

# **Mechanics of Document Formatting:**

## **Comparing strategies between undergraduates and graduates**

Jennifer Lui, Graduate Advisor: Sonia Arteaga, Research Advisor: Sri Kurniawan

UCSC SURF-IT 2011

### **Abstract**

We conducted a user study to investigate the strategies implemented by undergraduate and graduate students while formatting a Microsoft Word 2010 document. Each participant was recorded while they edited a document with formatting and layout errors using the Think Aloud Method. After transcribing the recordings, Open-Coding was used to extract themes associated with how people format documents. Next, these themes were incorporated into Thematic Networks, a method that displays a hierarchy of themes, which shows how global themes were developed. As a continuation of this qualitative data analysis, we developed a Hierarchical Task Analysis that demonstrated the process participants took to format the document. When comparing the two groups, we found that there were differences and similarities in their formatting methods. Graduates were more efficient in fixing inconsistencies, but detected fewer because they did not read the content. Otherwise, the direction and the reasoning behind making changes were similar. Results from this study will provide insight on the design of a worker prompt for Mechanical Turk workers who will format documents produced by blind people.

### **Author Keywords**

HCI; Microsoft Word; document formatting; formatting strategies; qualitative task analysis; Amazon Mechanical Turk

### **Keywords**

Human Computer Interaction (HCI), word processor, formatting

### **Introduction**

Understanding how people format documents is important for developing informative prompts that will result in well-formatted documents returned by workers in work-force environments such as Amazon Mechanical Turk. To our knowledge, few studies exist on how people edit documents. In order to develop applications that will use a work force in which direct one-on-one contact is not feasible, the clarity of how we request formatting and how we convey expectations to the workers are essential to perform the task correctly. Applications that can make use of such knowledge are quick editing jobs, formatting jobs, and automatic formatting requests for the purpose of assisting people with disabilities. This is similar to designing writing prompts for tests, which have been treated with special attention since it produces work that will

be evaluated [1]. People can easily interpret a topic in a variety of ways so test writers cannot assume that every reader is coming across the same thought. The goal is to have the responder match the meaning envisioned by the questioner.

To address and add to the body of knowledge on current formatting techniques using Microsoft Word 2010, we designed a user study to answer questions about formatting mental models and techniques that we had. In particular, we were interested in answering how education level affected formatting strategies, specifically between undergraduates and graduates. To explore their mental models regarding why these students changed certain text and how they changed it, we analyzed video recordings of their voice and screen movements/actions.

Participants were presented with a paper that had inconsistencies with font and position of text, which were found to be commonly made by blind users [2]. That paper explored them in order to develop tools to assist blind writers in the creation of electronic documents. It found that the writers cared about appearance of their documents, but they ineffectively used tools in the word processor such as spell checker and templates.

Direction for this research study followed a paper that showed the importance of designing user studies to ensure better feedback from Amazon Mechanical Turk (AMT) workers [3]. User studies are important because they can help formulate a better procedure for later studies and they can improve collection of desired results. That paper ran tasks on Amazon Mechanical Turk with the goal of having workers correctly rate Wikipedia articles. By running two experiments that asked users to rate Wikipedia articles and then comparing them to ratings by expert administrators, they found out that wording the question was essential to receive ratings that were similar to expert ratings and reduce invalid responses.

AMT is an online crowd sourcing labor market where businesses post jobs called Human Intelligence Tasks (HITs) [4]. These are work that computers cannot complete such as identifying objects in photos or filtering relevant information. Businesses use this service since there is a big pool of workers who accepts low monetary rewards for doing tasks that require little effort and time. Another benefit is that the requesters only need to pay for good work.

One software that uses AMT is Soylent, which has workers assist with the revision of a paper from shortening text, fixing spelling and grammar errors, to formatting citations [5]. To produce high quality work, each follows a technique called Find-Fix-Verify, which has one set of workers find errors, another set to fix those errors, and finally one other group to pick the best revision. We are doing something similar, but we are dealing with formatting and layout.

Another program that uses AMT workers is VizWiz, which allows blind people to pose visual questions, which are answered by sighted people [6]. The blind user takes a picture with their phone, dictates a question, and then receives multiple spoken choices from workers in real-time. Examples of questions include what color is an object, what denomination is this bill, and what temperature is set on the oven.

The results of our user study will provide insights on documents edited by workers on AMT, who will receive documents created by blind people and will be expected to return correctly formatted documents. If this is successful, then blind people can use our application (currently in development) to edit their documents by having an accessible tool to send their documents to AMT since current tools do not make catching formatting errors accessible or easy.

## Methods

To answer our research questions I carried out a user study designed by our group member. This study consisted of:

- People 18 years old and older formatted a document modified by myself
- The participants were asked to use the Think Aloud method while formatting the document
- The recorded audio of the participant speaking and the video of screen movements were analyzed
- A set of questions at the end of the study which participants answered

The study was conducted in a lab test environment (meeting room where only the participant, myself, and at times another researcher were in). For the study, the participants worked on the same laptop computer, which had Microsoft (MS) Word 2010 running on Windows 7. We chose to go with this setup instead of having the participant use their own computers in their natural environment because we wanted to control for the type of Microsoft Word version they were using since this could change interaction styles and may make comparison among the participants difficult. Most of the participants were familiar and comfortable using MS Word 2010 so this was not a problem in terms of introducing effects related to having them use a version of Word they were not familiar with using. However, there were a few participants who did not use MS Word so their data were not used in our analysis. The goal of this work was to investigate how undergraduates and graduates formatted documents on MS Word 2010.

A total of seven undergraduates and seven graduates at UC Santa Cruz were recruited to perform the user study. Their age, gender, and major were not taken into account, but they were from engineering or science background. Each participant was given a document with inconsistencies and asked to correct the document for formatting, style, and whitespace to their liking. The document was two-page long and included text from George Orwell's *1984* [7]. It was divided into one chapter with one part and another chapter with two parts. Some words were changed to lessen recognition of the text. Inconsistencies were added to allow evaluation of their strategies. These types of inconsistencies were indicated as those that blind people commonly made [2]. They are:

- Alignment errors: the justification of a paragraph is inconsistent or improper
- Capitalization errors: the capitalization of multiple consecutive letters is unintentional
- Character formatting errors: text is inappropriately bold, italicized, or underlined

- Font size errors: words have inappropriate font size
- Font selection errors: the font choice makes text hard to read
- Indentation errors: the indentation of text is inconsistent
- Line spacing errors: the spacing between adjacent blocks of text is irregular
- Word spacing errors: the spacing between consecutive words or consecutive sentences is irregular

Subscripts were also included since they could also potentially occur and current technologies do not appear to help in easily detecting these.

The user looked at the document, while a screen recorder by Snagit 10 recorded the user's on-screen movements and voice. They were given as much time as they needed to complete the task of formatting the document. Each user used the Think Aloud Method to explain their changes and thinking. From the screen recording, we were able to get time stamps for the different activities. As the researcher, I observed the user and reminded them to speak out loud when the user forgot to explain their actions and thoughts or was quiet for a long time (indication that they were forgetting to use Think Aloud Method). Once the user was finished with correcting, the video file was saved. Then, using the Sound Recorder to record audio, the user was asked some follow-up questions such as:

- What kind of errors were you looking for?
- Did you have any difficulties in what or how to format?
- What was your strategy in formatting?

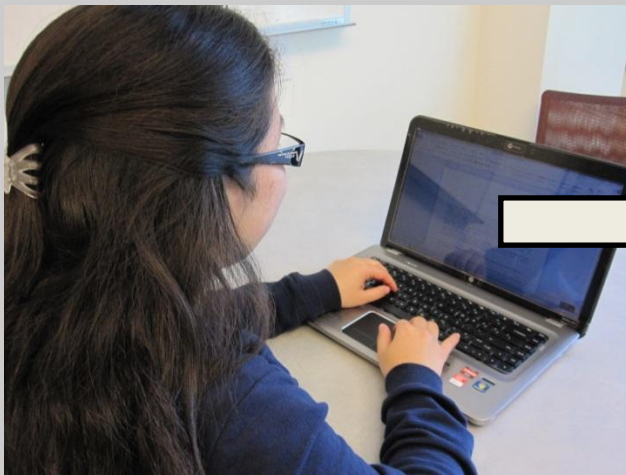
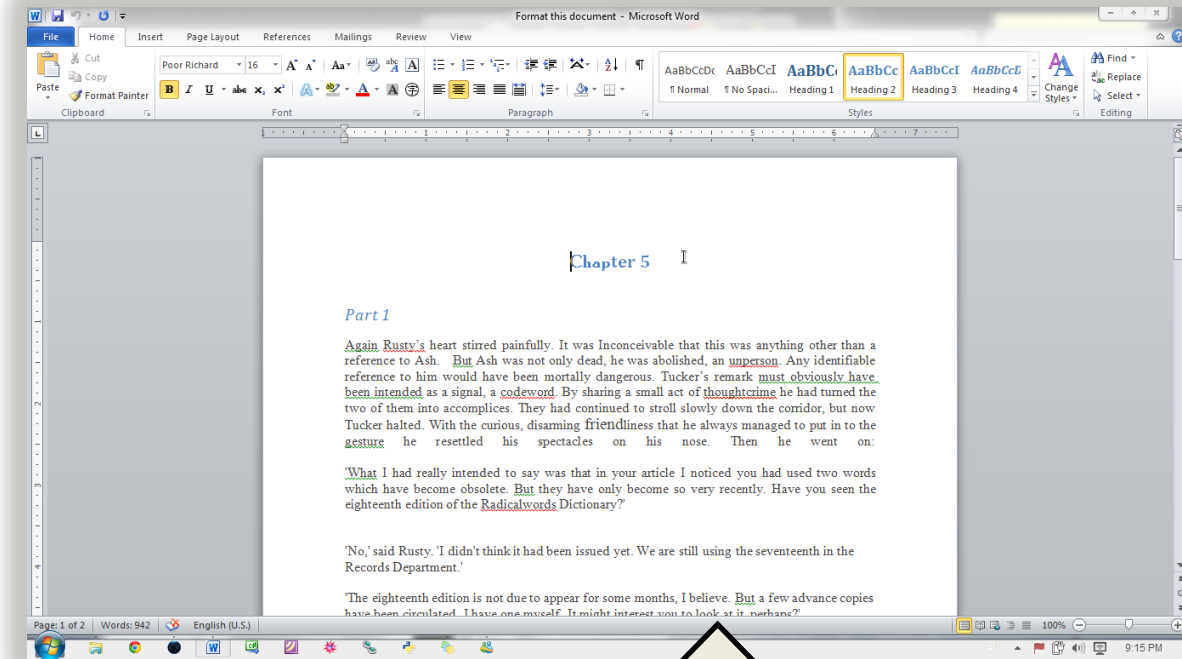
This file was also saved.

The qualitative data analysis allowed for a comparative analysis for the two groups of participants, i.e. undergrads and graduate students. I wanted to see whether graduates used a more efficient formatting strategy that would eliminate multiple inconsistencies in a few steps since being older, they would have more experience with using this word processor. The recordings were transcribed meaning that any movements or talking were noted as well as when they occurred. For example, at time 1 minute 30 seconds, participant highlighted Chapter 5 heading and chose "Times New Roman" from the Ribbon because "there are two different fonts present."

Open-Coding was then performed on the transcriptions to extract themes. These themes were then used to construct Thematic Networks. Thematic Network Analysis was done to develop web diagrams that showed how themes were generalized [8]. The themes from Open-Coding were referred to as basic themes. Organizing themes were constructed to encompass basic themes, and then a global theme encompassed organizing themes.

This was followed by the Hierarchical Task Analysis (HTA), which explained the process and tasks involved in formatting a document [9]. It can be seen as a how-to-guide that our

participants took in formatting the document they were presented with. This is developed with general steps that have subtasks. It answers the question of what a learner needs to do and think in order to achieve this goal.



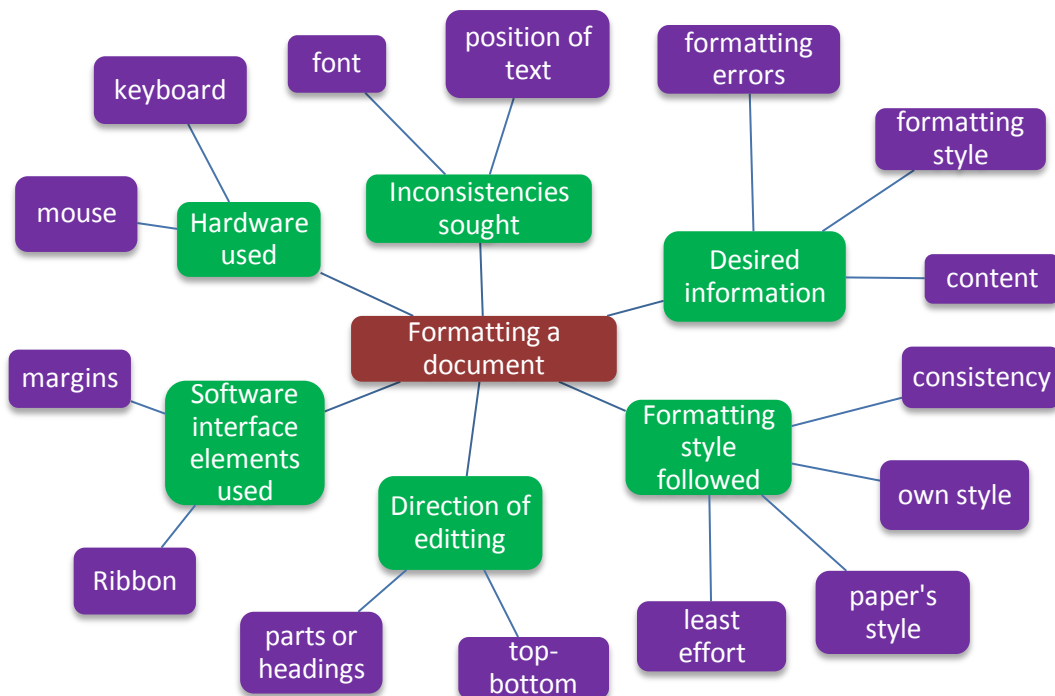
**Figure 1.** Zoomed in view of the document to be formatted.

The participant worked on a laptop with a document that had inconsistencies. As seen above, these inconsistencies included multiple font faces in the chapter heading, odd word spacing in

the last line of the first paragraph, and multiple line spacing between the second and third paragraph.

## Results

After analyzing our recordings, we came up with a Thematic Network for how undergraduates and graduates format documents (see Figure 2). We found that the participants used a bottom-up approach, which linked specific elements together by relevancy to form larger, general systems. The purple boxes represent the basic themes that were extracted through Open-Coding. These themes were grouped into organizing themes, represented by green boxes. These organizing themes were centered on the type of inconsistencies they fixed, the information they wanted to know, the style they followed, the direction in which they edited, the software interface elements used, and hardware tools they used. Then these themes were connected to a global theme, the red box, of formatting a document.



**Figure 2.** Thematic Network for themes seen in both undergraduates and graduates.

In terms of inconsistencies sought, they looked at font face, font size, font color, character format like bold, italics, and underline, and capitalization. They also looked at alignment, subscripts, line spacing, and word spacing.

Desired information included questions that came up as they were formatting. They asked what are the different types of formatting errors and if they needed to check for grammar or spelling. They also wanted to know if the style that should be followed was a book, MLA, or essay. Another question was which words should be emphasized.

For formatting style followed, they tried to make everything consistent. They either followed their style, the paper's style, or a combination of both. They also went with the style that required the least effort, meaning what was mostly present. When using the Ribbon, they hovered over different options to see which was best. An example was choosing between Heading 1 and Heading 2 in the Quick Styles. They often scrolled back and forth when trying to make one part consistent with another. This was seen with chapter and section headings. After reading the context, they decided whether an inconsistency such as something in italics was done purposely. Some of the graduates mentioned that they formatted so that the text was visually pleasing or easily readable. They were focused on efficiency, which meant fixing several inconsistencies in one step by highlighting everything and pressing buttons on the Ribbon, located on the top panel.

The direction of editing was top-bottom, which meant formatting headings and body text in the order presented in paper. Some formatted the text of the first chapter and then fixed the headings for the chapters and sections. Then they continued to the text of the next chapter.

In terms of software interface elements used, they used the Ribbon, specifically the Home tab, to fix almost all mistakes. After highlighting using the mouse, they clicked on buttons such as Bold, Italics, Underline, Subscript, Font Color, Align Text Left, and Justify. They also selected options in the drop-down menus in Font and Font Size. To change styles, they clicked on Quick Styles. Few also clicked on Line and Paragraph Spacing and the Show/Hide ¶. The Format Painter was often used by graduates, who used it to copy the format of one block of text and applied it to another. After highlighting, they used the Ribbon to check information about the format of the text. They were also able to do that by leaving the cursor in the text. This helped to verify their intuition. Some also used the Page Break under the Insert tab. Some also moved the horizontal margins to indent paragraphs. One undergraduate zoomed out to see all the pages in order to see the overall presentation, but most did not use the zooming feature. Some clicked on the words with green or red squiggles to see suggestions.

In terms of hardware used, they used the mouse to highlight text, click the buttons on the Ribbon, navigate to error, and hover over the text they were reading. The keyboard was used to highlight text, navigate to error by using arrow keys, check for information about text, and make corrections by using shortcuts.

### 1. Look through document.

- Determine what is most consistent.
- Get a feel of inconsistencies like font, spacing, and position of text.
- Decide on desired style.

### 2. Make changes.

- Look in a top-down manner.
- Refer to these guidelines to decide on how to make changes.
  - Use context.
  - See if inconsistencies were done for emphasis.
  - Look at different options before deciding.
- Use Ribbon or keyboard to fix errors.
  - Apply to headings and then body text.

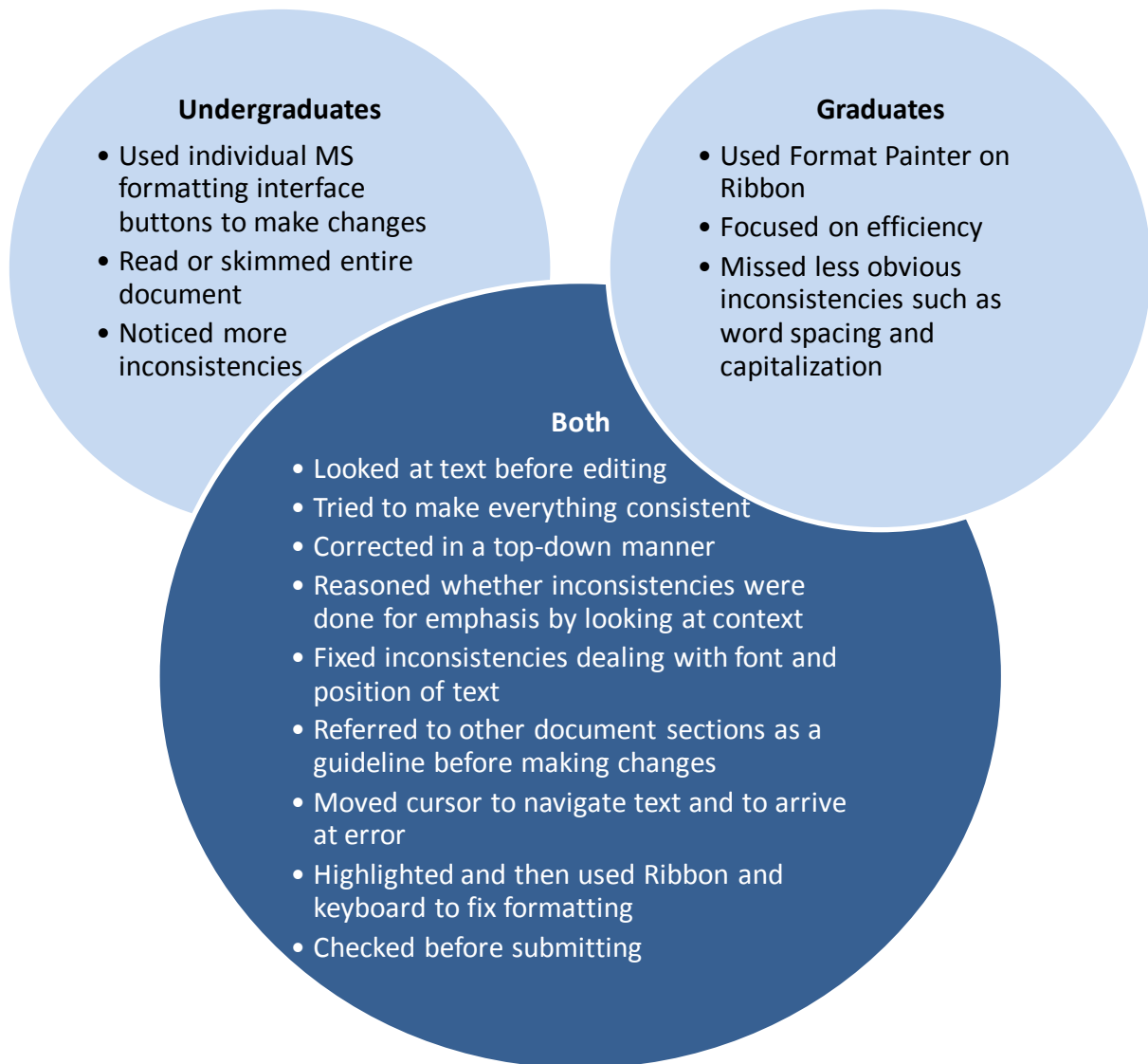
### 3. Check before submitting.

- Follow step 2 until there are no errors left.

**Figure 3.** Procedure that undergraduates and graduates took to format their document.

The representation created in Figure 3 after doing Hierarchical Task Analysis (HTA) showed the major steps participants took followed by the smaller steps. They had an overview of document before making changes, edited from top to bottom where they fix what they find wrong, and repeated this until they saw no more problems.





**Figure 4.** Similarities and differences between the approaches taken by undergraduates and graduates.

Figure 4 shows a comparison between the education levels. The dark blue circle represents the similarities while the light blue circles represent the unique themes for each group. The students had more in common with each other, than differences. The similarities can be seen through the procedure they followed while the differences laid in software use.

## Discussion

Undergraduates and graduates think alike when they format a document as seen in the Thematic Network, the Hierarchical Task Analysis chart, and the theme comparison chart. They look for similar types of inconsistencies in a top-bottom fashion. They shared the same goal of making all the text consistent to their respective parts. They used context clues or referred to other sections to determine how something should be changed. When they spotted an inconsistency, they would highlight it, and use the Ribbon to correct it. Most participants fixed inconsistencies one by one while others selected the entire text or chunks of text and applied changes to alignment or font size.

However, they used different software interface tools and different approaches to find inconsistencies. Graduates copied formatting from one text and applied it to another by using Format Painter. The Format Painter copied alignment, character format, font size, and other specificities. Undergraduates, on the other hand, applied those changes by clicking on multiple buttons on the Ribbon, compared to one button by graduates. Undergraduates read or skimmed the text, which allowed them to catch subtle inconsistencies such as multiple word spacing and capitalization, which were missed by graduates because they looked for ones that stuck out. Overall, education level does not affect the way the participants corrected documents, but may if comparing wider education levels like high school students and graduates. Graduates took an average of 15.70 minutes while undergraduates took 18.97 minutes. Most participants were not able to correct all inconsistencies.

We noticed that when people were given the freedom to edit the document in whatever way they want, their results varied with those that closely matched up to those that were totally different from the original document. Although many people went with the easiest style or the style that was presented, some followed their own style. For those latter cases, they added indents, replaced single quotation marks with double, picked different heading styles, modified font face and size, deleted paragraph spacing, and changed alignment. Some of those who decided to do that stumbled into more work and would undo them or would forget to undo them, which introduced more inconsistencies. By providing guidelines, we believe workers will return a properly edited document.

Future work will consist of using this information to develop the worker prompt for Amazon Mechanical Turk (AMT) workers related to formatting documents for blind users. The worker prompt will provide guidelines as to what type of inconsistencies should be fixed and what style should be followed. This future study will also investigate the reliability and cost effectiveness of AMT worker output.

## Acknowledgements

This research was sponsored by the National Science Foundation, SURF-IT ([surf-it.soe.ucsc.edu](http://surf-it.soe.ucsc.edu)) Research Experiences for Undergraduates Program. NSF grant Award No. CNS-0852099. In

addition, thanks to the University of California, Santa Cruz, my advisors, Sonia, and Sri, and user study participants.

## References

- [1] L. Ruth and S. Murphy, "Designing topics for writing assessment: problems of meaning," *College Composition and Communication*, vol. 35, no. 4, pp. 410-422, Dec. 1984. [Online]. Available: JSTOR, <http://www.jstor.org/stable/pdfplus/357793.pdf>. [Accessed Aug. 30, 2011].
- [2] D. G. Evans, T. Diggle, S. H. Kurniawan, and P. Blenkhorn, "An investigation into formatting and layout errors produced by blind word-processor users and an evaluation of prototype error prevention and correction techniques," *IEEE Transactions of Neural Systems and Rehabilitation Engineering*, vol.2, no. 3, pp.257-268, Sept. 2003.
- [3] A. Kittur, E. H. Chi, and B. Suh, "Crowdsourcing user studies with Mechanical Turk," in *CHI 2008 Proceedings*, 2008, pp. 453-456.
- [4] M. Buhrmester, T. Kwang, and S.D. Gosling, "Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data?" *Perspectives on Psychological Science*, vol. 6, no. 3, pp. 3-5, Jan. 2011.
- [5] M. S. Bernstein, et al, "Soylent: a word processor with a crowd inside," in *UIST 2010 Proceedings*, 2010, pp. 313-322.
- [6] J. P. Bigham, et al, "VizWiz: nearly real-time answers to visual questions," in *UIST 2010 Proceedings*, 2010, pp. 333-342.
- [7] G. Orwell, "1984," 1949, <http://www.george-orwell.org/1984/index.html>.
- [8] J. Attride-Stirling, "Thematic networks: an analytic tool for qualitative research," *Qualitative Research*, vol. 1, no. 3, pp. 385-405, Dec. 2001.
- [9] D. Aspinall, "Task analysis: HCI Lecture 5," 2007, <http://www.inf.ed.ac.uk/teaching/courses/hci/0708/lecs/tasks.pdf>.