Chain Reactions: The Impact of Order on Microtask Chains

Carrie J. Cai
MIT CSAIL
Cambridge, MA
cjcai@mit.edu

Shamsi T. Iqbal
Microsoft Research
Redmond, WA
shamsi@microsoft.com

Jaime Teevan
Microsoft Research
Redmond, WA
teevan@microsoft.com

ABSTRACT
Microtasks are small units of work designed to be completed individually, eventually contributing to a larger goal. Although microtasks can be performed in isolation, in practice people often complete a chain of microtasks within a single session. Through a series of crowd-based studies, we look at how various microtasks can be chained together to improve efficiency and minimize mental demand, focusing on the writing domain. We find that participants completed low-complexity microtasks faster when they were preceded by the same type of microtask, whereas they found high-complexity microtasks less mentally demanding when preceded by microtasks on the same content. Furthermore, participants were faster at starting high-complexity microtasks after completing lower-complexity microtasks, but completion time and quality were not affected. These findings provide insight into how microtasks can be ordered to optimize transitions from one microtask to another.

Author Keywords
Microtasks; crowdsourcing; selfsourcing.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

INTRODUCTION
Microtasks are small units of work that contribute progress towards a larger goal [11]. Recently, microtasks have been used to enable crowds of unknown workers to complete large tasks [7][15], but historically task decomposition has been used to support personal information management [4] and task sharing among collaborators [36][39]. For example, a person can organize a personal photo collection by completing a series of pairwise comparisons [45].

While individual microtasks require limited time to complete, in practice multiple microtasks are often completed one after another as part of a longer *task chain* within a single session. For example, there is evidence that language learners voluntarily fetch multiple flashcards during micro-moments of down time [10][16]. On crowd platforms, batches of tasks done in chains are not only prevalent, but also preferred by crowd workers because they leverage a worker's growing familiarity with the task [12].

The way that consecutive microtasks are chained affects the extent to which people complete tasks productively and continue to engage in the task at hand. Prior research suggests that task interruptions and delays can slow down performance [27][28]. Boredom and fatigue from completing long chains of tasks can also lead to under-performance and task abandonment [14][38]. However, past studies have demonstrated situations in which people continue doing tasks until a certain milestone has been reached, both on personal tasks [10] and crowd work [24].

A way to keep people engaged during a chain of microtasks is to order the tasks in a way that minimizes cognitive load, which could lead people to complete tasks more easily, more efficiently and with greater enjoyment. The set of microtasks performed by an individual can vary: multiple *operations* may need to be performed on a single piece of content (e.g., describing and categorizing an item), multiple pieces of *content* may require the same operation (e.g., transcribing audio segments), or operations may have different *complexity* levels (e.g., performing easier vs. harder search tasks). For example, an email-organizing application might present consecutive emails from the same thread to preserve continuity of the content, or it might instead group on the same operation (e.g., rate all emails, then categorize all emails). In crowd work, microtasks are often routed to workers in different orders [28], leading to potentially diverse experiences within a single session.

To understand task chaining, we focus on the writing domain. Not only is writing an important and common part of information work [21][25], but it also offers a particularly interesting case for effective task chaining because subtasks in writing vary widely in both content and complexity, from low-level proofreading to meaning-rich rephrasing and tone modification. Moreover, writing is a canonically difficult task to start doing [18][23], a hurdle that could potentially be addressed using microtasks.

We identify 11 common writing microtasks, and use crowd workers to evaluate the effect of chaining on microtask *continuity* (continuing microtask chains of the same complexity level), microtask *transitions* (transitioning across microtask complexity levels), and microtask *ease-in* (using simpler

CHI'16, May 07-12, 2016, San Jose, CA, USA
© 2016 ACM. ISBN 978-1-4503-3362-7/16/05⋯$15.00
DOI: http://dx.doi.org/10.1145/2858036.2858237

microtasks to ease people into more complex microtasks). Using operation, content, and complexity level as key properties in forming microtask chains, we find that microtasks have carry-over effects on subsequent microtasks within the same chain. Ordering affects 1) Continuity within the same complexity level: low-complexity microtasks chained on the same *operation* contribute to faster completion, while high-complexity microtasks chained on the same *content* are perceived to be less mentally demanding; 2) Transitions between complexity levels: microtasks are completed faster and are perceived to be better aided by preceding microtasks of the *same* complexity than by those of a *different* complexity; and 3) Easing in to complex tasks: people develop a sense of momentum and are faster to engage with a high-complexity microtask when it is preceded by lower-complexity microtasks. As more and more tasks are transformed into microtasks, these findings can be used to chain microtasks in a way that enhances performance and minimizes mental demands.

RELATED WORK

There is good evidence that the order in which tasks are done impacts people's ability to perform them. This is because transitions between consecutive cognitive tasks have measurable effects on ongoing mental processes. Psychology literature suggests that switching between different tasks results in slower and more error-prone performance than focusing on a single task, due to time spent re-configuring to a new task [32][47]. Beyond multi-tasking, transitions can also be interrupt-driven. Research shows that interruption from a secondary task harms subsequent performance on the primary task [5][19] and increases stress [30], even though people tend to underestimate the cost of these interruptions [27]. Complex tasks are vulnerable to interruptions [11] because people performing them build up a mental state that can be disrupted by the state of a secondary task [8].

Suboptimal transitions between tasks can occur not only between different activities or during an ongoing task, but also within a single task that has been broken into microtasks, depending on how the microtasks are chained. Lasecki et al. found that interjecting contextually relevant microtasks and delays can hurt worker performance [28]. While iterating on similar tasks can help people build familiarity and expertise [15][34], long chains of similar tasks can lead to boredom and fatigue [29][38].

Recent work has sought to improve people's experiences with microtasks by inserting micro-diversions to provide timely relief during long chains [14]. Large organizations have also explored re-designing assembly lines to build task specialization while still enabling task switching and creativity [3]. Others have aimed to mitigate the effects of task interruptions, by locating more optimal moments for interruptions [2][26]. On crowd platforms, priming effects [33] and monetary interventions [48] can also improve performance. In our approach, we examine how to optimize transitions through the ordering of existing microtasks.

The process of writing consists of uniquely demanding subtasks that require transforming abstract concepts into prose that someone else will understand [23]. Task ordering appears particularly critical in this domain; the process of editing text can lead a writer to develop new insights about the text [18], and the very process of engaging in a task can increase self-efficacy and the motivation to continue [6]. This suggests individual writing microtasks are not done in isolation, but rather are affected by a person's experience with neighboring microtasks. While recent work has explored ways to scaffold complex writing tasks by breaking them into subgoals [7][35][44] , in our work we assume an existing collection of microtasks and investigate the effects of ordering. For instance, simpler microtasks could be performed first, with the potential side effect of helping people ramp up to more difficult microtasks. Indeed, workflows that create short-term goals have been shown to help procrastinators by increasing the perceived likelihood of success [1][17].

In summary, prior research has shown that task transitions are important, even for microtasks within a single task. Writing, in particular, is a domain where thoughtful chaining seems likely to be important. The work in this paper builds on previous studies by examining how microtask chains can support continuity, help people transition between different microtasks, and allow them to ease in to the overall process.

PILOT INTERVIEWS

To understand the challenges people commonly face when transitioning between writing tasks and how they currently attempt to overcome these barriers, we interviewed 10 people about their writing process. We selected participants from a range of professions (3 graduate students, 2 corporate managers, 2 professors/lecturers, 1 lawyer, 1 IT specialist, 1 product manager) to elicit common themes across different work-based writing tasks. Participants were selected on the basis that they reported regularly writing long-form documents, including memos, legal contracts, project specifications, lecture notes, and research papers. We asked each participant how they transition into writing from other tasks, and how they transition between different processes while writing. Several key themes emerged.

Ramping up to high-complexity writing tasks requires substantial cognitive energy. Compared to other tasks such as email or programming, the production of prose was described to be more mentally taxing due to the seemingly vast space of phrasing possibilities, the lack of frequent, objective feedback, and the difficulty transforming abstract thoughts into a linear sequence of words. As a result, many reported difficulty getting started writing due to the large activation energy required: "The main difficulty is the start-up inertia. I have a pretty short attention span. If I get discouraged I'll move away from it quickly."

People frequently engage in a series of low-complexity tasks with verifiable results as a way of building momentum and making initial progress. Examples include proofreading, text formatting, and addressing previous to-dos. For some, these tasks help them build momentum while expending low effort: "It's an easy way to check things off the to-do list, doesn't require as much thinking." For others, low-complexity tasks also yield concrete progress because they are to some extent self-verifiable: "I do the lowest hanging fruit first. There's a right answer for whether this sentence is grammatically proper."

Simple tasks on nearby content help people transition to more complex writing tasks. Many people make light edits on text they have already written as a means of regaining familiarity with the text: "The nice thing about a little polishing is that it establishes the context but doesn't take a lot of energy." For others, starting with lighter editing tasks also helps build focus for more difficult tasks: "It takes less energy to get into editing, but in the process it gets me into the more focused mode. That progress helps me with parts where I do have to think a lot and have to put in some effort." Participants said that they typically work on the content that is closest to where they intend to start writing, as a way of getting into the mental state of that content.

Suboptimal task transitions can lead to fatigue or disengagement. Some interviewees expressed the need to transition quickly from simple editing to deeper-level writing, as too much time spent on simple tasks can lead to the indefinite postponement of more meaning-rich writing: "I don't want to get super tired because then I would write nothing new." Others were wary of making small edits immediately after writing new text, due to a fear of losing momentum: "It takes me a lot of energy to focus on the high level stuff, so if I get out of that zone, it's hard to bring me back in." Several people took breaks from complex tasks by doing easy tasks on another part of the document (e.g. font formatting) or elsewhere entirely (e.g. deleting a series of emails), as a way of still staying productive.

These initial interviews suggest that people who frequently write for work make use of a number of strategies to transition into and between writing tasks. Some of the people we spoke with performed low-level operations to build focus and momentum, while others started by editing nearby content as a way of transitioning smoothly into a more complex task. These insights helped us develop our research questions, described in the following section.

RESEARCH QUESTIONS

Our observations about the role of text content and task operation in easing transitions between low and high complexity tasks motivate the following research questions:

1) **Continuity:** How are performance and subjective experience of a microtask affected by whether preceding microtasks share the same *operation* or the same *content* as the current microtask?

2) **Transitions:** How are performance and subjective experience of a microtask affected by whether previous microtasks are of the *same* complexity or *different* complexity as the current microtask?

3) **Easing in:** How are performance and subjective experience of a *complex* microtask affected when it is preceded by *simpler* microtasks?

To address these questions, we begin by identifying a number of different writing microtasks that are of varying degrees of complexity. We then present the results of three crowd-based studies designed to explore these three questions using the microtasks we identified.

SELECTING WRITING MICROTASKS

The studies in this paper explore chains of microtasks. In each study, the microtasks may vary with respect to their:

1) **Operation:** There are a number of operations that could make up a writing microtask. We focus specifically on editing tasks that modify or build upon preexisting text.

2) **Content:** The preexisting text that the operations are applied to can vary. We identify a corpus of sentences for editing with similar style and reading level.

3) **Complexity:** Operations vary by complexity, depending on the level of involvement the task requires with the text and its meaning. We establish low, medium, and high-complexity groups of microtasks.

In this section, we describe how we selected the operations and content so as to control the amount of variation in each study, and how we measured microtask complexity.

Microtask Operations

For the studies presented in this paper, it was necessary to identify sets of easy and hard operations that were considerably different from each other in complexity, but similar within each set. We first selected a number of common operations (see Table 1) along basic rhetorical dimensions of writing: mechanics and semantics [40][43]. Tasks in the *mechanics* category include checking for technical errors such as spelling and punctuation. In contrast, *semantics* tasks involve more in-depth consideration of meaning, such as shortening or rephrasing a sentence. We also included several tasks that are often implicitly executed by the writer, but explicitly completed in many crowdsourcing workflows: Awkcheck (identify whether a sentence sounds awkward), WordChoice (provide a better word to replace a given word), and SelectBest (select the best word to replace an existing word). These may be meaning-rich, but still fast to complete when supported by a workflow. For example, the Find-Fix-Verify [7] workflow could be analogously expressed using Awkcheck (Find), WordChoice (Fix), and SelectBest (Verify). After several iterations, we identified a list of eleven microtasks, shown in Table 1.

Note that the operations we selected are just a subsample of the potential operations. Our goal was not to be exhaustive, but rather identify an interesting range for study, with sev-

	Operation	Example Prompt
Mechanics	**punctuationCheck ***	Fix the punctuation error in this sentence, if one exists.
	duplicateCheck *	Fix the spot that has the same word twice in a row, if one exists.
	capitalizationCheck *	Fix the capitalization error in this sentence, if one exists.
	spellCheck	Fix the spelling error in this sentence, if one exists.
Crowd	**awkCheck ***	How awkward does this sentence sound? (Rate on 5-point scale)
	selectBest *	Select the best re-wording of this sentence. The word to replace is in brackets.'
	wordChoice	Replace the word in brackets to improve this sentence.
Semantics	**changeTone ***	Rewrite this sentence in an emotional tone.
	paraphrase *	Paraphrase this sentence as if you were saying it to a five-year-old.
	shorten *	Without changing it meaning, shorten this sentence so that it is at most {2/3 of original length} words long.
	nextSent	Write a sentence that could make sense if it came after this sentence.

Table 1. Eleven microtasks commonly performed in writing. After we analyzed the complexity of these microtasks, those with asterisks were selected for final studies.

eral operations at similar complexity levels. Pilot studies suggested these were a reasonable set to pursue.

Microtask Content

The operations we identified were designed to be performed on sentences. We gathered sentences from CHI 2015 paper abstracts to use in our studies. We chose this source because abstracts are prominent and complex, condensing a body of work into a single paragraph. We targeted sentences that were comprehensible on their own but required considerable mental processing to understand. Starting with 3701 sentences, we removed those with technical terminology that might make the sentence difficult to understand without external knowledge. Of the remaining 3115 sentences, we kept only those that were relatively complex by calculating the Automated Readability Index [42] and selecting 300 sentences at around the 75^{th} percentile, with an Automated Readability Index of 18 – 23. Several example sentences are shown in Table 2.

Microtask Complexity

To identify the complexity of each operation applied to this content, we conducted a between-subject study on Amazon Mechanical Turk. We analyzed how microtasks compared along the following dimensions of complexity: semantic processing, time, mental demand, interest, meaningfulness, and open-endedness. We then used these results to select sets of high and low complexity microtasks for further study.

Example sentences from CHI 2015 paper abstracts
Selective exposure, the preferential seeking of confirmatory information, can potentially exacerbate fragmentation of online opinions and lead to biased decisions.
This paper describes a study of algorithmic living with Trace, a mobile mapping application that generates walking routes based on digital sketches people create and annotate without a map.
We found that collective action publics tread a precariously narrow path between the twin perils of stalling and friction, balancing with each step between losing momentum and flaring into acrimony.

Table 2. Examples from the 300-sentence corpus used in our studies. Sentences in the corpus are at the 75^{th} percentile of reading difficulty among CHI 2015 paper abstracts, determined using the Automated Readability Index.

Complexity Metrics

We examine *semantic processing*, or the extent to which meaning is processed during a task, as a key measure of microtask complexity. This is based on evidence that semantic processing is associated with the depth and elaborateness of mental processing [13]. We also measure *time* spent on a microtask and perceived *mental demand*. Because complexity is sometimes associated with involvement on a task, we also include questions on *interest* and *meaningfulness*, based on existing research on motivation and involvement [31][49]. Lastly, we include a question on *open-endedness* because pilot interviewees described it as a key component of their writing difficulties.

Method

Each participant was randomly assigned to an operation, and completed that microtask operation on a particular sentence. After the microtask, the participant answered subjective questions about the microtask, followed by a multiple-choice quiz about the meaning of the sentence. The subjective questions included the mental demand portion of the NASA-TLX [22], and Likert scale items about interest, meaningfulness, and open-endedness. The multiple-choice quiz was modeled after a common question on the TOEFL exam, which asks the person to select which of three sentences is most similar in meaning to the original sentence. We used accuracy on the quiz as a measure of semantic processing during the microtask. In addition to the eleven microtasks, we also included a baseline condition (quiz-only), where only the quiz was completed without any preceding microtask. The original sentence was hidden during the quiz for all aside from the quiz-only condition.

In early iterations, we found that objective measures on writing tasks, such as task completion time and semantic processing, fluctuate considerably depending on the nuances of the sentence, even if those sentences are of similar readability. To eliminate confounds resulting from sentence variations, in each study of this paper we hold the sentence constant on any microtasks being compared, but randomly sampled from the corpus otherwise.

We restricted participants to those on Amazon Mechanical Turk residing in the United States with at least a 95% ap-

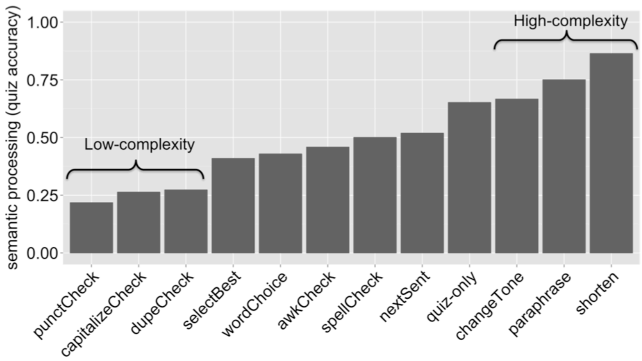

Figure 1. The level of semantic processing (quiz accuracy) by operation. In all conditions except the control (labeled *quiz-only*), the original sentence was hidden during the quiz. On average, 22 participants completed each operation. Based on these results, we selected low, medium, and high-complexity operations for further study.

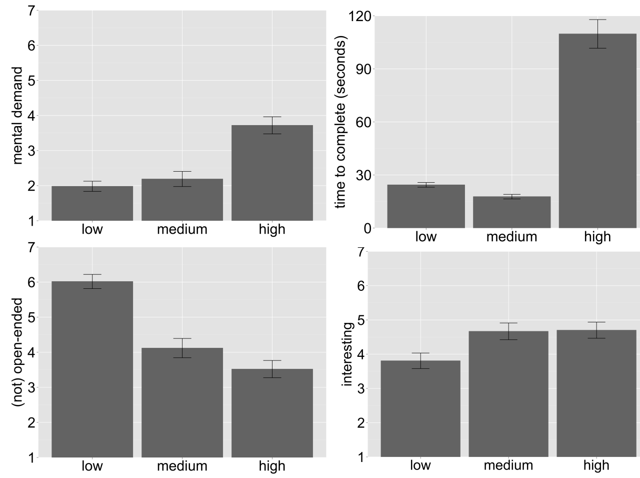

Figure 2. Medium-complexity microtasks are similar to low-complexity microtasks on mental demand, task time, and semantic processing, but similar to high-complexity microtasks on open-endedness, interest, and meaningfulness.

proval rating. Each person was compensated $0.30 and repeat participation was disallowed. To avoid selection bias, the same HIT preview page was displayed to all participants regardless of which condition they were placed in. 264 participated in the study.

Given that semantics-level microtasks involve more in-depth consideration of meaning than mechanics-level microtasks, we expect these tasks to exhibit greater semantic processing (higher quiz performance), impose more mental demand, take longer time to complete, and be considered more open-ended than mechanics-level tasks.

Analysis

In all studies, we tracked any loss of browser window focus, and excluded timing data for those who were away for more than fifteen seconds. Extreme outliers were also removed. For any timing data that was not normally distributed, but was log-normal, we applied a log-transformation before running significance tests. Accuracy metrics (e.g., quiz performance) were evaluated using a logistic regression. For Likert scale items, we report on ANOVA results, but non-parametric tests yielded empirically similar results.

Results

Using this approach, we were able to identify microtasks of different complexities, as shown in Figure 1. The task of shortening a sentence (shorten) led to the best performance on the quiz, whereas the task of checking for punctuation error (punctCheck) led to the worst performance. Interestingly, mechanics tasks led to below random quiz performance. Because the wrong quiz answers typically contained words from the original sentence, whereas the correct answer modified words while kept the original meaning intact, it is possible that those in the mechanics conditions simply guessed by selecting answers containing words they had seen.

As expected, we found significant differences between all pairs of mechanics microtasks and semantics microtasks ($p<0.05$), with the exception of spellCheck and nextSent. Semantics microtasks had significantly higher semantic processing and longer task time, were perceived to be more open-ended and imposed greater mental demand than mechanics microtasks. SpellCheck and nextSent did not align well with these categories, possibly because spellcheck sometimes activates a moderate level of semantic processing [20], and nextSent may be open-ended enough to be completed without fully understanding the original sentence. No differences were found within each set of mechanics tasks and semantics task s.

In analyzing the crowd-based microtasks (awkCheck, wordChoice, selectBest), we find that even though all three microtasks performed similarly on semantic processing, wordChoice took longer to complete, likely because it involved generating new content. Excluding wordChoice (as well as spellCheck and nextSent for reasons stated above), we compared how the crowd-based microtasks {awkCheck, selectBest} compared to mechanics and semantics operations. While awkCheck and selectBest had lower semantic processing, mental demand, and completion time than semantics microtasks (all $p<0.05$), they were also more interesting and open-ended than mechanics microtasks (all $p<0.05$), and more meaningful than mechanics with marginal significance ($p=0.06$) (Figure 2).

Based on the results, we chose a subset of these operations for further study. Our goal was to select sets of operations at opposite ends of the complexity spectrum, such that the sets are substantially different in complexity, but with similar enough operations within each set to be interchangeably used in our studies. We thus select {capitalizeCheck, punctuationCheck, duplicateCheck} to be *low-complexity microtasks* (L), and {paraphrase, toneChange, shorten} to be *high-complexity microtasks* (H). We also select {awkCheck, selectBest} to be *medium-complexity microtasks*

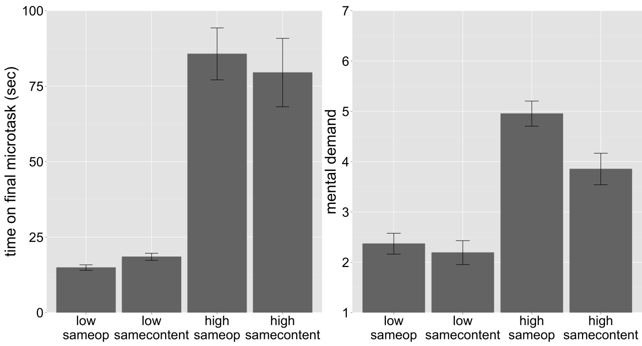

Figure 3. In low-complexity chains, the final microtask was completed faster in the *same-operation* condition. In high-complexity chains, participants found the final microtask less mentally demanding in the *same-content* condition.

(M), since they are more interesting than low-complexity tasks, but easier than high-complexity tasks.

MICROTASK CONTINUITY

Using the microtasks identified in the previous section, we evaluate how ordering affects microtask chains through a series of three studies, spaced across several weeks. We first assess microtask continuity: how task chaining affects a chain of microtasks with the same complexity. We asked participants to perform a series of microtasks, and evaluated their experience on the final microtask. We varied whether preceding microtasks had shared the same *operation* or the same *content* as the final microtask. In addition, we examined whether these effects depend on the complexity-level of the microtask.

Method

The study followed a 2 (complexity) x 2 (chain type) between-subject design, where complexity was either high (H) or low (L), and chain type was either same-operation or same-content. In all conditions, the chain consisted of three consecutive microtasks. In the same-operation conditions, participants did the same task (e.g., paraphrase, paraphrase, paraphrase) on three different sentences. In same-content conditions, participants did three different tasks (e.g., changeTone, shorten, paraphrase) on one single sentence. We used three tasks per chain because we had three operations per complexity level to work with. The study used a between-subject design because our goal was specifically to evaluate the effects of one microtask on the next, which could be potentially confounded in a within-subject design containing consecutive conditions.

In all microtasks aside from the final one in each chain, sentences were randomly sampled from the sentence corpus, and the order of sentences and task types was randomized for each participant. For the purpose of comparing performance on the final task, the sentence and task type of the final task was held constant. The final H task in a chain of high-complexity tasks was always paraphrase, and the final L task in a chain of low-complexity tasks was always duplicateCheck. We chose these particular microtasks because, compared to other microtasks of the same complexity, they demonstrated the most consistent performance across the complexity metrics previously described. Participants were compensated $1.00 for the study. 183 people completed the study.

Measures

We used measures of time, quality, mental demand, helpfulness and enjoyment to understand continuity.

- *Time.* We measured time spent on the final task.

- *Quality.* In L chains, we computed whether the participant correctly fixed the mechanics error on the final task. In H chains, the sentences produced by participants on the final task were each rated on a 5-point Likert scale by three different workers on Mechanical Turk.

- *Mental demand.* Participants completed the mental demand portion of the NASA TLX about the final task.

- *Helpfulness.* Participants rated on a 7-point Likert scale to what extent the first two tasks helped them do the final task.

- *Enjoyment.* Participants rated on a 7-point Likert scale to what extent they enjoyed the full microtask chains.

Because L microtasks are easier and encourage speed, we expected these microtasks to benefit more from same-operation chains which enable people to perform a series of similar microtasks in a row. Conversely, because H tasks are more cognitive and semantically rich, we hypothesized they would benefit from same-content chains, which allow people to focus and build on similar content across multiple microtasks.

Results

Overall, we found that tasks were affected by the preceding tasks, and that complexity-level was an important factor in mediating the effects of chain type. Low-complexity microtasks took less time when preceded by same-operation microtasks, but high-complexity microtasks were perceived to be less demanding when preceded by same-content microtasks (Figure 3). Specifically, we found a significant interaction effect between complexity and chain type on task completion time ($F(1,149)=6.6$, $p<0.05$). In low-complexity chains, participants completed the final microtask faster when it was preceded by same-operation tasks (mean=14.91 sec) than by same-content tasks (mean=18.49 sec), a difference of 3.58 seconds ($p<0.05$). However, no significant time difference was found on high-complexity conditions. In addition, no difference in quality was found.

In analyzing mental demand, we observed a marginally significant interaction effect between complexity and chain type ($F(1,179)=3.35$, $p=0.06$) (Figure 3). We therefore examined high-complexity and low-complexity conditions separately. On high-complexity chains, those in the same-content condition found the final task significantly less mentally demanding (mean=3.85) than those in the same-operation condition (mean=4.95, $p<0.01$). This suggests that the semantic meaning extracted during a H task builds

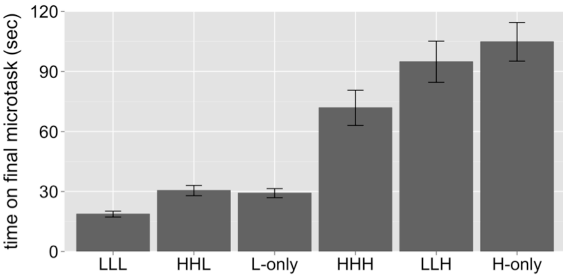

Figure 4. Same-complexity microtasks led to significantly faster completion time on the final microtask, compared to different-complexity microtasks and control conditions.

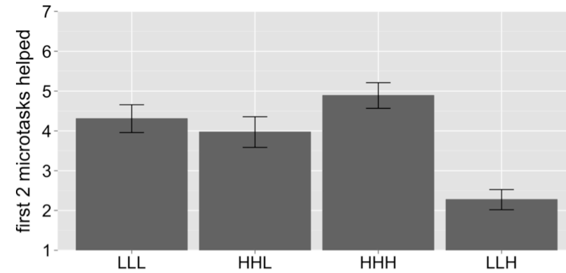

Figure 5. H microtasks were perceived to be significantly more helpful than L microtasks for completing final H microtask when chaining across complexity.

up a mental state [8] about the sentence that is then utilized on subsequent tasks requiring access to the same state.

Lastly, we found that initial microtasks were perceived to be significantly more helpful to the final microtask in high-complexity chains (mean=5.21) than in low-complexity chains (mean=4.46, $F(1,179)=9.09$, $p<0.005$). However, overall enjoyment was greater on low-complexity chains (mean=5.98) than high-complexity chains (mean=5.40, $F(1,179)=7.43$, $p<0.05$). This is consistent with our earlier finding that low-complexity microtasks can be completed without as much scaffolding or semantic processing.

In summary, low-complexity microtasks were completed faster when preceded by same-operation microtasks. However, high-complexity microtasks felt less mentally demanding when preceded by same-content microtasks.

MICROTASK TRANSITIONS

In the previous study we found that a person's experience with a microtask was affected by whether preceding tasks were of the same *operation* or the same *content*, given a constant complexity-level. However, microtask complexity might also vary within a chain of tasks. In writing, one might perform diverse modifications on the same section of content, with some that are more cognitively complex than others. In this study, we investigate the effects of transitioning between microtasks of the *same* complexity or *different* complexity, when the content is the same.

Method

The study followed a 2 (complexity) x 3 (chain type) between-subject design, with all microtasks performed on the same content. The complexity of the final microtask was either high or low, and the chain type was either same-complexity, different-complexity, or a control which had no microtasks before the final microtask. Specifically, the six conditions were: LLL (same-complexity), HHL (different-complexity), L-only (control), HHH (same-complexity), LLH (different-complexity), and H-only (control). We included L-only and H-only as control conditions to understand how performing a microtask with no preceding tasks compares to one that transitions from other microtasks, since these complexity shifts could impose a mental switch cost [47].

Similar to the previous study, microtask order was randomized for all except for the final microtask. The same content was held constant throughout the chains, and each participant did three different operations within each chain, where the last H microtask was paraphrase and last L microtask was duplicateCheck. The measures were the same as those described under the Microtask Continuity section, and participants were compensated $1.00 for the study. A total of 173 people participated in the experiment.

Since L microtasks demand relatively little semantic processing, we do not expect performance to be affected by whether the task is preceded by H tasks (HHL) or L tasks (LLL). Furthermore, if transitioning from H to L imposes an additional switch cost, HHL might perform worse than L-only. In contrast, because H tasks are more cognitive in nature and require semantic processing, we expect lead-up H tasks (HHH) to serve an advantage over lead-up L tasks (LLH). If transitioning from L to H imposes an additional switch cost, LLH might also perform worse than H-only.

Results

Overall, we found that the complexity of initial microtasks mattered, and their helpfulness was also perceived differently depending on the complexity of the final microtask. A two-way ANOVA found a significant main effect of chain type on task time ($F(2,139)=12.09$, $p<0.005$). Post-hoc Tukey tests indicate that same-complexity chains led to faster completion times on the final microtask (mean=46.46 sec), compared to control (mean=64.60 sec) and different-complexity (mean=61.47 sec) chains ($p<0.05$, Figure 4). No differences in quality were found.

Furthermore, we saw a significant interaction effect of complexity and chain type on perceived helpfulness ($F(1,116)=11.49$, $p<0.0005$) (Figure 5). For chains ending in a high-complexity microtask, those in the same-complexity condition (HHH) found the lead-up tasks to be significantly more helpful (mean=4.89) than those in the different-complexity condition (LLH) (mean=2.27, $p<0.0005$). However, no difference was found between chains ending on a low-complexity microtask. Hence, for chains ending on complex microtasks, same-complexity

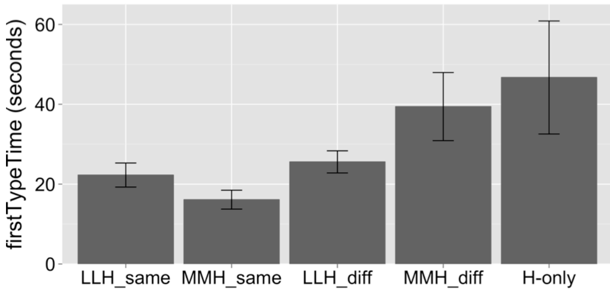

Figure 6. Typing started significantly sooner when an H microtask was preceded by M microtasks on the same content.

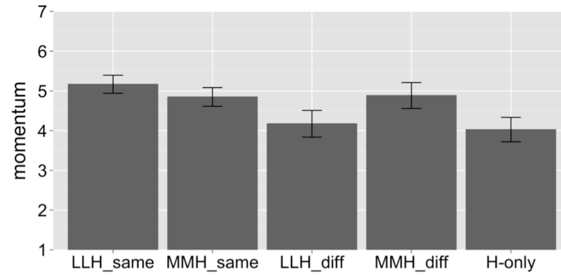

Figure 7. Participants felt they had significantly greater momentum going into an H microtask when it was first preceded by L microtasks on the same content.

chains were perceived to offer a cognitive benefit over different-complexity chains.

Lastly, H-only tasks were significantly less enjoyable (mean=4.08) than L-only tasks (mean=5.71, p<0.005). Even though H tasks may be valuable to future H tasks, they demand an upfront ramp-up of meaning and focus that demands effort. On average, those in the H-only condition spent 33 seconds longer completing the task than those who first completed other H tasks (HHH condition). The equivalent time lost was only 11 seconds in the L-only condition.

We found no evidence that transitioning between complexities dampens performance compared to starting off immediately with the final microtask. No significant difference was found on either time or quality between different-complexity and control conditions. We hypothesize that starting immediately with the final microtask is itself a task switch, since it still demands mentally transitioning from the user's previous activity.

In summary, same-complexity microtasks led to faster completion times and were perceived to be more helpful when the final microtask was complex. Low-complexity microtasks were completed slower when preceded by high-complexity microtasks than when preceded by low-complexity microtasks, possibly due to a depletion of mental resources [41]. However, we found no evidence that complexity-switching is worse than starting immediately on the microtask, possibly because the latter still incurs a switch cost. Completing an initial high-complexity microtask may also be particularly arduous, even if that effort pays off on future high-complexity tasks. These findings raise the question of how microtasks could be better chained to ease people into meaning-rich microtasks.

MICROTASK EASE IN

In the previous study, we found that initial H tasks were valuable to the completion of further H tasks of similar content, but they were also cognitively demanding and less enjoyable to do compared to L tasks. In this study, we examine whether performing lower-complexity microtasks can help people *start* a high-complexity microtask sooner, while simultaneously accomplishing a unit of work itself.

Method

To understand the impact of leading up with various kinds of simpler microtasks, we conducted a between-subject study with a 2 (complexity) x 2 (content) design, where the complexity of the first two microtasks was either low (L) or medium (M), and the content was either same or different from the final microtask. We also include a H-only condition with no lead-up microtasks to serve as a control. All conditions ended on a high-complexity microtask. We include M microtasks as a potentially interesting class of lead-up tasks because they were perceived to be more interesting and open-ended than L microtasks, yet still faster and easier to complete than H microtasks (see Selecting Writing Microtasks). We also compare same-content to different-content lead-up tasks because same-content was previously found to help build context toward a high-complexity task (see Microtask Continuity).

The four conditions were: LLH_same, MMH_same, LLH_diff, MMH_diff, and H-only. Similar to previous studies in this paper, the final (H) microtask was a paraphrase task across conditions, the order of the first two task operations was randomized, and the first two sentences were also randomly sampled in the different-content conditions. Participants were compensated $0.50 for the study. 201 people participated in the study.

Measures

As with the previous studies we used measures of time, quality, and mental demand. We also looked at the amount of time it took to start a task and measures of participant-reported momentum and warm-up.

- *Time, Quality, Mental Demand* on the final task, as before.

- *FirstTypeTime*. We measured the duration between the final task appearing and the person starting to type.

- *Momentum*. Participants rated on a 7-point Likert scale to what extent they felt momentum going into the final task.

- *Warm-up.* Participants rated on a 7-point Likert scale how mentally warmed up they felt going into the final task.

To prevent the questionnaire itself from confounding the experience of completing a task chain, all Likert-scale questions were asked after the final task.

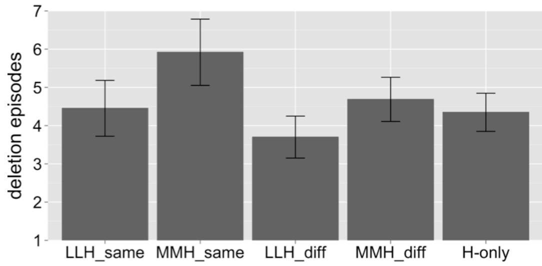

Figure 8. After performing M microtasks on the same content, participants made many deletes during the H microtask.

Because tasks chained on the same content could help build awareness of meaning, we hypothesize same-content lead-up tasks to yield stronger benefits than different-content lead-up tasks. M lead-up tasks may be more helpful than L lead-up tasks, due to greater meaning extracted.

Results

Lead-up microtasks had an effect on the final microtask, and these effects varied depending on their properties. We found a significant effect of condition on FirstTypeTime ($F(4,195)=5.4$, $p<0.0005$), the results of which are shown in Figure 6. Participants initiated typing significantly sooner in the MMH_same condition (mean=16.11 sec) than the H-only condition (mean=46.71 sec), a difference of 30.6 seconds ($p<0.005$, Tukey's post-hoc test). While typing does not necessarily mean faster engagement, it does suggest they were able to start transforming thoughts into text sooner, possibly due to context gained from the lead-up tasks. Whereas 43% of MMH_same participants started typing within 10 seconds of the H task being displayed, only 11% started typing in the H_only condition, and 38% in the LLH_same condition. The difference between LLH_same and H-only was marginally significant after post-hoc correction ($p=0.08$). No difference in quality was found.

Additionally, participants in the LLH_same condition reported a significantly greater sense of momentum going into the final task (mean=5.17), compared to H_only participants (mean=4.03, $p<0.05$) (Figure 7). The concrete, self-contained nature of low-complexity tasks may have helped build initial rhythm. Lastly, users in the LLH_same condition felt significantly more mentally warmed up (mean=5.25) than those in the LLH_diff condition (mean=3.87, $p<0.05$), suggesting that some awareness of content could be developed even during low-level mechanics tasks. However, given that participants did not perceive L microtasks to help with H microtasks in the prior study (see Microtask Transition results), these benefits may be subtle, and may be less obvious to the average user.

Although same-content lead-up tasks contributed to greater momentum and earlier action taken, we found no significant difference in task completion time on the final H task. Compared to L microtasks, the open-ended nature of H microtasks may mean that task time is not only attributed to efficiency, but also immersion in the task. Our post-hoc evaluations found that those in the MMH_same condition made a relatively large number of deletions during the H task (mean=5.9), suggesting a certain level of involvement (Figure 8) that may contribute to longer task times.

	SUMMARY OF FINDINGS
Microtask Continuity	**RQ 1:** When chains have the same complexity across microtasks, what is the effect of continuing on the same *operation* vs. same *content*?
	- In low-complexity chains, same-operation resulted in faster completion time on the final microtask than same-content. - In high-complexity chains, same-content resulted in lower mental demand on the final microtask than same-operation.
Microtask Transitions	**RQ 2:** When chains use the same content across microtasks, what is the effect of transitioning from microtasks of the *same* complexity vs. *different* complexity?
	- Same-complexity resulted in faster time on the final microtask than different-complexity and control conditions. - Same-complexity was perceived to be more helpful than different-complexity to a final high-complexity microtask. - Starting directly with a high-complexity microtask was less enjoyable than starting directly with a low-complexity microtask.
Microtask Ease-In	**RQ 3:** How is a person's experience on a *complex* microtask affected when it is preceded by *simpler* microtasks?
	- People started typing sooner if they first completed medium-complexity microtasks on the same content. - People perceived greater momentum if they first completed low-complexity microtasks on the same content. - People felt more mentally warmed up if they first completed low-complexity microtasks on the same content, compared to low-complexity microtasks on different content.

Table 3. Summary of findings from studies on Microtask Continuity, Microtask Transitions, and Microtask Ease In.

In summary, high-complexity microtasks were initiated sooner when preceded by medium-complexity microtasks, and were perceived to have greater momentum when preceded by low-complexity microtasks. We observed these effects only in same-content chains, suggesting that some content awareness is developed even on low-complexity microtasks. These findings are consistent with ease-in strategies described by participants in our pilot interviews, and lend support to prior work that shows concrete, short-term goals can help people get started on larger tasks [1][4].

DESIGN IMPLICATIONS

We have seen that microtasks have carryover effects on subsequent microtasks, both in chains that continued with the same complexity and that transitioned across complexities. Table 3 summarizes our results. Given that the overall time taken to complete a chain of tasks was short (median=132 seconds), the ordering effects we observed may have far-reaching cumulative effects. In this section, we discuss how our findings can be used by researchers and practitioners to improve microtask chains for writing.

Build momentum using same-operation L microtasks. At the start of a microtask chain, when users may be less focused, performing a series of low-complexity microtasks on the same operation could help build speed and rhythm. On low-complexity tasks, we found that same-operation chains led to higher efficiency compared to same-content chains (see Microtask Continuity). Preserving the same type of operation while switching on content may be particularly apt at the start of a chain, a time when users have relatively little context and are still acquainting to the task.

Transition using L and M microtasks on similar content. As high-complexity microtasks may be painstaking to do at the start of a chain (see Microtask Transition), low-complexity microtasks could help build momentum and awareness of meaning, while simultaneously completing a unit of work. In our study (see Microtask Ease-In), M microtasks enabled people to start typing sooner on a subsequent H microtask, and L microtasks were perceived to build momentum, but only in cases where content was similar to that of the final H task.

Avoid switching content during a series of H microtasks. Because the rich meaning extracted during a H task builds up semantic understanding about the content, interleaving content during H tasks may disrupt the mental state [8] associated with the task. Instead, staying on the same content could enhance experience on subsequent H tasks requiring access to the same state, building continuity in the task at hand. Participants rated H microtasks less mentally demanding when preceded by *same-content* microtasks than when preceded by *same-operation* microtasks (see Microtask Continuity). In crowd work, where several operations may be needed for multiple pieces of content, presenting similar content consecutively (e.g., adjacent audio segments) could help preserve the context gained in earlier tasks. In personal tasks, organizing tasks by region of content (e.g., email thread or lecture topic) could alleviate cognitive effort amidst competing priorities.

Consider switching content on transitions from H to L. Although transitions into H microtasks are supported by tasks with similar content, transitions into L microtasks demand less context and awareness of meaning. In fact, performance was slower when low-complexity tasks were preceded by high-complexity tasks on the same content, than when they were preceded by low-complexity tasks on the same content (see Microtask Transition). After completing a meaning-rich H task, shifting the content could potentially give a renewed perspective or "fresh eyes" to a low-level task such as proofreading, though without further investigation, this implication remains preliminary.

DISCUSSION AND FUTURE WORK

Our findings demonstrate the substantial impact that small changes in microtask ordering could have on users. While same-operation chains aid in the efficiency of simple microtasks, same-content chains may help alleviate mental burden on more complex microtasks. Ordering low-level microtasks first may help people start meaning-rich tasks on the same content. Easing the transition into meaningful work is not only important in personal tasks, where users often have to cope with interruptions, but also on crowd platforms where workers regularly switch tasks [28].

Although our studies provide direct implications for editing, the general classes of measures that were investigated (operation, content, and complexity) can be found in diverse domains, and provide a baseline for future work in this area. Aside from information management and crowd work, task chains are also common in educational exercises [9], micro-volunteer work [46], and systems for rehabilitation [37]. As more tasks are transformed into micro-components, systems should consider the user's cumulative experience and memory when ordering microtasks.

Given that individuals often engage in simple tasks as a way of starting more complex tasks, effective chaining could aid not only transitions between microtasks, but also transitions into larger tasks. One could imagine leveraging different ease-in chains for different purposes. For example, since medium-complexity microtasks led to faster action initiation, they could be used in mobile scenarios, when a user may otherwise be unmotivated to start a meaning-rich task. Alternatively, low-complexity microtasks can help build initial momentum in a desktop scenario, where the user simply needs support re-gaining focus.

In a complex system, microtask ordering may be subject to interdependencies and global constraints. Our findings could be combined with those additional constraints to inform which microtask ought to be presented next, given a history of recently completed microtasks and their properties. Although we used sentence-level units, the microtask properties described could be applied to longer text, such as keeping same-content chains within the same region of a document. We leave document-level tasks to future work.

Lastly, the task chains in our studies were relatively short, and were evaluated on text that was not originally authored by the participants. We held constant the chain length and text across conditions in order to study ordering effects alone. Future work should investigate ordering effects in longer chains, and explore situations that leverage an individual's long-term familiarity with the content.

CONCLUSION

In this paper, we showed that microtasks have non-negligible effects on other microtasks within the same chain, and demonstrated how key properties of a microtask (operation, content, and complexity) mediate the efficiency, mental workload, and activation of future microtasks. Through a series of studies focusing on writing, we found that small changes in task ordering can significantly affect microtask continuity, transitions, and ease-in. Taken together, these findings have important implications for designing effective microtask chains for writing.

ACKNOWLEDGEMENTS

We thank the selfsourcing team at Microsoft Research, including Nick Greer, Rhema Linder, Niloufar Salehi, Ke Tran, and Rajan Vaish for their valuable input. We are also grateful to Rob Miller, William Li, and Ramesh Sridharan for their feedback.

REFERENCES

1. David S. Ackerman and Barbara L. Gross. 2005. My instructor made me do it: Task characteristics of procrastination. *Journal of Marketing Education* 27, 1: 5–13.
2. Piotr D. Adamczyk and Brian P. Bailey. 2004. If not now, when?: the effects of interruption at different moments within task execution. *Proceedings of the SIGCHI conference on Human factors in Computing Systems*, ACM, 271–278.
3. Paul S. Adler, Barbara Goldoftas, and David I. Levine. 1999. Flexibility versus efficiency? A case study of model changeovers in the Toyota production system. *Organization science* 10, 1: 43–68.
4. David Allen. 2015. *Getting things done: The art of stress-free productivity*. Penguin.
5. Brian P. Bailey and Shamsi T. Iqbal. 2008. Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management. *ACM Transactions on Computer-Human Interaction (TOCHI)* 14, 4: 21.
6. Albert Bandura. 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review* 84, 2: 191.
7. Michael S. Bernstein, Greg Little, Robert C. Miller, et al. 2010. Soylent: a word processor with a crowd inside. *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, ACM, 313–322.
8. Jelmer P. Borst, Niels A. Taatgen, and Hedderik van Rijn. 2015. What Makes Interruptions Disruptive?: A Process-Model Account of the Effects of the Problem State Bottleneck on Task Interruption and Resumption. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, 2971–2980.
9. Eric Butler, Erik Andersen, Adam M. Smith, Sumit Gulwani, Zoran Popovic, and W. A. Redmond. 2015. Automatic Game Progression Design through Analysis of Solution Features. *Proc. of the SIGCHI Conf. on Human Factors in Computing (CHI'2015)*.
10. Carrie J. Cai, Philip J. Guo, James R. Glass, and Robert C. Miller. 2015. Wait-Learning: Leveraging wait time for second language education. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, 3701–3710.
11. Justin Cheng, Jaime Teevan, Shamsi T. Iqbal, and Michael S. Bernstein. 2015. Break it down: A comparison of macro-and microtasks. *Proceedings of CHI*.
12. Lydia B. Chilton, John J. Horton, Robert C. Miller, and Shiri Azenkot. 2010. Task search in a human computation market. *Proceedings of the ACM SIGKDD workshop on human computation*, ACM, 1–9.
13. Fergus IM Craik and Robert S. Lockhart. 1972. Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior* 11, 6: 671–684.
14. Peng Dai, Jeffrey M. Rzeszotarski, Praveen Paritosh, and Ed H. Chi. 2015. And Now for Something Completely Different: Improving Crowdsourcing Workflows with Micro-Diversions. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, ACM, 628–638.
15. Steven Dow, Anand Kulkarni, Scott Klemmer, and Björn Hartmann. 2012. Shepherding the crowd yields better work. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ACM, 1013–1022.
16. Darren Edge, Elly Searle, Kevin Chiu, Jing Zhao, and James A. Landay. 2011. MicroMandarin: mobile language learning in context. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 3169–3178.
17. Joseph R. Ferrari, Judith L. Johnson, and William George McCown. 1995. *Procrastination and task avoidance: Theory, research, and treatment*. Springer Science & Business Media.
18. Linda Flower. 1979. Writer-based prose: A cognitive basis for problems in writing. *College English*: 19–37.
19. Tony Gillie and Donald Broadbent. 1989. What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological research* 50, 4: 243–250.
20. W. Graves, J. Binder, M. Seidenberg, and R. Desai. "Neural correlates of semantic processing in reading aloud." In M. Faust (Ed.), Handbook of the Neuropsychology of Language. Malden, MA: Wiley-Blackwell.
21. Nick Greer, Jaime Teevan, and Shamsi T. Iqbal. 2016. An Introduction to Technological Support for Writing. *Microsoft Research Tech Report MSR-TR2016-001*.
22. Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology* 52: 139–183.
23. Suzanne Hidi and Pietro Boscolo. 2006. Motivation and writing. *Handbook of writing research*: 144–157.
24. John Joseph Horton and Lydia B. Chilton. 2010. The labor economics of paid crowdsourcing. *Proceedings of the 11th ACM conference on Electronic commerce*, ACM, 209–218.
25. Panagiotis G. Ipeirotis. 2010. Analyzing the amazon mechanical turk marketplace. *XRDS: Crossroads, The ACM Magazine for Students* 17, 2: 16–21.
26. Shamsi T. Iqbal and Brian P. Bailey. 2008. Effects of intelligent notification management on users and their

tasks. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 93–102.

27. Shamsi T. Iqbal and Eric Horvitz. 2007. Disruption and recovery of computing tasks: field study, analysis, and directions. *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 677–686.
28. Walter S. Lasecki, Jeffrey M. Rzeszotarski, Adam Marcus, and Jeffrey P. Bigham. 2015. The Effects of Sequence and Delay on Crowd Work. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, 1375–1378.
29. Ulf Lundberg, Marianne Granqvist, Tommy Hansson, Marianne Magnusson, and Leif Wallin. 1989. Psychological and physiological stress responses during repetitive work at an assembly line. *Work & Stress* 3, 2: 143–153.
30. Gloria Mark, Daniela Gudith, and Ulrich Klocke. 2008. The cost of interrupted work: more speed and stress. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, ACM, 107–110.
31. Andrew A. Mitchell. 1979. Involvement: a potentially important mediator of consumer behavior. *Advances in consumer research* 6, 1: 191–196.
32. Stephen Monsell. 2003. Task switching. *Trends in cognitive sciences* 7, 3: 134–140.
33. Robert R. Morris, Mira Dontcheva, and Elizabeth M. Gerber. 2012. Priming for better performance in microtask crowdsourcing environments. *Internet Computing, IEEE* 16, 5: 13–19.
34. Mohamed Musthag and Deepak Ganesan. 2013. Labor dynamics in a mobile micro-task market. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 641–650.
35. Michael Nebeling, Alexandra To, Anhong Guo, et al. 2016. WearWrite: Crowd-Assisted Writing from Smartwatches. *Proceedings of CHI*.
36. Christine M. Neuwirth, David S. Kaufer, Ravinder Chandhok, and James H. Morris. 2001. Computer support for distributed collaborative writing: A coordination science perspective. *Coordination Theory and Collaboration Technology, ed. GM Olson, TW Malone, and JB Smith, Lawrence Erlbaum Associates, New Jersey*.
37. Marion C. Panyan and R. Vance Hall. 1978. Effects of Serial versus Concurrent Task Sequencing on Acquisition, Maintenance, and Generalization. *Journal of Applied Behavior Analysis* 11, 1: 67–74.
38. Nathalie Pattyn, Xavier Neyt, David Henderickx, and Eric Soetens. 2008. Psychophysiological investigation of vigilance decrement: boredom or cognitive fatigue? *Physiology & Behavior* 93, 1: 369–378.
39. Ilona R. Posner and Ronald M. Baecker. 1992. How people write together [groupware]. *System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on*, IEEE, 127–138.
40. Bennett A. Rafoth and Donald L. Rubin. 1984. The impact of content and mechanics on judgments of writing quality. *Written Communication* 1, 4: 446–458.
41. Darryl W. Schneider and John R. Anderson. 2010. Asymmetric switch costs as sequential difficulty effects. *The Quarterly Journal of Experimental Psychology* 63, 10: 1873–1894.
42. R. J. Senter and E. A. Smith. 1967. *Automated readability index*. DTIC Document.
43. Gordon Taylor and Peggy Nightingale. 1990. Not mechanics but meaning: Error in tertiary students' writing. *Higher Education Research and Development* 9, 2: 161–176.
44. Jaime Teevan, Shamsi Iqbal, and Curtis von Veh. 2016. Supporting Collaborative Writing with Microtasks. *Proceedings of CHI*.
45. Jaime Teevan, Daniel J. Liebling, and Walter S. Lasecki. 2014. Selfsourcing personal tasks. *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, ACM, 2527–2532.
46. Rajan Vaish, Keith Wyngarden, Jingshu Chen, Brandon Cheung, and Michael S. Bernstein. 2014. Twitch crowdsourcing: crowd contributions in short bursts of time. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, ACM, 3645–3654.
47. Glenn Wylie and Alan Allport. 2000. Task switching and the measurement of "switch costs." *Psychological research* 63, 3-4: 212–233.
48. Ming Yin, Yiling Chen, and Yu-An Sun. 2014. Monetary interventions in crowdsourcing task switching. *Second AAAI Conference on Human Computation and Crowdsourcing*.
49. Judith Lynne Zaichkowsky. 1994. The personal involvement inventory: Reduction, revision, and application to advertising. *Journal of advertising* 23, 4: 59–70.