**Introduction**

The purpose of this study was to test the interaction design of two webmail interfaces as to how well each facilitates users’ activities. We also conducted this study to see how well each interface conformed to their corresponding guidelines provided by the companies that designed them as well as to interaction design principles in general.

**Methods**

We created anonymous email accounts for both Yahoo! and Outlook with the address “ixdtest” that users would use to complete the tasks. We decided on three tasks that would allow users to interact with aspects of the interface that are considered to be used most often:

1.) Replying to an email - We populated the inboxes with emails that contained subject lines such as “Please reply to this email” and simple questions such as “What is your major?”, “What is your class year?”, and “What is 2 + 2?” that users would have to reply to. This was done prior to the study. Users were then asked to start from the inbox, bring up the original email message to the screen, and then reply to that email by answering the said questions and clicking “Send.”

2.) Recovering an email from the trash - A previously deleted email with a subject line of “Recover this email” or “Restore this email” would be in the trash before the users began the study. The users were then asked to navigate to the trash and find a way to move it to the inbox.

3.) Changing the theme - Users were asked to find the “theme” settings and change the background/wallpaper of the interface.

Only individuals who had never before used Yahoo! Mail or Outlook Mail (or at least not used them in a very long time) were asked to participate in the study. To collect data, facilitators used QuickTime Player for Mac or Screencast-O-Matic for PC to record screencasts as users completed the tasks. Audio was also recorded during the screencasts via these apps for qualitative data.

Our group agreed on a certain measuring scheme when extracting the time data from the screencasts: We started timing users from the moment they moved the mouse “toward the goal in mind” and stopped timing the users when *they* felt they had successfully completed the task. For example, some users would change the theme multiple times before deciding on one that they preferred. Our reasoning is that since this is a user-centered study, our data could have added value by reflecting user perceptions. Yes, the users objectively switched the background to a different theme for their first click, but if they did not prefer that theme, then in their minds they did not “successfully” change the theme. We believe that framing the data in this manner provides insight into how the interfaces facilitate usability.

**Data Collection**

There were two main phases to the study. The first phase involved having the users complete all three tasks for the first time without having seen the interfaces before. These recordings were used to extract data for Learnability metrics. The second phase involved having the users complete the same three tasks on each of the interfaces a second time. This second set of recordings was used to extract data for Efficiency metrics. Users were offered time after the first phase to practice the three tasks before entering the second phase of the study. During all times of the study, facilitators observed the users’ actions to take note if any Errors made at any time. We defined an Error as an event where the user clicks on something expecting one thing to happen but getting an unexpected result.

For qualitative data, users were asked to vocalize their thinking, frustrations, confusion, mastery, and so on, during the recording, and they were also interviewed after the two phases for them to explain how or why they utilized or reacted to the interface in certain ways. Post-phase interviews consisted of inquiries into how users perceived the interface: “Were there any aspects that you found confusing?”; “Were there any aspects that you found helpful?”; “Why did you \_\_\_\_ when you \_\_\_\_?”; “In general, how would you describe \_\_\_\_ ?”

Users were told which tasks to accomplish before the recording started, however, facilitators also guided users on which tasks needed to be completed and whether actions were valid or not *during* each recording.

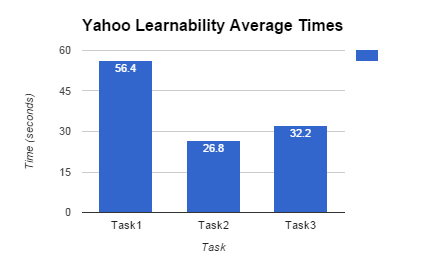
**Results**

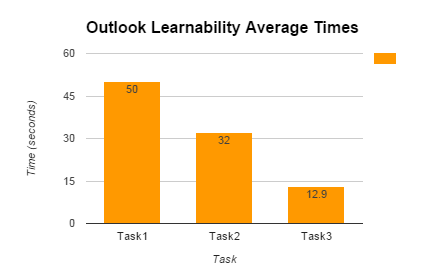
(The complete set of the data that is not averaged can be found in Appendix 1.)

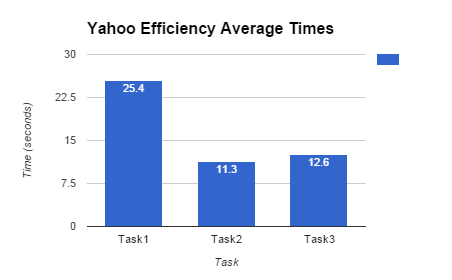
A total of 10 individuals participated in the study. Two of those participants were also facilitators, but had never used Yahoo! or Outlook before or at least had not used either in the past 5 years. Our team of 4 facilitators was together when we recorded the screencasts of the 2 facilitators who were participants. The remaining 2 facilitators conducted the post-phase interviews for qualitative data and, at a later date and time, also extracted the time data from those screencasts.

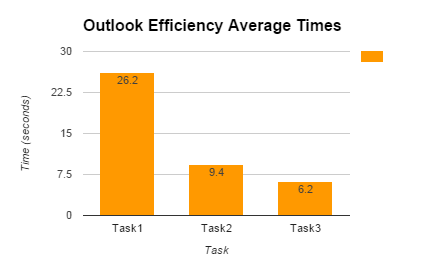
Each survey with a participant averaged between 5 - 12 minutes to record the tasks for both phases of the study as well as interview that person about their experience and perceptions. According to the data, there was no outright “winner” as to which interface facilitated these specific tasks more effectively. The graphs below show the averaged times spent for each task by participants, where Task 1 is Replying to an Email, Task 2 is Recovering an Email from the Trash, and Task 3 is Changing the Theme of the interface.

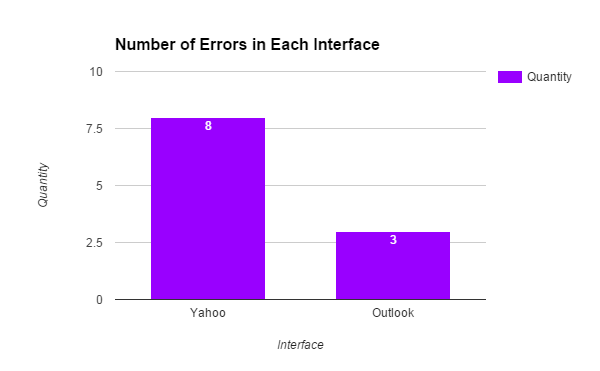
From these visuals we can see that in terms of average Learnability, participants spent less time on Task 1 and 3 in Outlook than in Yahoo! but spent less time on Task 2 in Yahoo than in Outlook. It is also interesting to note the difference between the Learnability and Efficiency results and how the times for each task were reduced by about half for Task 1 and by several factors for Tasks 2 and 3. These averages also include outliers in the data set.





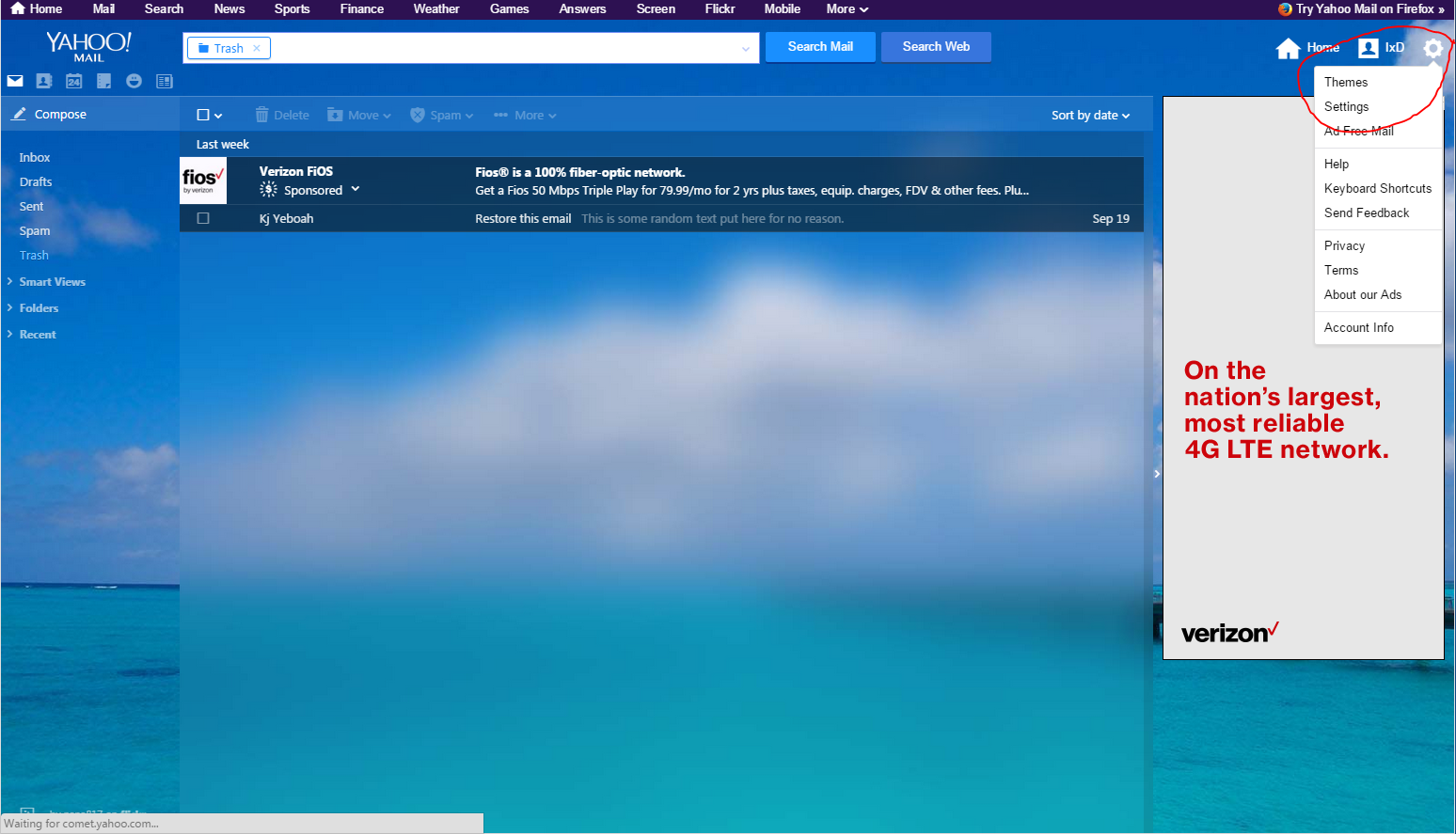






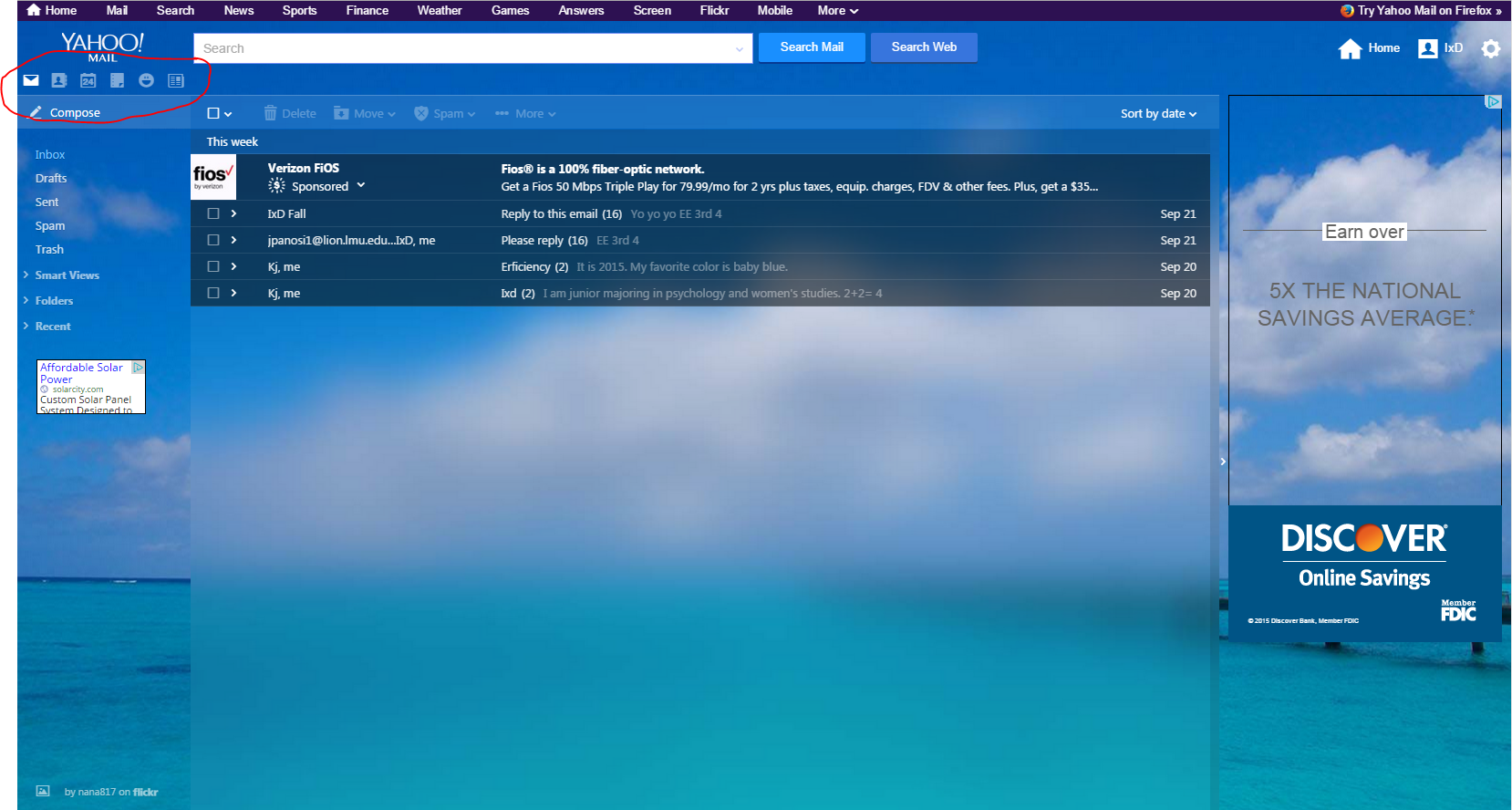
**Trends**

There were indeed some “repeat occurrences” between participants as they completed the tasks. For example, 6 of the 8 Errors associated with Yahoo! during both the Learnability and Efficiency phases were due to the settings gear icon on the top right of the screen.

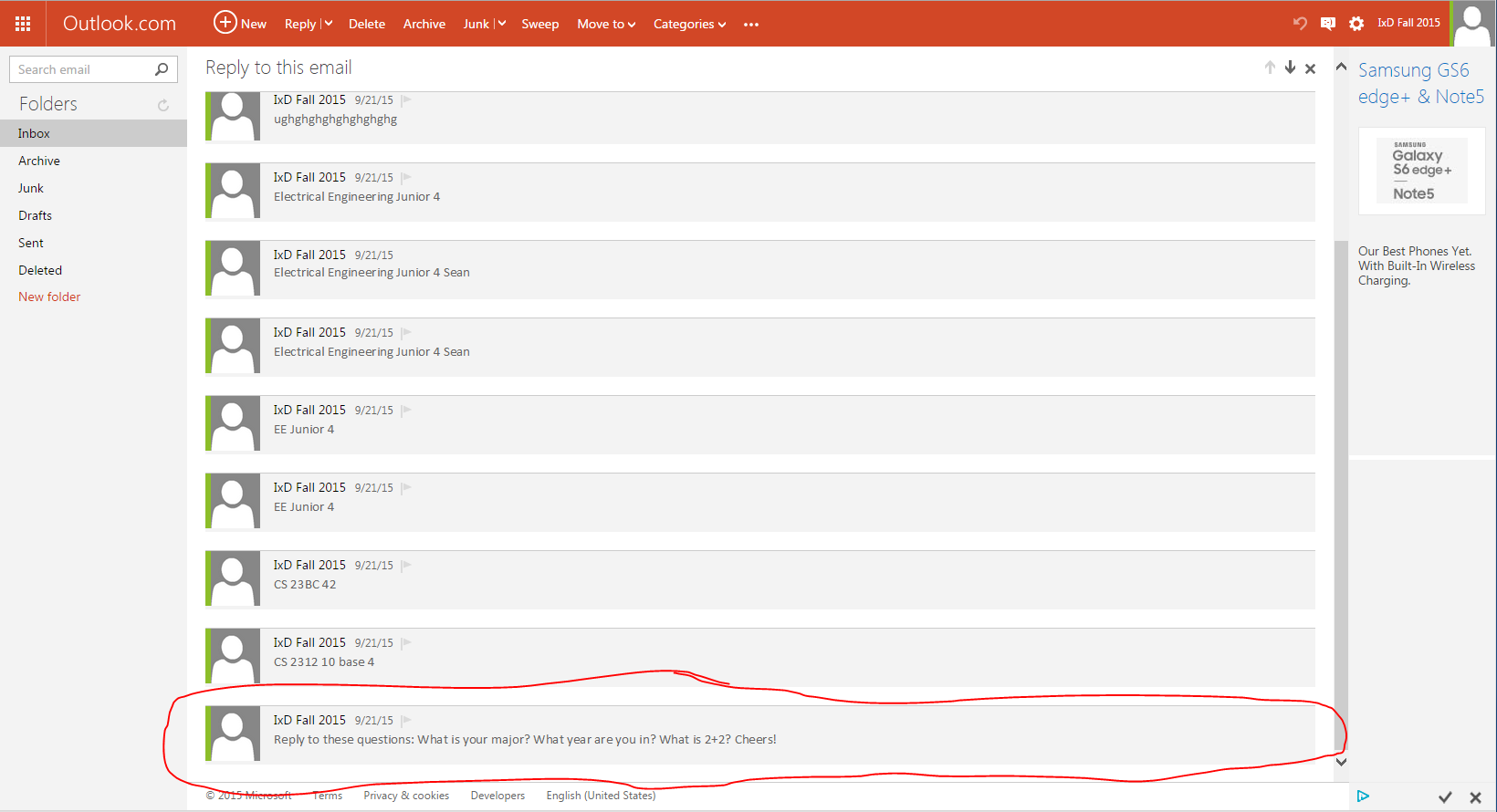


The gear icon is set so that the menu opens when the mouse hovers over it. Many users did not know this during the Learnability phase and had not adjusted to it even during the Efficiency phase. As a result, the participants would try to click on the gear icon, but the menu it reveals would appear and then immediately disappear. One participant had to click around 10 times just to access the menu. However, several participants also commented in the post-phase interview that the gear icon was crucial in helping them identify where the settings were because of its use in other interfaces such as Gmail.

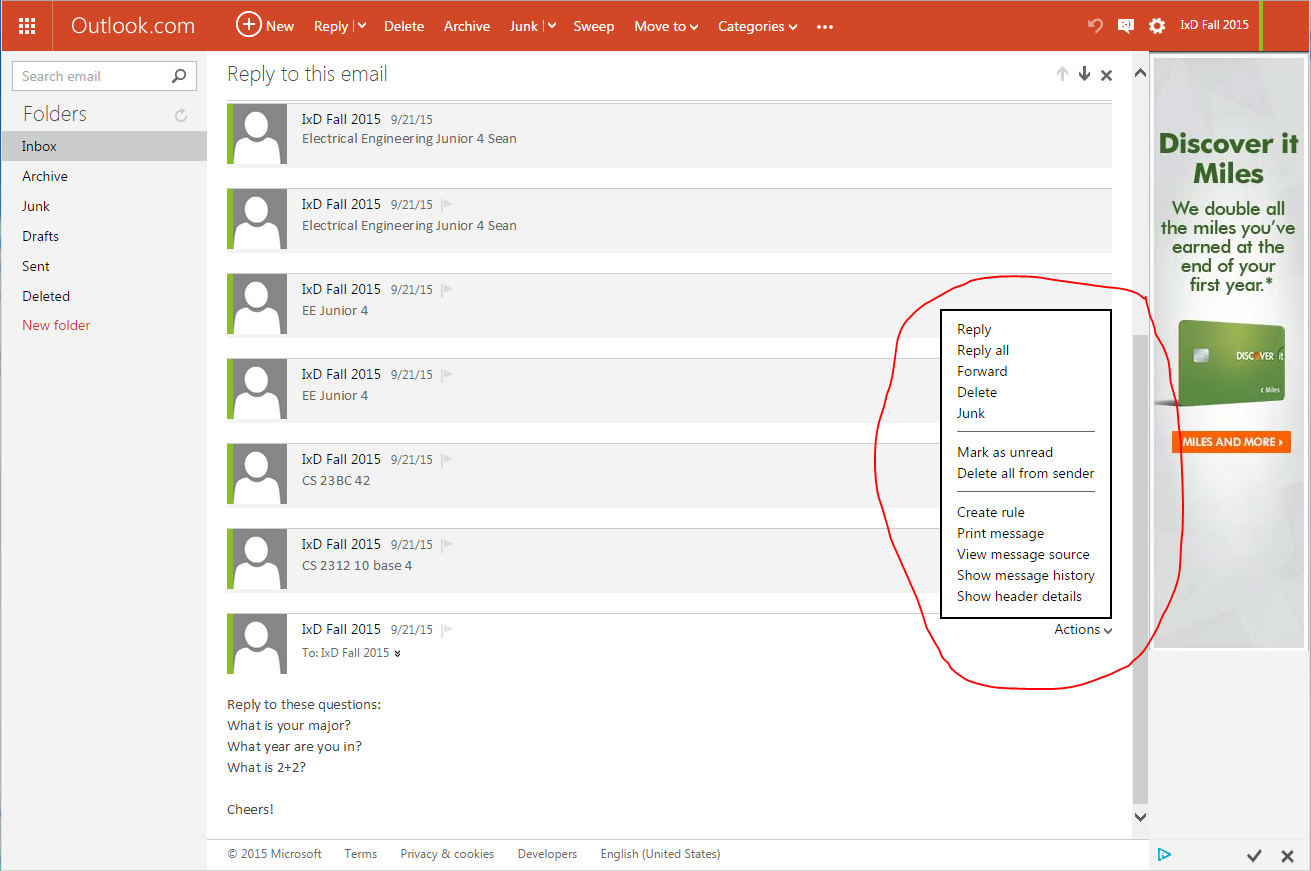
Another note is that several participants also started looking through the small transparent icons on the top left of the screen as soon as they were instructed to change the theme during the Learnability phase. Two participants even clicked on them just to see what would happen. After 4 or more seconds, the participants in this subset eventually moved to the right side of the screen to look for the theme settings.



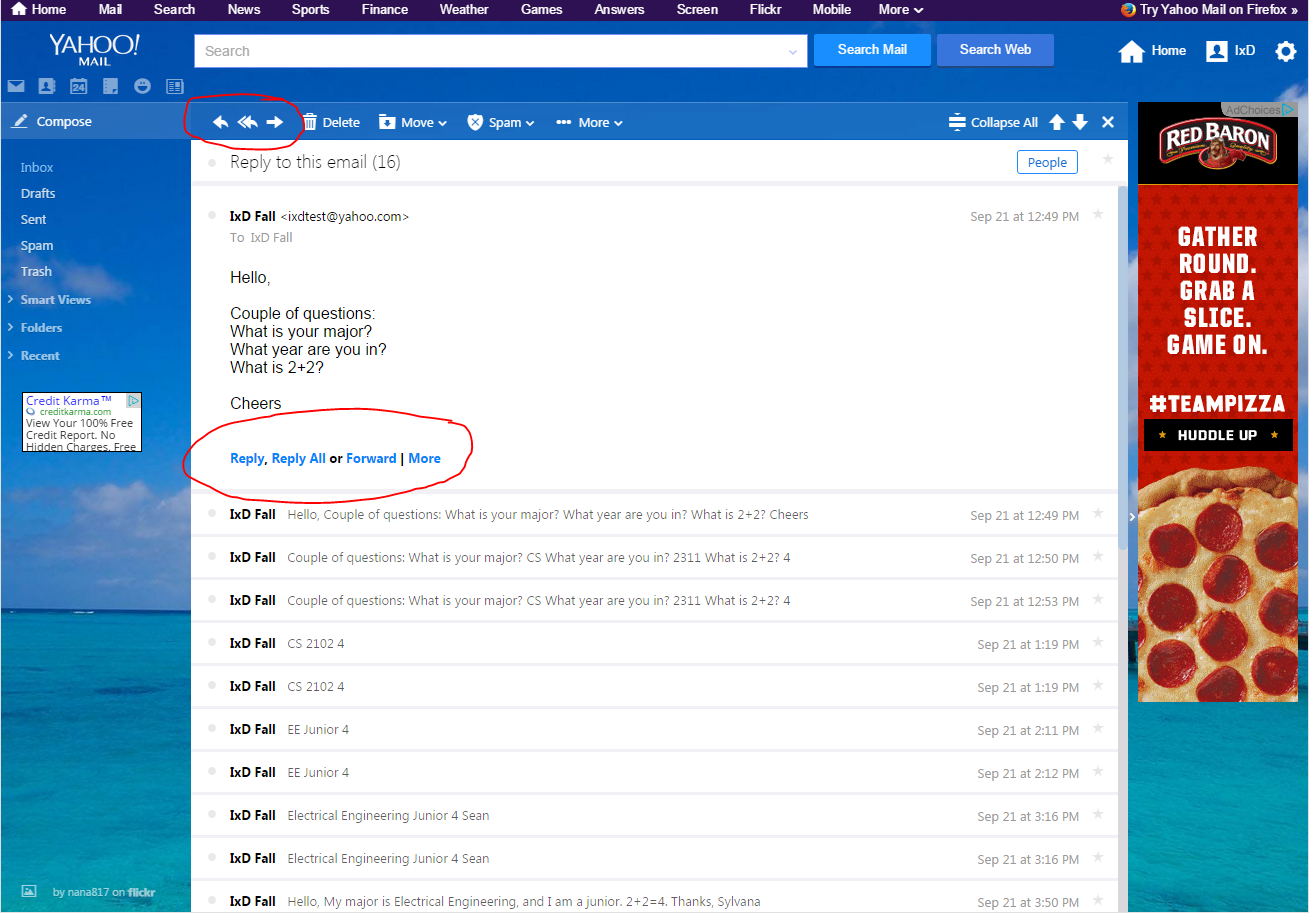
Outlook also had its caveats. For conversation threads in emails, Outlook utilizes an inverted thread feature, where the original message is displayed on the very bottom with consecutive replies being stacked above. Several participants stated this display was “weird” and threw them off because they expected the original message to be at the very top of the thread.



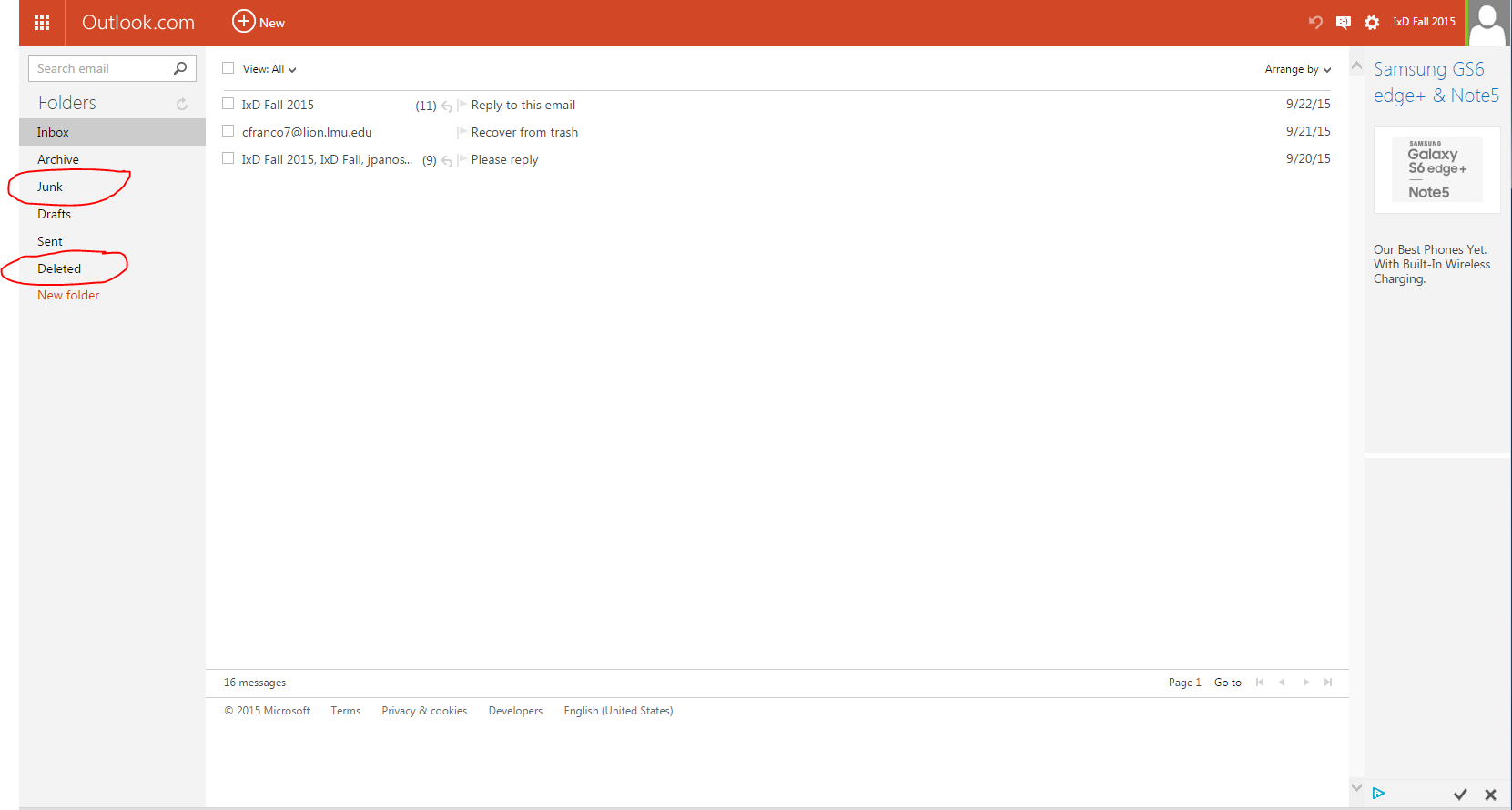
Another issue with this screen specifically that about half of the participants encountered was taking more than 3 seconds to find the “Reply” button. Notice that it is situated at the top of the screen, along with the other important email action options, opposite to the original message containing the content at the very bottom. The only other alternative for users is to click “Actions” on the right after opening an email.



Yahoo, on the other hand, had embedded Reply and Reply All options in the email itself. Still, several participants did not immediately locate these embedded options but instead used other means of sending the email after 3 or more seconds of searching.



Outlook also differs from other interfaces in its naming conventions. Instead of the standard “spam” and “trash” options on the left of the screen, Outlook has “Junk” and “Deleted” options in their place. This change threw off several participants and even led to an Error in one of the Efficiency surveys.



**Heuristic Evaluation/Conclusions**

We were able to find some form of interface guidelines for Yahoo!, but did not find documentation for the *design* of an interface for Outlook. The guidelines we did find for Outlook were about the development/software level:

Yahoo! - Yahoo Design Pattern Library

Outlook - Outlook VBA Reference (“for developing custom Outlook solutions.”)

According to Ben Schneidermann and Catherine Plaisant in their book *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, the number one rule in the “Golden Rules for interface design” is to *strive for consistency* (70). It states that terminology in menus, prompts, color, layout, fonts, and so on, should all be consistent with each other. Of course, this rule is regarding consistency throughout an individual application and not between several applications. Nevertheless, Outlook contained features that were not consistent with other widely used webmail interfaces. The inverted reply threads and the different naming conventions for spam and trash resulted in several errors by users and also led to them taking more time than what was probably necessary to accomplish a task. To reiterate, several users did mention that this layout was “weird,” and that they were used to different conventions from other applications. Furthermore, Schneidermann points out one of the top ten mistakes for presenting and accessing information on a web page is having “awkward or confusing navigation.” Unfortunately, Outlook’s separation of the Email actions from the actual content did lead to extended times and Errors. On the other hand, Yahoo! had also made a similar error in Schneidermann’s list of “putting information in unexpected places on the page” by embedding “Reply, Reply All, . . . “ options within the content of the email itself.

The author, being one of the participants of the study, can also attribute to expecting similar conventions across applications. One of the Errors made in the Yahoo Learnability phase was by the author when he was attempting to change the theme of the interface. He knew to click on the gear to access some form of settings, as is standard across webmail applications. However, upon opening the menu, the author overlooked the “themes” suboption and went straight to another suboption titled “settings.” This resulted in an extra 57 seconds to accomplish the task.

Perhaps, we can also enter into discussion about “Hidden Features,” as one of the participants put it. For example, only 2 participants in the study knew that they were able to drag an email from the Trash and drop it into the Inbox during the Learnability phase for *both* interfaces. One of those participants was a facilitator, and the second had accidentally been instructed by the facilitator to “drag the email into the inbox.” (The facilitator was supposed to use the more subtle language of “restore this email from the trash to the inbox,” which would have then required the user to explore different features.) The remaining 8 participants had all opened an email first and used the “Move” option at the top of the interface to move it to the Inbox. Hypothetically speaking, it would have probably taken much less time for users to accomplish the task of recovering an email from trash if they had known to drag the emails across the screen. In the Yahoo! Design Pattern Library, there is a page that describes best practices for Drag and Drop modules. Although there is a page for “Drop Invitation,” denoting the visual queue that a user can drop something into something else, there is no “Drag Invitation” literature. On their “Invitation” page, the literature explicitly states “The user needs to be informed of the existence of interactive features and invited to try them.” Despite this guideline, in both Yahoo! and Outlook, the mouse changes to a hand pointing with an index finger . . .



. . . inviting the user to click on the email, but not giving any visual queue that they are able to grab it. This may be the reason why many of the users first opened the email and then utilized the “Move” option at the top for both interfaces.

Several participants also discussed the layout of each of the interfaces. One of the participants, who was also a facilitator, stated the the Yahoo! interface seemed rather cluttered. They considered the window that viewed emails to be too small and attributed this smallness to the cluttering caused by the size of the search bar, adds, pop-ups, and a lot of small option icons. The author can also add to this from his experience of replying to an email during the Learnability phase. After clicking “Reply,” the text field appeared, but the original email with the questions disappeared off of the small box in the center of the screen that was displaying the thread. The author had to scroll up to read the questions and then scroll back down to type in the text field. One other participant also stated that there is “a lot going on” in the interface, with very small option icons for which a person is not able to deduce their function just by looking at them.

Regarding Outlook, a few participants commented on its simplicity. One of them stated that Outlook better utilized symbols to denote functions and actions on the interface, specifically for “Reply” and “Settings.” The participants were able to deduce their functions just by looking at them and quickly locate them on the page. Another user stated that compared to Yahoo!, the words and options in Outlook were “Big and Bright.” However, despite the advantages that simplicity offered in Outlook, there were some drawbacks. One participant had trouble finding the text field when replying to an email because there was a lack of formatting (too much white space) with no queues as to where the text field was.

Taking all of these aspects into consideration, it cannot accurately be stated that one interface is generally more effective than the other. Both excelled in certain areas and caused issues in others. What we can say, based on the quantitative and qualitative data, is that Outlook excelled in changing the theme due to the fact that one could only change the color and that the colors would immediately be available after clicking on the gear icon. Yahoo! and Outlook required similar amounts of time for replying to emails in both the Learnability and Efficiency phases, and Yahoo! also required less time to recover an email from trash during the Learnability phase (most likely because Outlook had different naming conventions). In terms of layout, one helpful aspect touched upon by a majority of users was the use of the gear icon in the top right corner for accessing theme options due to its standardization across many webmail applications.

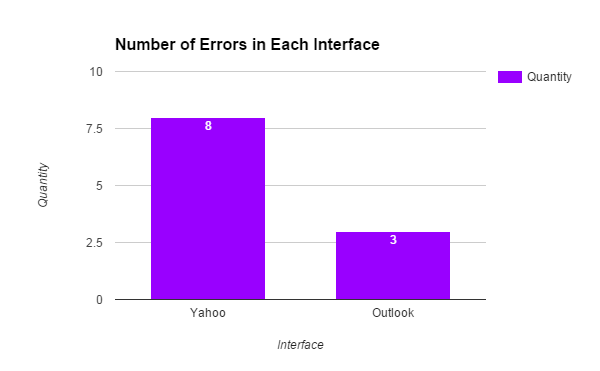
**Limits of the study**

All of the individuals who participated in this study were tech-savvy college students. Also, a majority of the participants were either Computer Science or Engineering majors. Given that this sample set is very small and extremely biased, it would be unwise to draw any far-reaching conclusions from this study about the effects of certain features of this interface.

Also, contrary to the expectations of the facilitators of this study, several of the participants stated that they do not usually restore emails from the trash and thus slightly invalidated our judgement of choosing tasks that are frequently performed by a broad set of users.

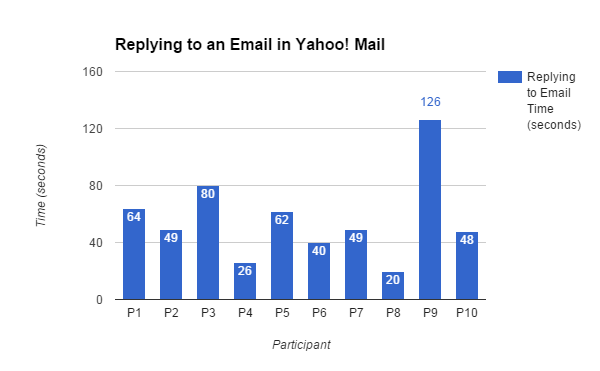
**Appendix 1**

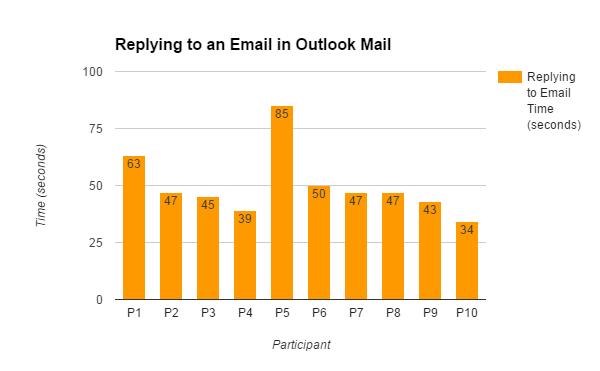
**Errors Metric**



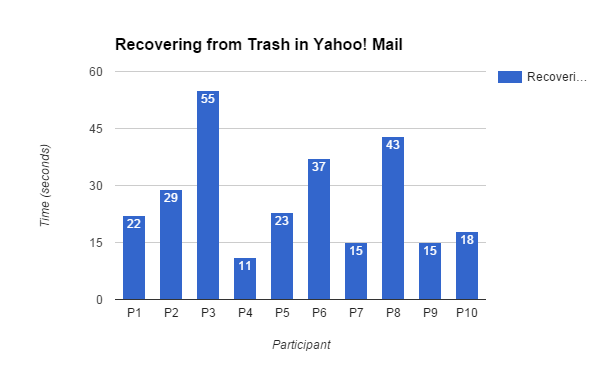
**Learnability Metric**

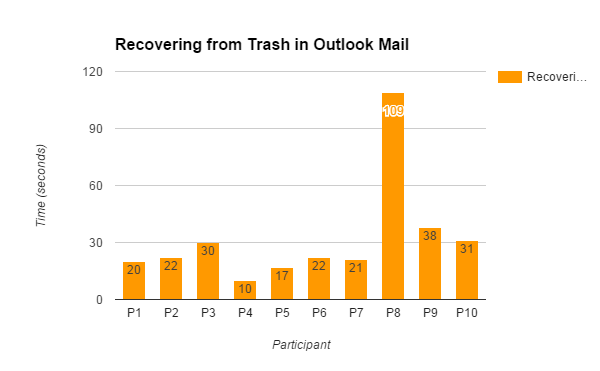
Replying to an Email



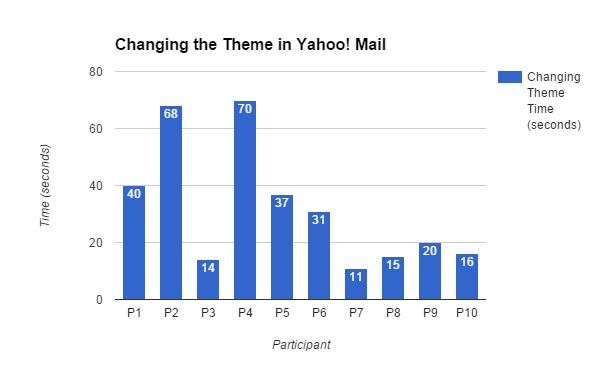


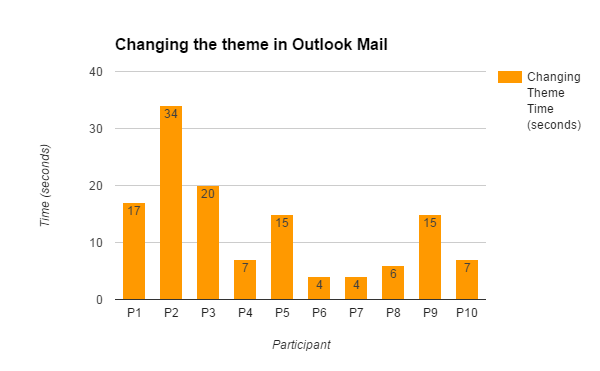
Recovering from Trash





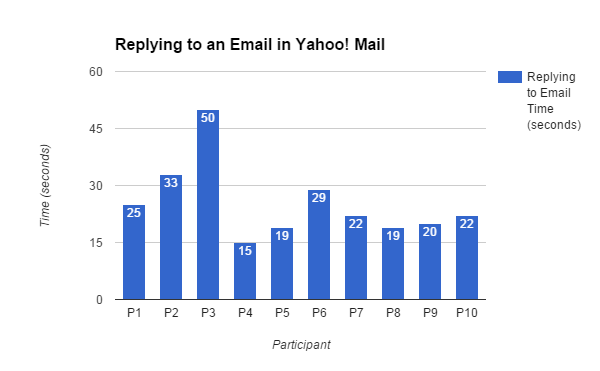
Changing Theme

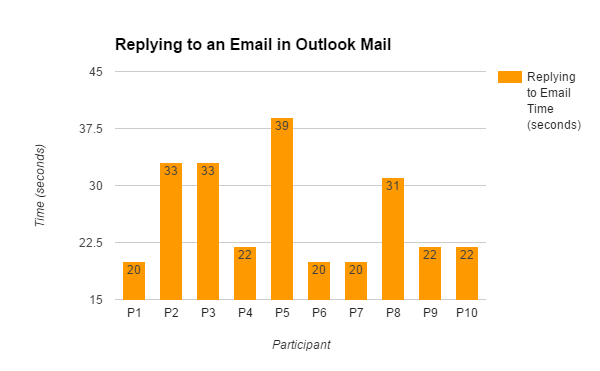




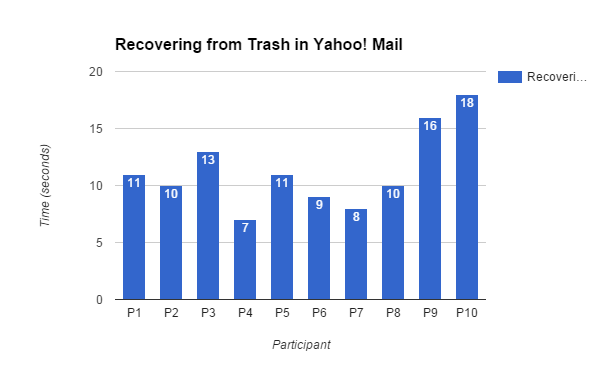
**Efficiency Metric**

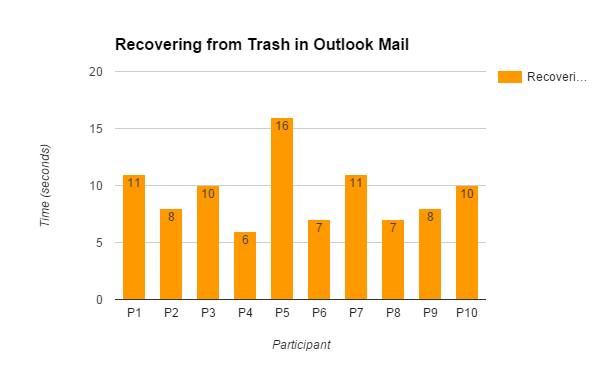
Replying to an Email





Recovering from Trash





Changing Theme

