Characterizing the Mobile Microtask Writing Process

Leave Authors Anonymous
for Submission
City, Country
e-mail address

Leave Authors Anonymous
for Submission
City, Country
e-mail address

Leave Authors Anonymous
for Submission
City, Country
e-mail address

ABSTRACT

Microtasking—breaking down complex tasks into small subtasks—mitigates distractions and leads to more effective work. Writing is an ideal process for microtasking because of its many subgoals (e.g., looking up a reference or writing an introduction sentence), but little is known about how this decomposition impacts writers' processes through the evolution of a document. In this paper we present findings from a controlled, week long user study to characterize how writers use mobile microtasks throughout authoring a document. We found that authors created microtasks for editing and inserting information that generally required minimal writing. These tasks were especially well suited for mobile devices, with authors completing tasks on commutes or waiting for meetings. Authors who microtasked felt that their final documents were more coherent than authors who instead edited their document directly. This was supported by automatic metrics of the document, showing that microtasking writers indeed wrote and edited more.

Author Keywords

microtasking; writing; user study

CCS Concepts

•**Human-centered computing** → *User studies; Text input;*

INTRODUCTION

How we work is becoming increasingly fragmented. Knowledge workers spend an average of less than 3 minutes on a single task, suggesting that constant task switching and managing interruptions has become the workplace norm [9]. These distractions can increase the amount of time spent on tasks while also damaging productivity and increasing stress [12]. Microtasking [5]—breaking down complex tasks into small, manageable, context-free subtasks—allows for workers to complete work while mitigating the consequences of interruptions and distractors. These microtasks can be easily accomplished in short moments, leading to more effective work, reducing time spent on tasks, and shortening ramp-up times after a distraction [20, 9, 5].

One common work process that benefits from being decomposed into smaller subtasks is writing. Effective writers break down their writing goals into smaller subgoals that can be quickly accomplished or revised [8]. For example, while writing a proposal, a writer might be deeply focused, to the point where looking up a particular reference, editing an awkward sentence or drafting an introductory sentence would only interrupt their flow [10]. Microtasks can help the writer effectively defer these peripheral tasks, completing them later when they were not otherwise involved in writing. Because these microtasks rarely require a full word processor [26]—looking up a reference or writing a sentence can be accomplished on a mobile phone—the writer can accomplish them at times they are away from their computer, such as on a commute home or waiting for a meeting to start.

Past work on mobile microtask writing has focused on automatically generating tasks for writers [13] or offloading tasks to crowdworkers [22]; however, less is known about how writers fit microtasking into their own writing process. Writing is a complex process filled with writer-defined goals and subgoals [8]. Automatically generated tasks take control away from writers while tasks completed by crowdworkers force writers to curate tasks able to be completed by a stranger (e.g., providing lots of context or avoiding personal sections) [22].

Our research focuses on the opportunities offered by mobile interfaces to support writing on-the-go. We are interested in answering questions around writers' use of mobile microtasking for writing: how do writers use mobile microtasking as a tool for writing when given the opportunity? What tasks do they create for themselves, how do they complete these tasks while mobile, and does the document look different because of it? Another hurdle for microtask writing is that completing writing microtasks requires context about the document [24, 8]. Providing this context can be key to effectively completing writing tasks [24]; however, it is important not to overwhelm the user. Providing the entire document for a microtask only makes navigation for the task difficult and frustrating, especially for a mobile phone or small screen [17]. How does context play a role in completing mobile microtasks? Such questions can be answered through a deeper exploration of how authors manage their own writing microtasks while authoring a document and how mobile interfaces are integrated into their experiences.

In order to characterize mobile microtask writing processes and how context plays a role completing writing microtasks, we ran a controlled, week long, user study with 24 participants. In a setup involving both focused writing on the desktop and

interactions on a mobile device, we analyzed how writers *created* microtasks to offload appropriate parts of writing to a mobile phone during a writing session on a desktop word processor, how they *completed* microtasks over the course of a week using a prototype mobile system that provided varying levels of context, and how their *integration* of these tasks back into their document impacted the final document quality. We compared this process with a baseline group of writers who instead used Microsoft Word Mobile to complete writing tasks and edited their document on their mobile device directly throughout the week.

We find that for *microtask creation*, quickly creating tasks for later while composing a document allowed writers to continue staying focused on writing. Many of the tasks that were created involved inserting information and adding or editing sentences, which were well suited for mobile microtasking.

The limited space of mobile phones was a constraint for *microtask completion* more so than context: most participants avoided writing extensively on a mobile keyboard by completing the tasks that required only a few lines of surrounding text to contextualize and complete. Those who used the mobile microtasking system gave longer response to their tasks compared to writers using Word Mobile, finding the experience of completing writing tasks easier.

When *integrating* microtasks, writers who used the mobile microtasking system to complete writing tasks ended up making more edits to their documents compared to writers who edited the document directly with Microsoft Word Mobile. This increased editing due to mobile microtasking led to a significant increase in document length and self-rated coherence for those who used our microtasking system, compared with writers using Microsoft Word Mobile. Our results provide an in-depth exploration of how microtasking integrates into writers' processes and exciting evidence of microtasks' usefulness to the writing process. Based on these findings we discuss design recommendations for mobile microtask writing systems.

RELATED WORK

We discuss prior work in how productivity tasks are accomplished while mobile, concepts of microproductivity that explore ways to adapt complex tasks to micromoments, processes of writing to understand opportunities for mobile interfaces, and the importance of context in task completion.

Use of Mobile Devices for Productivity Tasks

People are increasingly using multiple devices, such as phones or tablets, to complete their work. Some tasks lend themselves better to mobile work, such as email, with users often deferring tasks that are ill suited to a mobile environment [1]. At the same time, the portability of a mobile device lends itself well to the fragmented workday many knowledge workers experience, and logs of mobile and desktop use show that mobile devices are often leveraged for productivity tasks [23]. Cross-device synchronization poses a major challenge for shifting tasks between a mobile device and a PC [14]. We build on this work by examining mobile phone productivity in the context of completing writing tasks.

Microproductivity

Microproductivity is the process of breaking down large, complex tasks into small, manageable, context-free tasks that can be completed in short moments. These tasks are referred to as microtasks. Microtasking can lead to more effective work: increasing overall task quality and shortening ramp-up times after an interruption or break in work [5]. Microtasks are an effective strategy for breaking down work and completing small tasks using time previously thought of as unusable, such as while commuting to work, while waiting for a meeting to begin, or even while on social media [11]. It has opened up the possibility of completing work through smaller screens, such as a mobile device [27] or a smartwatch [22]. Mercury [27], a system that automatically generated coding microtasks on a mobile device, allowed programmers to make meaningful progress on a coding task with little time and attention, and to resume the task faster after a break. Play Write [13] allowed users to make progress on a writing task via completion of automatically generated microtasks on a mobile device. Users strongly preferred Play Write to the Word Mobile interface when multitasking, indicating the resilience of microtasking when multitasking [13].

While Mercury and Play Write focused on individuals completing microtasks, systems have also shown the benefits of microtasking across collaborators and crowdworkers. Wearwrite [22] allowed users to coordinate microtask completion by crowdworkers using a smartwatch. Microwriter [26] supported microtasks across collaborators, who found the ability of sharing tasks on a paper helpful for sharing ideas.

In a similar vein, crowdsourcing systems like Crowdforge [18], Soylent [3], and Mechanical Novel [16] all show the value of crowdsourced writing. Soylent allowed users to offload tasks while writing to a crowd that would complete these tasks in real time. Mechanical Novel broke down the entire writing process into tasks easily accomplished by crowdworkers, such as writing a section of a story or voting on the best next chapter of a story [16]. Lastly, work has also shown the feasibility of writing in collaboration with a machine. Clark et al. [6] explored slogan and creative story writing in a human-machine collaboration, finding that users enjoyed the sentences suggested by the machine, allowing them to think more creativity.

Work on microtask writing has focused on automatically-generated tasks [13], multi-author documents [26], or crowdsourcing writing tasks [3, 16, 22]. In each of these contexts, there is little focus on how writers use microtasks throughout their writing progress. In this paper we set out to understand how authors create their own writing tasks, rather than have them automatically generated, how authors complete these tasks, rather than a crowdworker or collaborator, and how authors integrate completed tasks into the document.

Writing Process

Flower & Hayes [8] proposed a cognitive process theory of writing. In this theory, writing is broken down into three core processes: planning, translating, and reviewing. Each of these processes have smaller subgoals, such as organization or goal setting in the planning process. These processes are hierarchical, meaning that while writing, a writer might switch

between each task at different levels of writing. For example, a writer can translate, or write, a single sentence as part of a broader process of planning and organizing a paragraph.

This decomposition of tasks has a natural parallel in microtasks, where writers can break down complex writing goals into small, manageable subgoals. These tasks can have different intentions behind them [28, 15]. Kaur et al. [15] identified how writers do this in academic papers and Wikipedia articles, developing a taxonomy of writing tasks broken down into five categories: adding content, surface-level issues, editing content, and references. In our work, we are interested in how writer-created microtasks fall into these different categories, providing insight into the types of tasks writers feel are appropriate to be deferred to a mobile device.

Conveying Context
While helpful in understanding the use of microtasks for writing, the cognitive process model also provides one potential hurdle for effective microtask writing: writing tasks are often heavily context dependent [24]; it is difficult to imagine writing an introductory sentence without knowing the preceding and following sentences, the goals of the paragraphs, and the trajectory of the document as a whole.

Salehi et al. [24] explored communicating this context in crowdsourced writing through answering questions posed by the crowdworker, commenting and editing a crowdworker's response, highlighting areas of the response that were good or needed work, or providing high level feedback about a response. They found that for early drafts providing high level feedback made the best responses, while for later drafts highlighting areas was most effective. Other microtask writing systems have provided the entire document [22], parts of the document [3], or restricted tasks to only those that needed little context, such as editing tasks [13] as opposed to content creation. While past work has explored communicating context to crowdworkers who have never worked on the document before but have access to a computer and few time constraints, in this paper we explore the most effective ways of providing context for a single author to complete their own microtasks on a mobile device having not worked on the document recently.

Work in search behaviour has also explored how to communicate the context of a document using only a short piece of text, referred to as a snippet [7]. Snippet length can dramatically impact the time it takes to complete a search task, with longer snippets (5-7 lines) helping for informational tasks (where users tried to find specific information) but hindered navigational tasks (where users tried to find a specific website). Building on this work, Kim et al. [17] evaluated snippet length for mobile browsers, finding that long snippets increased search time without improving accuracy due to users having to scroll more to see new websites. We extend this work on mobile browsing to mobile microtask writing, evaluating if the amount of context can overwhelm or force users into time-consuming interactions similar to snippet length in mobile search tasks.

CHARACTERIZING MOBILE MICROWRITING
We evaluated how authors used microtasks throughout their writing process with a controlled user study where participants wrote a document and completed microtasks for the document on their phone as part of the editing process. Part of the study involved using a prototype mobile writing system, which we describe in Section 3.3. We analyzed user behaviour during writing and completing microtasks, final document quality, and user feedback on the tool.

The study was designed to explore the following research questions:

1. **RQ1**: *Microtask Creation* - what kinds of tasks do writers create for themselves to complete on a mobile phone?

2. **RQ2**: *Microtask Completion* - how do writers complete tasks on a mobile phone and how does this compare with editing the document directly on a phone?

3. **RQ3**: *Content Integration* - how does integration of the content created while mobile impact the final document quality?

Participants
Participants were recruited at a large technology company via email lists and flyers. Participant's average age was 30.25 years (std=7.01). 10 participants were native English speakers, 11 were fluent, and 3 had professional working proficiency. All participants had completed at least a bachelor's degree. The majority of participants were researchers or software engineers, with 7 in managerial or business positions. 21 participants used their mobile phone for productivity tasks (e.g., answering email) at least weekly. Participants were compensated with a $50 giftcard. A total of 24 participants completed the study (M=14, F=10). Assignment was split across all 5 conditions–the 4 context microtasking conditions and the baseline–resulting in 5 participants assigned to the baseline condition (one context condition held 4 participants). This small number of baseline participants reduced the power for significance tests comparing the baseline to the microtasking writers. We report both qualitative and quantitative differences between the two experiences.

Experimental Design
The study had three parts: 1) a 15-minute in-lab writing session on a predefined topic using Microsoft Word where at the end they would create microtasks for future consumption, 2) a week-long in-the-wild session for completing microtasks either using our prototype mobile microtasking system (described in the next section) or Word Mobile and 3) a final 5-minute in-lab writing session where writers integrated the outcome of the microtasks into the original Microsoft Word document and finalized the document. The study was designed to accommodate participants' own productivity habits while still controlling for writing topics and amount of time given to a participant for writing.

Because microtasks are presented independently outside of the originating document, prior work had shown that this limited context of microtasks can often make it difficult to complete them on a mobile device [13]. In that work the context was fixed (limited to the surrounding sentence or surrounding paragraph), as the tasks were automatically generated. However,
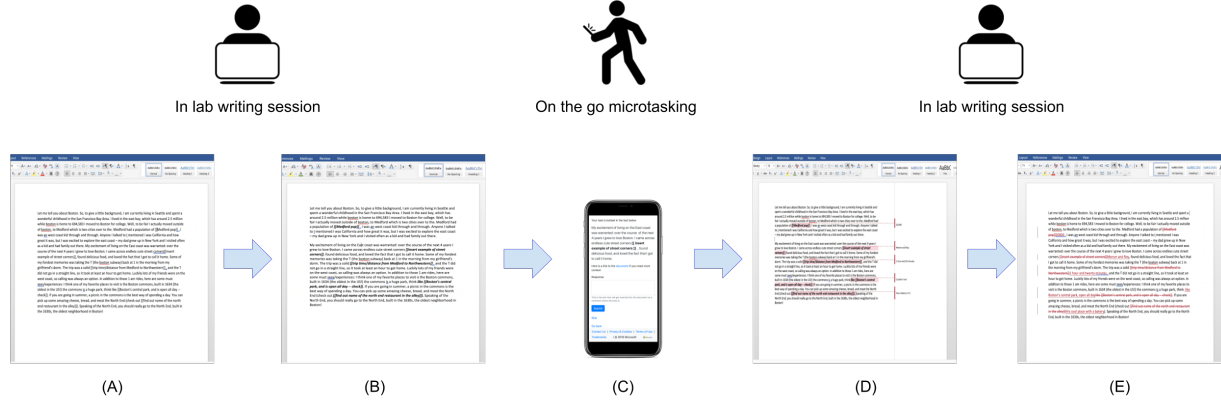
Figure 1: Document lifecycle for writers using microtask system. (A) Writers write a document for 10 minutes. (B) Writers create/convert notes into microtasks. (C) Writers complete microtasks on their mobile phone over the course of a week using the microtask system. (D) Completed microtasks are updated in the document as comments. (E) Writers edit the document with the completed microtasks.

for microtasks that are created by the users, we were interested in understanding how much additional context was needed to complete the task. With this goal in mind, we designed the study to be a 2x2 between-subjects study exploring two categories of context: global and local. Each category of context was further broken down into 2 levels of expressing this context.

For local context, participants could see either enough text to fit onto a single mobile screen without scrolling (glanceable) or enough text surrounding a task to overflow a mobile screen (scrolling). For global context, users were either provided the writing prompt they received when writing the document (prompt) or a link to their document (document).

To compare microtasking with a realistic baseline of writing, there was an additional condition where participants had no access to microtasks but could access and edit the document directly on their phone using Microsoft Word Mobile over the course of the week.

Participants saw one level of local context (glanceable or scrolling, see Figure 2) and one level of global context (prompt or document) for each microtask. Participants were assigned into one level for each type of context from the 4 possible combinations, or the baseline condition, for a total of 5 conditions. Both context levels were fixed for all of a particpant's microtasks.

System
We developed a prototype microtask writing system to characterize how microtasking influenced a writers' process and test different levels of context to provide. The writing microtask system has two subsystems: a Microsoft Word add-in ('the add-in') for capturing microtask creation, and a mobile web application ('the web app') for users to complete microtasks from their mobile device.
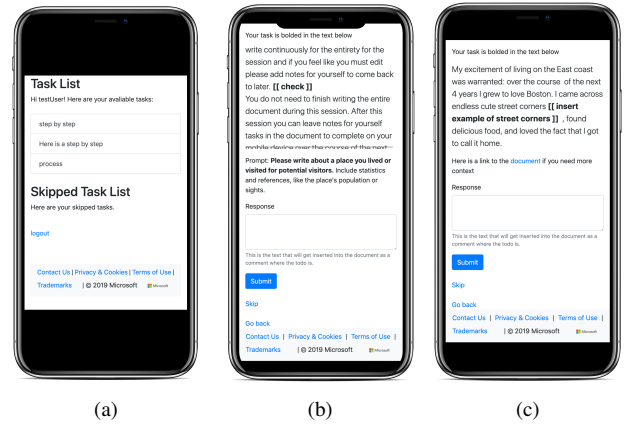

Figure 2: Views of the web app. (a) Task list user will see after logging in. (a) Task display with the prompt and scrolling local context. (a) Task display with the document link and glanceable local context.

Microsoft Word Add-in
The add-in is a VSTO Microsoft Word add-in that captures user actions (e.g., typing) in a document. Whenever a user inserts double brackets ([[ ]]) around a piece of text, the add-in saves the text inside the brackets as the microtask itself and collects the text surrounding the microtask as local context. The amount of text the add-in collects depends on the local context condition the writers is in (i.e., glanceable or scrolling).

For glanceable context, the system collects the surrounding 300 characters minus the amount of text within the task. The remaining characters are collected evenly from above and below the task. For example, if a task had 20 characters, the context includes the 140 characters ((300 - 20) / 2) = 140) above and below the task. Because truncated sentences could

be nonsensical or confusing and overwhelm any help that the text may provide, the surrounding text is truncated to the previous full sentence above and below the task. This is to guarantee that the full sentences did not expand over the 300 character limit, which is strictly the amount of text that the view in the mobile app allows for without scrolling.

For scrolling context, the allotted number of characters is 750. Rather than truncating to the nearest full sentence, the system instead expands to the nearest full *paragraph*. This was to provide much more complete surrounding text and to have the text overflow in the mobile web app view.

A token is provided to the Word add-in at launch time to authenticate to the web app and specify which user is writing. Each microtask created in the document is communicated to the mobile web app, along with the document and user. The add-in also polls the web app for any completed microtasks matching the current user and document. When a microtask is completed by the user in the mobile web app, this populates into the Word document as a comment over the location where the microtask was inserted. If the microtask had been deleted, then the comment is inserted as a comment over wherever the inline microtask was (see Figure 1).

Web app

The web app contains an API for the Word add-in to push new microtasks and poll for completed microtasks and an interface for users to complete their microtasks on their mobile phone.

The API for the web app authenticates users via a token submitted by the Word add-in specifying the user. When given a new microtask the web app API saves this microtask with the associated user and document. When the add-in polls for completed tasks, the web app only returns tasks created and completed (not skipped) by that user for the specific document the add-in is listening on.

The interface of the web app consists of a login page, a task list page, and a task page (see Figure 2). At the login page users log in with the username and password provided to them at the end of the first section of the study. This username and password is associated with the token the Word add-in used to communicate the current user to the web app. Once logged in, writers see all their tasks in an accordion style list. Each item in the list displays the task text, truncated at overflow. Users can select a task, bringing them to the task view page.

The task view page has a 300 pixel by 250 pixel frame that holds the surrounding text (local context) of the task. This is either overflowed, in the case of scrolling context, or not, in the case of glanceable context. The task itself also is included in this context, bolded to denote it as the current task. Below the local context frame is either the prompt of the document or a link to the document, depending on the global context condition.

Writers can complete a task by writing in the provided text box and hitting 'submit'. Alternatively, they can skip the task by hitting the 'skip' button below. If writers decide to skip they were prompted to select why they skipped from a selection box, or write their own reason in. When a task is skipped, it would move down to a skipped task List below the task list on the task list page (Figure 2). Users can select these tasks, where their skip reason is displayed along with the task, and can complete these tasks then. Only completed tasks (not skipped or unfinished ones) get populated back to the document.

Methodology

Before beginning the study every participant was given a brief overview of the study. They then signed a consent form outlining the potential risks and benefits of participation.

Initial in-lab writing session

After agreeing to the study, participants began with the 15-minute in-lab writing session. At the beginning of the writing session participants filled out a brief survey covering their writing and productivity habits. Questions included their English fluency, what they often write, their level of education, and how often they use mobile-based productivity tools.

Writing Task. Participants were instructed to write about a place they had lived in or visited, for readers who are potential visitors to this place. This writing prompt was chosen to provide participants with an interesting topic they were personally invested in and therefore more motivated to write about. Each participant used Microsoft Word 2016 and Microsoft SurfaceBooks to write their document in order to control for word processor version and keyboard differences. Participants were not allowed to use any other program on the computer or access the Internet during the writing session. The writing prompt asked participants to include relevant statistics and information, such as the population of the area, interesting surrounding destinations, or directions. These instructions were aimed at encouraging participants to leave notes for themselves to later search out information and include it in the document via microtasks. Participants had 10 minutes to write and were encouraged to write nonstop.

Microtask Creation. After the 10-minute writing session was over, participants were given 5 additional minutes to convert these notes into microtasks for themselves to complete on their phone over the course of the next week and/or add new microtasks. They were instructed on how to leave these notes so a microtask would be created, which involved placing double brackets ('[[ ]]') around the tasks. In the document this would change the text to be italicized and bolded (e.g., [[check this]] to ***[[check this]]***). For editing microtasks (e.g., rewriting an awkward sentence) participants were instructed to include the text to be edited in the brackets. Participants in the baseline condition were given the additional 5 minutes to leave tasks in the document however they liked to complete later using Word Mobile. Participants were also told that at the end of the week they would have 5 minutes to finalize their document in a second writing session.

At the end of the writing session, participants rated their documents on a 5-point Likert scale (1=Not at all, 5=Extremely) for coherence, grammaticality, creativity, and how entertaining the document was. These questions were drawn from past work comparing document quality in crowdsourcing [16].

Each participant completed the writing session with a researcher present who answered any questions they might have. After finishing the 15-minute writing session, participants were given a URL, username, and password that they could use to complete their microtasks on their mobile device and were instructed to save the URL to a bookmark or shortcut on their phone for easy access. Baseline participants were given a link to a Microsoft Word document and instructed to only access and edit it on a mobile device.

Editing on mobile device
Following the in-lab writing session, participants were given a week to complete writing tasks on a mobile device.

Participants in the baseline condition were given access to their document directly to complete any tasks they had left for themselves. Participants used Microsoft Word Mobile to complete tasks in the document. All completed tasks were immediately updated in the document, and participants were allowed to edit the document on their phone in any other ways they wished over the course of the week.

Participants in the microtask condition used our microtask writing system to complete their writing tasks (Figure 2), which we described in the previous section. Each participant had a different level of local or global context that defined their interactions with these microtasks. Participants logged into the microtask writing system with a username and password provided to them after the in lab writing session. Once logged in, participants saw all their microtasks and could select them in whatever order they desired (see Figure 2). After selecting a task, participants saw the task, their assigned level of local and global context, and a text box to respond to the task (for more details, see Section 3.3).

We gathered usage data on the website, including how often users logged in, viewed microtasks, how many microtasks they completed, how many they skipped, how long they spent on each microtask, and how much they wrote for each task. Participants were reminded over email 2–3 times over the course of the week to complete tasks. These reminder emails were always sent at 3 pm in order to catch participants before they left work for the day.

Final in-lab writing session
Following the week of completing microtasks, participants were given another brief writing session of 5 minutes. Participants came into the lab for this second session. All completed microtasks were included as comments in the document, and all skipped or unfinished microtasks remained in the document as participants had left them following the first writing session. Participants were instructed to finalize and polish their document, using whatever completed microtasks were helpful and removing all inline tasks. During this session participants were given access to the Internet.

Participants then completed an exit survey with questions on the ease of completing writing tasks based on the difficulty of use and mental burden questions in the User Burden Scale [25], ratings on documents quality (as described in Section 3.4.1) and a short informal interview about the tool and their ability to complete tasks. All participants were debriefed with information about the goals of the study and additional resources if they wanted to learn more.

Analysis
Each step of the study focused on answering one of our research questions. At each stage of the study writer behaviour (e.g., creating microtasks or editing the document) and document quality were compared between writers using the microtask system and those editing the document directly using Word Mobile (our baseline condition). Additional comparisons across the different context levels given to participants are also reported in the results.

Initial writing session: The analyses for the initial writing session focused on answering **RQ1**: What tasks do writers create for themselves to complete on a mobile phone?

After the first in-lab writing session the number of tasks created by each participant was saved and all participant microtasks were categorized manually into task type. Task type was broken down into two main categories adapted from [15]: inserting and editing content. In addition, initial document length, content and writer's self-ratings of document quality were saved.

Completing microtasks: The analyses for the the week-long period where participants edited the document (either by completed microtasks or by editing the document directly) focused on **RQ2**: How do writers complete tasks on a phone and how does this compare with editing directly on a phone?

The number of tasks users completed and skipped out of the total number they created was compared across conditions. We additionally analyzed the amount of time it took for participants to complete tasks and the length of responses for tasks. We defined time to complete a task as the difference between when a user viewed a task for the last time and when that task was submitted. Finally, we analyzed usability metrics of completing tasks (either on the web app or Word Mobile) based on questions from [25] (e.g., "The website demanded too much effort," rated on a 5-point Likert scale).

Final writing session: The final writing session allowed us to focus on **RQ3**: How does microtasking impact final document quality?

Final documents were saved following the final writing session. Each document was compared to its earlier draft, saved after the initial writing session. To characterize how a document changed, length differences and edit distance were calculated between the document snapshots. Edit distance, known as Levenstein distance [21], is a string matching method that calculates the number of edits (defined as deletions, additions, or substitutions) required from changing one string to another (e.g., the edit distance from 'banana' and 'nanas' is 3). Initial ratings of document coherence and grammaticality were compared with final ratings for each document.

Because the tested variables were not normally distributed or were ordinal data, the Mann-Whitney *U* test [19] was used for all statistical tests. All tests use the Benjamnini-Hochburg

correction for multiple hypothesis testing [2]. Effect size is reported as $r = Z/\sqrt{N}$ where $N =$ samples. Small, medium and large effects are observed at $r > .1$, $r > .3$, and $r > .5$.

RESULTS

Initial writing session

After analyzing our data we found that context levels displayed in the microtask writing system did not significantly impact completion rate, response length, or response time for microtasks. Because of this, we collapse across context conditions for the following analyses. We report on participant response to context levels and discuss the practical implications of these reactions for guidelines on providing context to authors for completing microtasks. A results summary is in Table 1.

Microtask creation
Writers created a median of of 8 tasks (std=3.41), with 149 tasks total using the microtask writing system. Writers in the baseline condition created a median of 5 tasks (std=1.72), with a total of 34 tasks. This was highly variable across participants, with some making only 3 or 4 tasks, while others creating upwards of 12 and 13.

Most tasks were insertion tasks for microtask writers (median=67%, std=26%) and baseline writers (median=67%, std=35%) with most tasks inserting small amounts of information that writers could not remember while writing (e.g., adding opening times for an attraction or the population of a city). Adding sentences or including more descriptions (e.g, adding a sentence about a place, or a concluding sentence) were also common. Editing tasks, which constituted the rest of the tasks for microtask (median=33%, std=26%) and baseline writers (median=33%, std=35%), focused on rewriting sentences, fixing typos, or in some cases editing and rearranging full paragraphs. For both editing and insertion tasks, most tasks required only surface level adjustments (e.g., edit a sentence or add in a fact). There was a small subset (<10) of tasks that required more in-depth writing (e.g., rewriting a paragraph or adding a conclusion).

Because we were interested in only understanding how writers created tasks but did not expect differences across any of the conditions (since all writers were able to create tasks while writing), we did run statistical tests to compare the number or type of task created across different participants.

Document state after the initial writing session
At the end of the initial authoring session, the average document length was 305.75 (std=75.30) words, corresponding to a little over half a page of single spaced text. Initial document length was similar between participants who left microtasks for later completion using the web add in (mean=308.58, std=71.76) and those in the baseline condition (e.g., who would edit the document directly later) (mean=302, std=96.27). Some participants in the baseline condition left comments on the document. For these participants we added the length of comments to the total word count of the document.

After the end of writing in the initial session, most writers rated their documents as somewhat coherent (median=3, std=0.49) and grammatical (median=3, std=0.75), while most writers
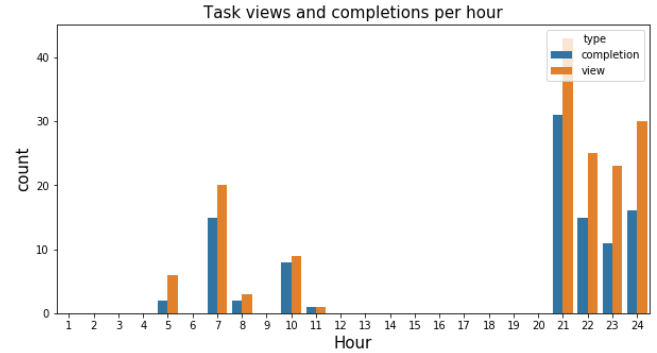

Figure 3: Task views throughout the day for microtasking participants. Participants editing the document directly followed this pattern but with edits done in the morning between 9 and 11 AM.

found their documents a little creative (median=2, std=1.02) and entertaining (median=2, std=0.8). All ratings were based on a 5-point Likert scale. The low scores for documents on creativity and entertaining suggest that participants did not feel like they had the time for spending effort to create exciting documents, and overall felt only OK about what they were able to complete in the 15 minutes given (shown by the middling scores for coherence and grammar).

Mobile editing

Task Completion while Mobile
Participants using the microtasking system viewed and completed tasks generally in the afternoon and evening (see Figure 2, reflecting participant responses that they completed tasks while on their commute. Participants editing the document directly (the baseline condition) generally followed this pattern, but with most edits done in the morning between 9 and 11 AM. The authors in the microtask condition completed an average of 73% of their tasks (median=75%, std=21%). The majority (34 out of the 44) of tasks participants did not complete they instead skipped (as opposed to never responding). Most of these skipped tasks were due to the tasks being unnecessary (14 tasks). Participants in the baseline condition, who completed tasks directly in the document, on average completed 94% (median=100%, std=11%) of their tasks. One baseline participant did not complete any tasks, and so was not included in the task completion analyses.

Overall the average length of responses was 12.99 words (median=6, std=22.02). A few microtasks generated substantial responses (greater than 100 words), leading to the high variance in response length. Microtasking participants wrote significantly more for their tasks (median=8, std=23.16) compared to baseline participants (median=2, std=13.09) ($U = 670.0, Z = -2.27$, $p < .05$, $r = 0.20$). Participants also tended to write more for adding tasks (median=9 words, std=23.41) than for editing tasks (median=5, std=22.36). The short length of responses overall suggests that users sought to write minimally on their phone.

Table 1: Summary of results. All results are reported as median (std) across participants. * = p<0.05.

Stage	Result	Microtasking	Baseline
Creation	# Tasks	8.00 (3.41)	5.00 (1.94)
	% Add	67% (26%)	67% (35%)
	% Edit	33% (26%)	33% (35%)
Completion	% completed	75% (21%)	100% (11%)
	Response length (# words) *	8.00 (23.16)	2.00 (13.09)
	Response time (# seconds)	58.00 (212.46)	NA
Integration	**Length difference (# words) ***	60 (116.70)	28 (17.76)
	Edit distance (# edits) *	460 (617.99)	308 (106.33)
	Coherency difference (Initial, Final)	**Initial=3, Final=4 ***	Initial=3.5, Final=3.5

Microtask users on average took 126.70 seconds (median=58, std=212.46) to complete tasks. Participants took slightly longer to complete insertion tasks (median=60, mean=129.87, std=195.97) than editing tasks (median=55, mean=92.47, std=163.13). Participants also varied widely in how long it took them to complete tasks, ranging from taking a median time of 26 second per task to taking 210 seconds per task. We did not have telemetry data that specified the time it took baseline participants to complete tasks using Word Mobile.

How context impacted task completion. Participants were able to complete tasks regardless of the amount of context they were given (e.g., glanceable or scrolling and document link or simply the prompt); however, participant responses highlight tradeoffs between the amount of surrounding text to give for tasks. 8/10 participants given scrolling local context reported never feeling like the microtask system web app forced them to remember too much information. This was also shown in participant responses saying they enjoyed the amount of context given for each task: P11 said, " having just a few lines visible while being able to scroll through the rest and access the entire document with one click was super convenient." However, 4/10 participants in the scrolling condition also said that the web app presented *too much* information at least a little bit of the time, while 8/10 participants given glanceable context (much less surrounding text to navigate on the phone) reported never feeling that the web app presented too much information. These responses highlight a tradeoff between providing more or less surrounding text for completing tasks. While participants were able to complete tasks regardless, it is important to recognize when the added burden of more text could be helpful, possibly leaving it up to writers when to be given more text while completing mobile microtasks.

The 10 participants who could access the document from the microtask system (separate from the baseline participants who edited the document directly on their phones) viewed their document 11 times for their 77 total tasks. While most participants did not mention the document being helpful for completing tasks, P11 and P23 felt that the document was helpful for checking some task context more than immediate surrounding context. Tasks for which the participants viewed the document did not differ noticeably from other types of tasks (e.g., surface level insertions and edits), but participants did take much longer for completing these tasks (median=169.5 seconds,

std=314.35 seconds) compared with other tasks (median=58 seconds, std=212.46).

Final writing session
In the final writing session, all participants (baseline and microtasking) were given 5 minutes to finalize the document. Participants in the microtasking conditions had to use this time to integrate any completed tasks (populated to the document as comments) into the document, whereas the edits made by the baseline participants were already in the document.

Document Evolution
Participants' documents generally were longer after this final session, with an median increase in length of 53.5 words (std=106.18 words) from the initial writing session. This increase was significantly larger for microtask writers (median=60, std=116.70) compared with baseline writers (median=28, std=17.76) ($U = 15.0$, $Z = -2.31$, $p < 0.05$, $r = 0.47$). This increase in length also came with more substantial changes in the document: microtask writers wrote final documents that had a significantly higher edit distance between final and initial draft (median=460 edits, std=617.99) compared with baseline writers (median=308 edits, std=106.33) ($U = 18.0$, $Z = -2.10$, $p < .05$, $r = 0.43$).

Self-Rated Document Quality. These edit differences also lead writers to feel more confident in their documents. Writers who used the microtask system had a significant increase in their rating of document coherence from the initial document rating (median=3) to their final document rating (median=4) ($U = 88.0$, $Z = -2.70$, $p < 0.01$, $r = 0.44$), while writers who edited the document directly did not (initial median=3.5, final median=3.5) ($U = 9.0$, $Z = -0.73$, $p = 0.24$, $r = 0.23$). Coupled with the increased editing and length changes in the document, this suggests that microtasking encouraged more refining of the document compared with the baseline. There were no significant differences in self-rated creativity, entertainment, or grammaticality across initial and final documents.

Document Engagement. The difference in document editing was also easily observed in the final writing session. Participants in the baseline condition rarely used the entire 5 minutes to edit their documents, while participants using the microtask system almost always did. Past work has shown that microtasks help users ramp up for completing more complex tasks [4]. We found some evidence of this: participants who used the microtasking system felt that the tasks had helped

them get started writing again, leading to more edits and higher confidence in the document. P1, for example, found the tasks, "helped prime me to keep writing. [It was] nice to have easily marked places to jump to, with content, instead of having to remember where the weak parts were."

Participant feedback
The participant interviews and ratings of the system contextualized their writing and behaviour during the study.

Task deferral allowed better flow while writing. Participants liked to defer tasks for a later time, P21 created tasks when, "wanting to spend more time thinking about the structure of the sentence," and as P19 put it, "adding tasks helped me focus more on writing rather than getting distracted with internet search. it was nice to get back to that later and search for information during a bus ride." One participant liked the way of deferring tasks so much that they continued using the Word add-in after the study was complete.

Participants create tasks suitable for mobile interactions. Many of the participants' tasks fit well with the constraints of mobile phones. The majority of tasks were simple insertion and editing tasks, with only a few tasks that required extensive writing for most participants. This was also reflected in participants' completed tasks. For example, P2 said, "while I created all kinds of tasks for myself, I only wanted to tackle those that I could do on my phone, leaving the more involved ones for desktop writing." This was generally due to participant's aversion to typing or navigating excessively on a phone. As P12 said, "I found that [completing tasks] hard since it required a lot of clicking and typing om [sic] my phone."

Just-right context with the Microtasking system. Generally participants found it easier to use the microtask interface than navigating the document directly. P18 (a baseline writer) reported, "It's hard to do it on the phone, I would open in [sic] on desktop if it was a work related document." While P21 and P20, both users of the microtask system, found the interface "very easy" and "... intuitive and the context around the tasks helped." These responses were supported by participants' rating of the usability of the microtask system, calculated using the difficulty of use questions in the User Burden Scale [25]. Baseline participants rated their mobile writing experience as significantly more difficult (median=2, std=0.97) than those using the microtask system (median=1.00, std=0.75) (on a 5-point Likert scale, 1 being the easiest) ($U = 991.5$, $Z = 2.20$, $p < .05$, $r = 0.23$).

DISCUSSION
In this paper we set out to characterize the mobile microtask writing process. What kinds of tasks do writers create for themselves to complete on a mobile phone, how do writers complete tasks on a mobile phone, and how does microtasking impact final document quality? In addition, writing tasks require context to complete [24], and so we set out to identify what sort of context is necessary for mobile microtasks in a single author document.

We answered these questions through a controlled, week-long user study where writers had an initial writing session to begin writing a document and create microtasks. Following the writing session, participants either completed microtasks over a week in the wild using a mobile microtask system (Figure 1) or edited the document on their mobile phone. Participants then integrated these completed tasks into their document in a final writing session.

Writers used microtasking differently, though overall they found deferring some actions as microtasks to be useful for staying focused on writing. Writers mostly created tasks that were for inserting snippets of information (e.g., a fact that the writer couldn't remember), adding short sentences (e.g., writing an introductory sentence), or editing surface-level issues (e.g., typos or unclear sentences). These tasks were well-suited for completing on a phone since they required minimal writing and context from the document. Writers took advantage of the mobile experience and micromoments [26, 13], such as while on their commute or waiting for a meeting, to complete tasks. At the same time, writers found navigating the document or completing writing intensive tasks difficult on a phone due to the small screen and keyboard (P2 and P12). This was also reflected in writers finding the microtask system, which restricted document context to immediate surrounding text, easier than editing the document directly in Word Mobile.

Authors who used the microtasking system also got more done on the document while mobile compared with authors editing the document directly. Authors wrote more for their tasks using the mobile microtasking system compared with authors using Word Mobile. Writers using the mobile microtasking system also enjoyed the few lines of context given and possible access to the document (P11), while writers using Word Mobile often found the experience of completing writing tasks difficult. This preference was also reflected in the final document. Writers using the microtasking system edited their document significantly more over the course of the entire study, leading to a significant increase in self-rated coherence compared with writers editing directly using Word Mobile. These results point to the exciting impact of microtasking in writing documents. Not only did microtasking help authors write smoothly, but it also lead to significant document improvement.

One explanation for this is that microtasking helped writers re-engage with the document quicker during the final writing session. The microtasking system populated completed tasks into the document as comments tied to the inline microtask left by the writer. During the final 5-minute writing session writers integrated these comments into their document. By having to integrate their completed tasks into the document, writers had a smooth transition back into the document, going from simple task completion (e.g., completing and integrating microtasks) to more complex tasks (e.g., editing the entire document). Our findings that this ramping up lead to more editing compared with baseline writers (where they only had the complex task of editing the document) is in line with past work on task ordering to help users ramp up to complex tasks [4].

Writing is made up of many cognitive processes [8]; however, many current writing tools help writers only with the revising process [8, 10]. This study offers an example of how writing tools can support writing processes at a higher level than purely

Table 2: Summary of main findings

Research Questions	Main findings	Section
RQ1	Writers create surface-level insertion and editing tasks most often, and these are the main tasks they feel comfortable completing on a mobile phone.	4.1
RQ2	Writers complete most tasks they assign themselves, with generally short responses reflecting the constraints of a mobile phone. Writers tend to write more for tasks using a microtask system rather than editing the document directly.	4.2
RQ3	Writers who used the microtasking system edited their document significantly more than writers who instead edited the document directly over the week. This increased editing was also associated with higher ratings of document coherence.	4.3

editing and formatting. The microtask task system gave writers control over creating, completing and integrating their own tasks, a novel contribution to writing microtask systems [22, 26, 13]. Writers were able to break down the complex task of writing into smaller subtasks, creating and deferring tasks to a later time for completion, encouraging a hierarchical writing process [8]. During the final writing session writers integrated these tasks back into the document, showing them how they could break down their writing and build it back up. Based on writer feedback and measures of document edits, it seems that writers enjoyed and benefited from this process. One participant liked the process so much, they continued using the tool after the study was completed.

Based on our findings we outline design recommendations below to leverage these benefits in mobile microtask writing:

Constrain the task completion interface Our results highlight tradeoffs in mobile productivity [1]; while the mobile experience allows for much more flexibility in completing tasks, it is important to keep in mind the system's constraints. Forcing participants to navigate an entire document, even one that is less than a page, makes completing tasks much more difficult. Surfacing tasks that take into account the constraints of a mobile system by, for example, providing restricted surrounding text and a simple interface for writing a response, leads to an easier and more effective microtasking experience.

We found this to be the case in our own study. The microtask system surfaced writing tasks, allowing writers to complete tasks without navigating through the entire document. Because most tasks did not require deep document context, the immediate surrounding text that the microtask system provided was usually enough to complete tasks.

Integrating completed tasks Writers completed tasks in varying ways, sometimes responding with a single word or a full paragraph. To address the variance in responses, adding completed tasks as comments to the document or in some other way that allows writers to integrate the tasks on their own is useful. This is especially important when authors don't have direct access to the document to edit while microtasking, since they might not be able to edit surrounding text in the document to fit with their response. In addition, we found that this integration had the added benefit of encouraging writers to engage more with their writing, leading to more edits and clearer documents compared with those who edited the document directly instead of microtasking.

LIMITATIONS & FUTURE WORK
This study did not control for task type, only characterizing the tasks after they were created. There was some indication based on participant interviews that different task types (like editing ones) required more surrounding context. A future study could explore the impact of immediate surrounding context while controlling for task type. This would be a strong extension to our findings on precedence of mobile phone constraints over amount of context for completing mobile writing microtasks. This study explored single-author documents only, but collaborative writing leads to interesting possibilities for microtasks [26] since the author who creates a task could be different from the one who completes it. An exciting extension of this work is to explore how writers collaborate on the same document, creating and completing microtasks between them.

CONCLUSION
This paper presents findings from a controlled, week-long user study to characterize how writers use mobile microtasks throughout authoring a document. Authors created microtasks for editing and inserting information that generally required minimal writing and context, making these tasks well suited for mobile devices. Microtasking writers edited their document and felt that their document improved in coherence more than writers editing their document directly on Microsoft Word Mobile. Writer responses to microtasking suggest that the experience of creating, completing, and integrating microtasks helps writers stay more focused on writing, easily complete writing tasks in short moments (e.g., while commuting), and get back into writing their document more easily. Based on these results, we discuss design recommendation for building writing tools that support writers' processes beyond editing and revising.

REFERENCES
[1] Patti Bao, Jeffrey Pierce, Stephen Whittaker, and Shumin Zhai. 2011. Smart phone use by non-mobile business users. In *Proceedings of the 13th international conference on human computer interaction with mobile devices and services*. ACM, 445–454.

[2] Yoav Benjamini and Yosef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57 (1995), 289–300. DOI:`http://dx.doi.org/10.2307/2346101`

[3] Michael S Bernstein, Greg Little, Robert C Miller, Björn Hartmann, Mark S Ackerman, David R Karger, David Crowell, and Katrina Panovich. 2010. Soylent: a word processor with a crowd inside. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. ACM, 313–322.

[4] Carrie J Cai, Shamsi T Iqbal, and Jaime Teevan. 2016. Chain reactions: The impact of order on microtask chains. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 3143–3154.

[5] Justin Cheng, Jaime Teevan, Shamsi T Iqbal, and Michael S Bernstein. 2015. Break it down: A comparison of macro-and microtasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 4061–4064.

[6] Elizabeth Clark, Anne Spencer Ross, Chenhao Tan, Yangfeng Ji, and Noah A Smith. 2018. Creative writing with a machine in the loop: Case studies on slogans and stories. In *23rd International Conference on Intelligent User Interfaces*. ACM, 329–340.

[7] Edward Cutrell and Zhiwei Guan. 2007. What are you looking for?: an eye-tracking study of information usage in web search. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 407–416.

[8] Linda Flower and John R Hayes. 1981. A cognitive process theory of writing. *College composition and communication* 32, 4 (1981), 365–387.

[9] Victor M González and Gloria Mark. 2004. Constant, constant, multi-tasking craziness: managing multiple working spheres. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 113–120.

[10] Nick Greer, Jaime Teevan, and Shamsi T Iqbal. 2016. *An introduction to technological support for writing*. Technical Report. Technical Report. Microsoft Research Tech Report MSR-TR-2016-001.

[11] Nathan Hahn, Shamsi T Iqbal, and Jaime Teevan. 2019. Casual Microtasking: Embedding Microtasks in Facebook. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 19.

[12] Shamsi T Iqbal and Brian P Bailey. 2005. Investigating the effectiveness of mental workload as a predictor of opportune moments for interruption. In *CHI'05 extended abstracts on Human factors in computing systems*. ACM, 1489–1492.

[13] Shamsi T Iqbal, Jaime Teevan, Dan Liebling, and Anne Loomis Thompson. 2018. Multitasking with Play Write, a mobile microproductivity writing tool. In *The 31st Annual ACM Symposium on User Interface Software and Technology*. ACM, 411–422.

[14] Tero Jokela, Jarno Ojala, and Thomas Olsson. 2015. A diary study on combining multiple information devices in everyday activities and tasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3903–3912.

[15] Harmanpreet Kaur, Alex C Williams, Anne Loomis Thompson, Walter S Lasecki, Shamsi T Iqbal, and Jaime Teevan. 2018. Creating Better Action Plans for Writing Tasks via Vocabulary-Based Planning. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 86.

[16] Joy Kim, Sarah Sterman, Allegra Argent Beal Cohen, and Michael S Bernstein. 2017a. Mechanical novel: Crowdsourcing complex work through reflection and revision. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 233–245.

[17] Jaewon Kim, Paul Thomas, Ramesh Sankaranarayana, Tom Gedeon, and Hwan-Jin Yoon. 2017b. What snippet size is needed in mobile web search?. In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*. ACM, 97–106.

[18] Aniket Kittur, Boris Smus, Susheel Khamkar, and Robert E Kraut. 2011. Crowdforge: Crowdsourcing complex work. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 43–52.

[19] Henry B Mann and Donald R Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics* (1947), 50–60.

[20] Gloria Mark, Victor M Gonzalez, and Justin Harris. 2005. No task left behind?: examining the nature of fragmented work. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 321–330.

[21] Gonzalo Navarro. 2001. A guided tour to approximate string matching. *ACM computing surveys (CSUR)* 33, 1 (2001), 31–88.

[22] Michael Nebeling, Alexandra To, Anhong Guo, Adrian A de Freitas, Jaime Teevan, Steven P Dow, and Jeffrey P Bigham. 2016. WearWrite: Crowd-assisted writing from smartwatches. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 3834–3846.

[23] Antti Oulasvirta and Lauri Sumari. 2007. Mobile kits and laptop trays: managing multiple devices in mobile information work. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 1127–1136.

[24] Niloufar Salehi, Jaime Teevan, Shamsi Iqbal, and Ece Kamar. 2017. Communicating context to the crowd for complex writing tasks. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 1890–1901.

[25] Hyewon Suh, Nina Shahriaree, Eric B Hekler, and Julie A Kientz. 2016. Developing and validating the user burden scale: A tool for assessing user burden in computing systems. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. ACM, 3988–3999.

[26] Jaime Teevan, Shamsi T Iqbal, and Curtis Von Veh. 2016. Supporting collaborative writing with microtasks. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. ACM, 2657–2668.

[27] Alex C Williams, Harmanpreet Kaur, Shamsi Iqbal, Ryen W White, Jaime Teevan, and Adam Fourney. 2019. Mercury: Empowering Programmers' Mobile Work Practices with Microproductivity. (2019).

[28] Diyi Yang, Aaron Halfaker, Robert Kraut, and Eduard Hovy. 2017. Identifying Semantic Edit Intentions from Revisions in Wikipedia. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Copenhagen, Denmark, 2000–2010. `DOI: http://dx.doi.org/10.18653/v1/D17-1213`