clips were created using Mouse Recorder software (https://www.mouserecorder.com), which records keystrokes that later can be played back exactly as recorded on any area that has been clicked on. Thus, the timing between keystrokes, as well as any backspaces and changed content is preserved when the recording is replayed. When viewed within the shared document used for the editing task, the replay of these recordings was indistinguishable from real typing.

The mouse recorder additionally logs all keystrokes, allowing us to quantify differences in fluency across revision versions. Based on the keypress data obtained in Elliott and Horton (in prep), we created fluent versions of the scripted revisions to have average raw keypress speeds below 0.3 seconds, no more than two pauses over 200 ms, and at most one or two low-level typing errors (e.g., transposed characters that are immediately corrected). The fluent recordings also omitted any significant edits that included over three repeated backspaces or keypresses. In contrast, we created disfluent recordings to have an average keypress speed ranging from 0.3 to 0.5 seconds between keystrokes, and one to five low-level typing errors and also included at least two significant content edits (i.e., backspace events with over three consecutive keypresses).

Procedure. Each participant was told they would be working with another student to edit a series of four short paragraphs within a shared Google Doc, and that each paragraph would potentially contain a number of grammatical and stylistic errors. Their task was to improve each paragraph by making revisions as needed. For each paragraph, each partner was assigned specific color-coded sentences to edit, and the edits made by each partner were always visible to the other person as they happened. In actuality, participants worked alone, and the contributions of the partner were conveyed through the prerecorded keystroke files entered into the Google Doc by

the experimenter, which made it seem as if the partner was simultaneously engaged in the editing task. For half the participants, these recordings showed the partner making their revisions *fluently* (i.e., relatively fast with few pauses or errors), while other participants saw the partner's edits being produced *disfluently*. Importantly, the actual content of the final edited sentences was always held constant.

Once the revisions were complete for each paragraph, participants then clicked on a link that took them to another document where they were asked to independently judge the quality of their partner's revisions to the previous paragraph. For these judgments, participants were asked to "consider the overall readability, clarity, and grammatical correctness" of each of the revised sentences produced by the partner (which, again, were always pre-scripted and identical across conditions). Participants made this judgment on a 0-100 scale, where "0" was "fundamental errors and stylistic issues remain; needs complete rewriting" and "100" was "free of grammatical or stylistic errors, no further editing needed." Participants were led to believe that their partner was separately rating the quality of the participants' revisions as well. After providing these ratings, participants were then instructed to return to the main shared Google Doc to continue the collaborative editing task for the next paragraph. The collaborative editing and independent rating tasks alternated until all four paragraphs were complete.

Results

Table 3 presents the average partner revision quality ratings computed across both revision sufficiency and partner fluency. To analyze these data, we fit a linear mixed-effects model in R using *lmer*. Confirming the success of our manipulation of revision sufficiency, participants gave significantly higher ratings for sufficient revisions (M=83.52; SD=16.86) than

insufficient revisions (*M*=72.05; SD=21.91), *b*=11.47, *p*<.03. More importantly, though, there was also a significant main effect of partner fluency. Participants gave significantly higher ratings for revisions produced by a fluent partner (*M*=81.08; SD=20.97) than for revisions produced by a disfluent partner (*M*=74.49; SD=19.19), *b*=6.59, *p*<.03. However, these two factors did not interact, *b*=4.38, *p*=.10.

Table 3. *Average ratings of the partner's revision quality (0-100 scale), by partner editing fluency and revision sufficiency (SDs in parentheses)*

	Revision Sufficiency	
Partner Fluency	Sufficient	Insufficient
Disfluent	79.13 (16.2)	69.85 (20.8)
Fluent	87.92 (16.4)	74.25 (22.8)

Discussion

In a collaborative editing task, participants were given the opportunity to observe the real-time contributions of an online partner, who appeared to engage with the task in a relatively fluent or disfluent way. Subsequently, when rating the overall quality of the partner's edits, participants gave higher ratings for revisions produced by the fluent partner, even though the final content of these revisions was kept identical across both fluent and disfluent partners. We believe that being able to observe the dynamics of the partner's typing allowed participants to

make epistemic inferences about the partner's overall competence in carrying out the task, which influenced their perceptions of the quality of this partner's contributions.

As remote online platforms increasingly become the standard for collaborative work, it will be important to develop a more complete understanding of the cognitive and social mechanisms behind communication within these contexts. In order to adapt concepts that have been well addressed in the literature on face-to-face collaboration, we need an understanding of the factors that influence communicative behaviors when many of the typical 'tools' of face-to-face communication are removed. Without access to gesture, facial expression, tone of voice, or physical surroundings, users must rely on other sorts of visible cues to coordinate effective collaboration. Here, we have begun to examine how typing patterns can be used to inform representations of an online partner. This sets the stage for exploring how similar behaviors may shape how people interact in real-time collaborative writing contexts.

References

Elliott, A. V., & Horton, W. S. (in preparation). Typing speed and fluency as cues to uncertainty in the real-time production of written messages.

Epp, C., Liphold, M., & Mandryk, R. L. (2011). Identifying emotional states using keystroke dynamics. In *Proceedings of the SIGCHI Conference of Human Factors in Computing Systems* (CHI '11) (pp. 715-624). https://doi.org/10.1145/1978942.1979046

Kalman, Y., Scissors, L. E., & Gergle, D. (2013). Online chronemics convey social information. *Computers in Human Behavior*, *29*, 1260-1269. https://doi:10.1016/j.chb.2012.12.036