

Redesigning 'Move to' - An Email Triage Function Design in Mobile Email Client

Qinyu Ding, Manouj Kumar Govindaraju, Feiyang Li, Jinqi Li, Aishwarya Navale

December 7th, 2018

Human-Computer Interaction CIM - 795

CLIENT

Abstract

Email has become a ubiquitous modern communication tool in present days. Though adopted by millions of people, email clients still lack a proper interface design for providing quality email management experiences. This study explores potential factors that influence the user experience in managing emails and proposes a solution for the current problems that have been observed in email filing.

The initial part of this study was the exploratory study in which three new designs to perform email filing in smartphones were proposed: drop-down design, side-scrolling design and pop-up design. In this study, the usability of three designs were compared in terms of effectiveness, efficiency and satisfaction. For this, a within-subject experiment was carried through with one independent variable (design) containing three levels (side-scrolling design, pop-up design and drop-down design). 18 participants performed assigned tasks in each of the designs and the observed results were used to determine the usability of each of the designs.

The exploratory study was followed by an experimental study in which the most usable design from the exploratory study was compared with the email filing feature of Microsoft Outlook. The independent variables were design (drop-down and Outlook) and number of folders (5, 10, 15). A 2x2 within-subject experiment was conducted. Similar to the experimental study, the results obtained from 18 new participants who performed the assigned tasks on both the designs were used to determine the usability of the designs.

In both experiments, the task completion time, recorded errors and satisfaction towards the designs were the dependent variables. The participants' subject evaluations towards each design was collected using 7-point Likert scale surveys and post-test questionnaires. In the exploratory study, data analysis results showed that, the drop-down design had better performance in terms of usability. Similarly, the result analysis of the experimental study revealed that the drop-down design had better usability than Outlook's design. The outcome of the analysis also showed that, scalability influenced this aspect. It was found that the usability of drop-down design was likely to go less than the Outlook's design when the number of folders increases. Taken together, this entire study revealed that the new drop-down design showed the potential to be a feasible alternative for building email management solution. This study will open avenues for further exploration and development of interface solutions for email management.

Keywords

Triage, folders, email filing

1. Introduction

Electronic mail (email) is a means or system for transmitting messages electronically between one person and another person or a group of people over the internet (Merriam-Webster, 2003). Over the past decade, there has been an exponential increase in the number of people who use email as the primary medium of telecommunication. Also, the volume of emails handled by each and every user has increased expeditiously. As the volume of emails in the inbox increases, the problems related to email management also increases. Such problems include cluttered inbox and the increased difficulty in information search (Whittaker & Sidner, 1996). Effective email management is the key to tackle these problems. But, the interface and the features that support email management have not evolved relatively well (Neustaedter, Brush & Smith, 2005).

One of the important email management methods is the practice of filing emails into folders. In this process, a user might move emails into folders for two purposes: organization of emails and information retrieval for later use (Neustaedter, Brush & Smith, 2005). In most email client applications, the interface has a 'move-to' button to trigger a list of target folders. Though the steps involved in performing this email management action are appropriate, the interface still fails to provide the effective means for performing email management (Neustaedter, Brush & Smith, 2005). For example, in most email applications, users find it difficult to identify the 'email triage' feature. Usually it takes at least three steps in the email interface to move an email into a folder. Therefore, the primary goal of this research was to propose an alternate interface design to solve the problems associated with filing emails into folders. The research was conducted in two phases.

CLIENT

The first phase hereinafter referred to as 'experiment 1' was the exploratory study, in which three new designs for filing emails were introduced namely side-scrolling design, drop-down design and pop-up window design. The results and observations of this study were used to determine the most usable design, evaluated in terms of effectiveness, efficiency, and satisfaction (*Ergonomics of human-system interaction*, 2018).

The second phase hereinafter referred to as 'experiment 2' was the experimental study, in which the best design identified from experiment 1 was compared with the Outlook design. This comparison involved two independent variables namely the design and scalability of the design. The two designs were the most usable design from experiment 1 and the Outlook design. The scalability of these designs were determined by testing each design with 5, 10 and 15 folders. The results of this experiment were again interpreted in terms of effectiveness, efficiency, and satisfaction. The implications of these results for effective email management was also discussed.

2. Background

There has been a tremendous growth in the volume of emails coming in one's inbox every day. Thus, different email classification methods have been introduced to help organizing emails. Using folders is one of the most popular ways for email triaging and refinding. As number of functions in email applications increases, menus are used to provide a well-organized interface. Studies have also provided evidence that structure and layout of an interface have a significant effect on both users' performance and experience.

2.1 The volume of emails in the inbox

There are several types of research pertaining to increased volume of emails in inbox and its associated problem. These problems may probably be due to the perpetuated email usage. The most commonly discussed problem is the pain of managing the emails when the volume of emails in the inbox is huge.

Emails should aid in communication and in task completion. When users spend time managing the inbox, it becomes a task by itself (Whittaker, Mathews, Cerruti, Badenes & Tang, 2011). When not managed, the huge volume of emails in the inbox may become a source of stress (Barley, Meyerson & Grodal, 2011). Dabbish and Kraut (2006) examined the relationship between email volume and user feelings and found in their experiment that higher the volume of emails in the inbox, higher the associated problems which include the increase in information search times and elevation of stress.

Several factors were found to be contributing to the huge volume of emails in the inbox. One of the important factors is that users do not have enough time to manage emails (Lantz, 1998). Another factor is that technology has been growing continually to offer higher inbox storage spaces (Whittaker & Sidner, 1996). In such cases, the users are not mandated to handle emails in their inbox (Carmel, Halawi, Lewin-Eytan, Maarek & Raviv, 2015). This ultimately leads to an increased volume of emails in the inbox.

2.2 Folders in email triaging

It is a common practice that an email user will seek for old emails either to retrieve information or to respond with an action. In such situations, the user is required to find the specific email within the client application. Users usually employ two strategies to retrieve an email: search and scroll, and the use of folders (Whittaker, Mathews, Cerruti, Badenes & Tang, 2011). The current technology allows users to search and scroll features to find an email. These features may seem to be effective than organizing emails into folders (Whittaker, Matthews, Cerruti, Badenes, & Tang, 2011). However, when the volume of emails in the inbox increases, the search feature generates a large list of emails. This is when filing emails into folders becomes a better solution (Whittaker, Matthews, Cerruti, Badenes, & Tang, 2011).

Creating folders, tagging emails to folders and re-finding emails from the folders are still preferred by a significant set of people (Whittaker, Matthews, Cerruti, Badenes & Tang, 2011). Venolia et al., (2001) in their survey revealed that more than 60% of the respondents used folders at least once a week to file emails (Venolia,

CLIENT

Dabbish, Cadiz & Gupta, 2011). This feature is preferred by users probably because it aids in faster information retrieval and at the same time promotes the organization of emails (Jones, Phuwanartnurak, Gill & Bruce, 2005).

However, as per some other researches, folders simply did not serve as the ultimate solution for information retrieval. In one such research, the researchers Grbovic et al. (2014) suggested several constraints for folders in order for it to be effective (Grbovic, Halawi, Karnin, Mareek, 2014). They suggested that the number of folders should be less than 20 to perform better than search and scroll (Grbovic, Halawi, Karnin, Mareek, 2014). One other suggestion is that the folder names should be easily understandable to the users (Grbovic, Halawi, Karnin & Mareek, 2014). When the folder names are not easily interpretable, the information within the folder becomes obscure, and ultimately the user tends to lose this information (Carenini, Ng & Zhou, 2005).

2.3 Menus in emails

Menus in any application work better when there is a semantic relation between the menu icon and the menu items. Here the menu icon acts as the visual cue for the user (Brumby & Zhuang 2015). In email applications, especially in smartphones, the three-dots icon (kabob or overflow menu) gives access to functions such as 'move-to', 'snooze', 'mark important', 'mute', 'print' and so on. These items are semantically unrelated. Also, the menu items lacks a visual cue. According to researchers, these two factors hugely influenced the efficiency and ease of use (Brumby & Zhuang 2015).

With regards to the orientation of the menu items, it has been proved by research that the search time for a user's eye to find a particular menu item is less in a vertical list than in a horizontal list (Backs, Walrath & Hancock, 1987). On the contrary, another research suggests that horizontal lists may be more effective when the most important list items are placed on the left (Williams, Hennig, 2015).

2.4 Hierarchy of menu

It is plausible for large commercial applications to have multiple menus and multiple levels within them (Kurtenbach, Fitzmaurice, Owen & Baudel, 1999). But email clients are supposed to be simple and the interface corresponding to primary actions should be straightforward, but they do not provide email management controls in simple manner. In most email client applications, the menus have eight items or less per level and are two levels deep which are acceptable in order to maintain selection speed and accuracy (Zhao & Balakrishnan, 2004). But, when primary actions are positioned under menus, accessibility gets compromised. Users would want fast and easier access to commonly used features (Kurtenbach, Fitzmaurice, Owen & Baudel, 1999). The placement of the move-to feature under the three-dot menu will make it difficult for the user to identify this feature. Especially, a novice user would need a simple and straightforward interface for a commonly used feature (Savage, Habinek & Barnhart, 1982). This also requires functional controls to be easy to find but also requires the design to meet the user's needs in as few steps as possible, reducing the number of redundant actions. Richard B. Segal and Jeffrey O. Kephart's (1999) research shows that predicting users' archive folders, reducing the number of clicks and the effort for locating informs efficient email archiving.

2.5 Work memory and positioning of menu

Another study shows that the interface which requires fewer clicks significantly decreases users' time-cost of locating items and reduces their frustration (Bao, Herlocker & Dietterich, 2006). The essence of message classification is information processing. Since human work memory has a limited capacity for cognitive processing, the performance in filing email into folders might be influenced by mental load. Studies have shown that lowering anxiety levels can help reduce the user's mental load and increase short-term memory (Gwizdka, 2013). In the case of a multi-level menu, the user may need to remember the contents of the mail for a short time and make a judgment as to which folder the mail is to be placed in. A lower frustration level is needed to improve the performance for such tasks. According to Norman's suggestion, designers should provide the user with the required operational

CLIENT

knowledge to reduce the user's short-term memory needs (Norman, 2013). While completing the task, what users need is recognition rather than recall.

2.6 Scalability of Menu

The breadth of the menu is one of the most critical task characteristics that cannot be ignored in the design of a hierarchical menu system (Jacko, Salvendy, & Koubek, 1994). It was examined formally in a study by Miller (1981) where the structure of a menu hierarchy was hypothesized to affect the speed and accuracy of target acquisition. Positive response times were found to increase linearly with the number of items displayed (Neisser & Beller, 1965). In addition, users' perceptions of the complexity of tasks would also be affected by the breadth of the menu (Campbell, 1988). Thus, scalability of the menu should be taken into consideration in study related to new menu design.

3. Methods

3.1 Experiment 1

3.1.1 Participants

Participants for this experiment were recruited from Miami, south florida. This sample size was 18, which was slightly less than 20 participants used for a previous menu study (Tsiodoulos, 2016) in order to assign the conditions to the 6 participant groups in a balanced manner. The participants consisted of 11 females and 7 males whose ages ranged from 20 to 40 ($M = 21.78$, $SD = 4.76$). All the participants were right handed and they were well aware of smart phones and email applications.

Table 1. Email clients usage information of participants

Operating phone system	%	Currently using email client	%
iOS	77.78%	Gmail	55.56%
Android	22.22%	Outlook	44.44%
		Default application on the phone	33.33%
		NetEase Email	11.11%
(Note: Some participants chose more than one option as response to this question.)			

CLIENT

Frequency of checking emails	%	Frequency of triaging emails	%
A few times a day	66.67%	Never	27.78%
Once a day	22.22%	Hardly ever	22.22%
A few times a week	5.56%	Yearly	5.56%
Once a week	5.56%	Monthly	22.22%
		Weekly	16.67%
		Daily	5.56%

3.1.2 Apparatus

Prototypes of the three designs were created with Figma 3.0 and were transferred using Figma Mirror 2.5 to an iPhone 7 running iOS 12. The inbuilt screen recorder of the iPhone 7 device was used to record all the operations of a participant. The test was conducted in a quiet classroom with the presence of all moderators. After the test, the recorded clips were transferred to the Adobe Premiere Pro CC 2018 software to calculate the task completion time. Anova within SPSS 24.0.0.0 was used to analyze the collected data.

A survey based on the 7-point Likert scale was created to evaluate each design (Tsiodoulos, 2016). The survey had two questions, respectively measuring how easy the design was to use, and how much the participant liked the design. A post-test survey was also created in order to gather feedback about the three designs for comparison.

3.1.3 Procedure

A participant was first given a brief introduction to the study, explaining the title, the purpose, and the study procedure. An Informed Consent Form was then presented to inform about the voluntariness of this experiment (MacKenzie, 2013). Then the participant filled out a demographic survey, followed by a three-minute warm-up session prior to the experiment to help him/her get familiar with all the design interfaces. In a previous study, Foster and Foxcroft (2011) provided a warm-up period of approximately five minutes for their participants. Since interfaces used in this experiment as well as the tasks performed were simpler than those in Foster and Foxcroft (2011), the length of time for the warm-up session was reduced to three minutes.

Following the warm-up period, the participant proceeded to perform the assigned tasks. The participants were presented with three designs with 5 trials in each design. In a trial, the participant opened the email, read the email body to find the priority designated for the email and filed the email into the corresponding folder. Similarly, the participant performed all the five trials of a design. Then the participant completed a survey containing two questions to evaluate the design. Likewise, the participant performed the tasks for the other designs and completed respective surveys. In the end, the participant reported their preference among all three designs in terms of the ease of operation, the efficiency, the aesthetic pleasure and their preference in general through a post-survey containing five questions.

3.1.4 Design

Experiment 1 followed a within-subject experiment design with one independent variable containing three levels. The independent variable was the ‘move-to-folder’ function design with three levels which were:

- A ‘move-to’ button which opened an accordion-panel containing a side-scrolling list of folders.

CLIENT

- A 'move-to' button which opened a drop-down list of folders.
- A 'move-to' button which triggered a pop-up window containing a list of folders.

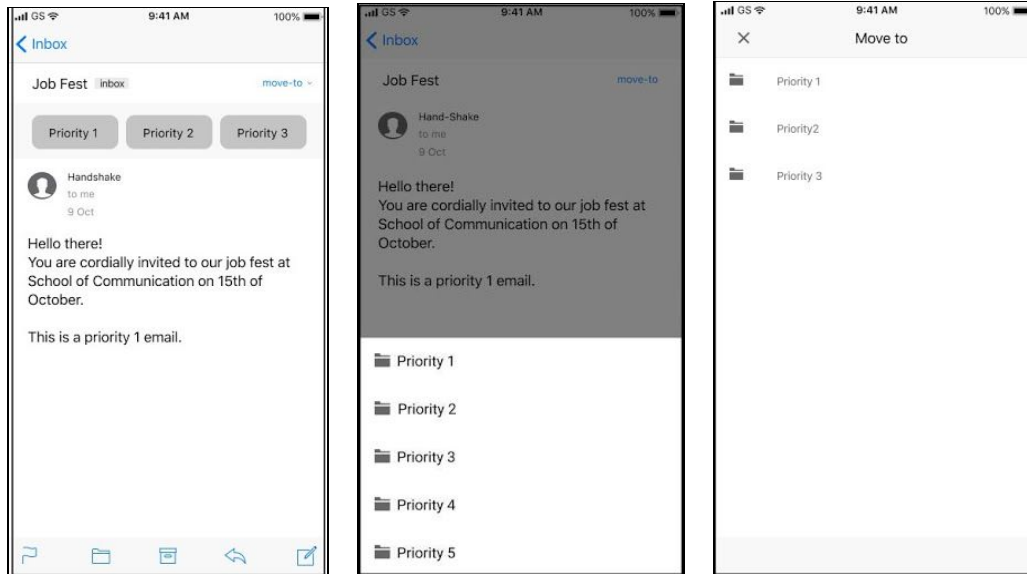


Figure 1: Side-scrolling design, drop-down design and pop-up design for 'move-to' button

The participant population was divided into six groups with three people in each group. There were 15 emails in total, distributed into three sets. The sets were then assigned to each of the group. To minimize the order effect, a balanced-Latin-Square order was used to determine the order of designs and the set of emails for every group (Williams, 1949). As the number of conditions was three, which is not even, a pair of 3×3 Latin squares were used to achieve balance. In this way, "the differences not represented in one square were represented twice in the other." (Williams, 1949). In addition, the order of emails within each set was randomized before presented to a participant.

Dependent variables of this experiment were effectiveness, efficiency, and satisfaction, which are three key measures used by ISO 9241-11 (*Ergonomics of human-system interaction*, 2018) to describe the definition of usability. The efficiency was measured by task completion time. The effectiveness was measured by the error rate. When the participant filed an email into a folder different to the assigned priority in the email body, it was counted as an error. The overall satisfaction for each design was measured by the results of all surveys.

The control variables were kept constant throughout the experiment. IOS style guidelines were referenced to design the typography of the interface. The font 'SF Pro' was used for all text content. For the primary text which includes titles, paragraph text, links, and button text, the font size was 17pt. For secondary text, the font size was 15pt. The font color for the email subject, sender name, and email content were in black (#000000). The preview text, timestamp, and secondary texts were in a lighter black color (#8C8C8C). The buttons and links were in blue (#007AFF). A 100% line height was used for the paragraphs. The sender name and subject lines had a line height of 22pt. Each list item had a height of 98pt.

3.2 Experiment 2

A further experiment was conducted to compare the winning design from experiment 1 with one of the standard design for "move-to" function - Outlook design. Enlightened by the one of the limitations identified in the experiment 1, scalability of menu design was taken into account in experiment 2.

CLIENT

3.2.1 Participants

Participants for this experiment were recruited from Miami, south Florida. This sample size was 18 of which 14 were female and 4 were male. The participants’ ages ranged from 18 to 50 ($M = 24.78$, $SD = 8.37$). Except one, all the other 17 participants were right handed. They all were well aware of smart phones and email applications. More than half of the participants never or hardly ever triaged their emails, but about 22.22% of participants triaged their emails frequently, even daily.

Table 2. Email clients usage information of participants

Operating phone system	%	Currently using email client	%
iOS	77.78%	Gmail	61.11%
Android	22.22%	Outlook	55.56%
		Other email clients	27.78%
		(Note: Some participants chose more than one option as response to this question.)	
Frequency of checking emails	%	Frequency of triaging emails	%
A few times a day	72.22%	Never	11.11%
Once a day	22.22%	Hardly ever	50%
Once a week	5.56%	Monthly	11.11%
		Weekly	5.56%
		Daily	22.22%

3.2.2 Apparatus

Apart from the surveys about each design and the post-test questionnaire, the same apparatus used in experiment 1 were used in experiment 2. The survey about each design had three questions, respectively measuring how easy the “move to” button was to be found, how easy the design was to use, and how much the participant liked the design. A post-test survey was also created in order to gather feedback about all the two designs to compare them.

3.2.3 Procedure

A participant was first given a brief introduction to the study, including the title, the purpose, and the study procedure. An “Informed Consent Form” was then presented to inform about the voluntariness of this experiment

CLIENT

(MacKenzie, 2013). After filling out a demographic survey, the participant underwent a three-minute warm-up session, keeping consistent with Experiment 1.

Following the warm-up session, the participant was presented with the first design, with first condition and the first trial. A trial required the participant to open the email, read the email body to figure out the priority of the email and finally file the email into its respective folder. Once the participant completed all the three trials of a condition, he/she proceeded to the next condition of the same design. Once all the three conditions of a design were completed, the participant completed a survey containing questions about the satisfaction with this design. Then, the participant moved to the second design where the same task completion procedure as the previous design was followed and then a similar survey related to this design was conducted. Finally, a post-test survey was conducted in which the participant reported his/her preference in terms of the ease of operation, the efficiency, and aesthetic pleasantness among the two designs.

3.2.4 Design

Experiment 2 involved within-subject design with two independent variables namely:

1. The design - 2 levels (drop-down design and Outlook design)
2. The number of folders - 3 levels (5,10 & 15 folders)

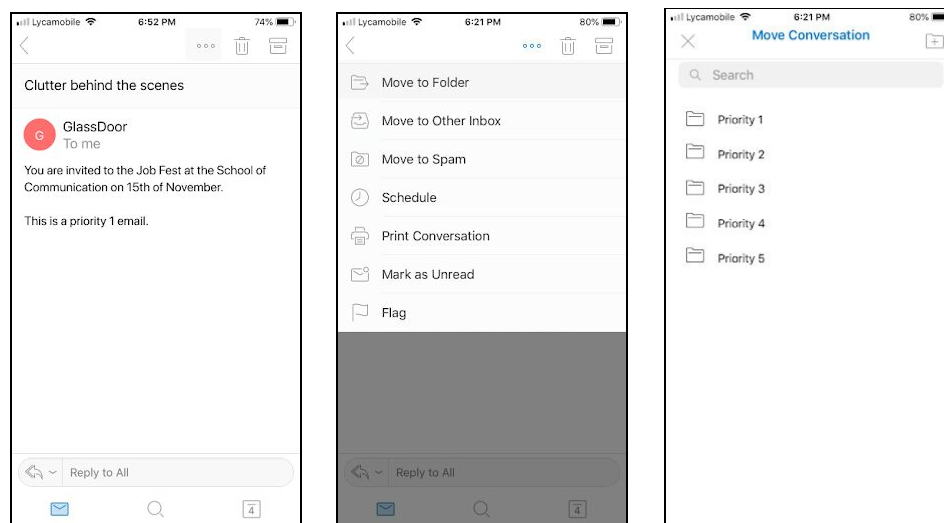


Figure 2: Outlook Design

The reasons for choosing 5, 10 and 15 as the number of folders were:

- 5 folders were used during the exploratory study, which was also the maximum number of folders displayed in one view of the drop-down design. Both the Outlook and drop-down design can handle 5 folders without the need for scrolling.
- The maximum number of folders displayed without scrolling in Outlook was 11. Thus, both the designs would require scrolling when the number of folders was 15.
- 10 is the median between 5 and 15. Here, only the drop-down design required scrolling. .

Participants were divided into two groups with nine people in each. There were 18 emails in total, distributed across six conditions. There were three emails per condition as opposed to five in the previous experiment in order to compensate the increased number of conditions in this experiment which will lead to very large number of trials. The three emails of each condition had priorities which were evenly distributed in the folder list to test the scrolling effect on the designs. Each participant first went through 5/10/15 folders on one interface

CLIENT

design, then through 5/10/15 folders on the other design. A 2×2 balanced Latin-square was used to determine the order of design to be presented to a participant for the purpose of minimizing the order effect (Mackenzie, 2013). The three emails of each condition had priorities which were evenly distributed in the folder list to test the scrolling effect on the designs. The dependent variables of experiment 2 was the same as the dependent variables of experiment 1. The control variables were also the same as experiment 1 and were kept constant throughout the experiment.

4. Results

During both the phases of experiments, all the participants completed the test and their responses were recorded. To compare the effect of the designs on efficiency and effectiveness, a one-way within-subjects repeated measure ANOVA was conducted in experiment 1 whereas a two-way within-subjects repeated measure ANOVA was conducted in experiment 2. The significance of the differences in satisfaction with three alternative designs was tested by the Friedman Test and Binomial Test in experiment 1. Also post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction to test where the significance lies. Since there were only two designs tested in experiment 2, Wilcoxon Test along with Binomial Test were used to determine the significant difference in satisfaction between the designs.

4.1 Experiment 1

4.1.1 Efficiency

The mean task completion time was 2.09s for drop-down design, 2.39s for side-scrolling design and 2.31s for pop-up design (Figure 3). The main effect of designs on task completion time was statistically significant ($F_{1,89} = 9.083$, $p < .05$). The results of post hoc analysis indicated that the mean score of drop-down design ($M=2.097$, $SD=0.072$) was statistically significant than the side-scrolling design ($M=2.396$, $SD=0.102$). However, the mean score of pop-up design ($M=2.316$, $SD=0.163$) did not statistically differ from either the drop-down or side-scrolling designs. It was thus identified that the time taken to complete tasks when using the drop-down design was significantly lower than side-scrolling and pop-up designs.

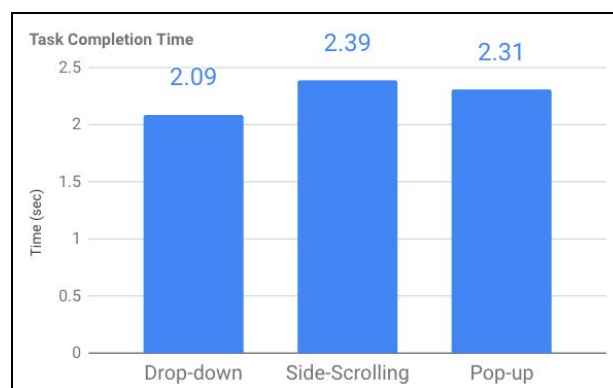


Figure 3: Mean task completion time

CLIENT

4.1.2 Effectiveness

Overall, there were four errors that were recorded and were all from the pop-up design. The error rate in the pop-up design was about 4.4% (Figure 4). There was no error in the other two designs.

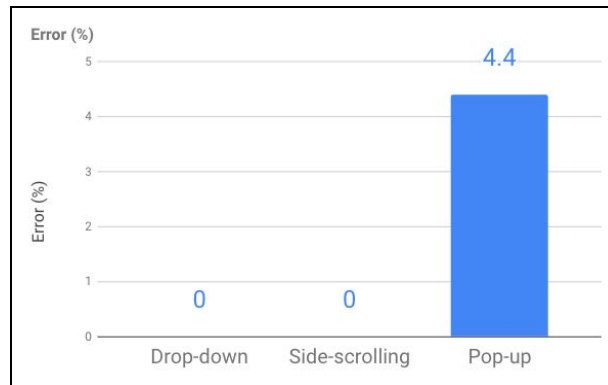


Figure 4: Error rate for each design

4.1.3 Satisfaction

A complete analysis of all the participants' surveys of experiment 1 was made to identify the most satisfying alternative design. Questions from all the surveys were of four categories: overall preference, preference in terms of efficiency, preference in terms of ease of operation and preference in terms of aesthetic pleasure.

Overall, the drop-down design was the most preferred one among all three alternative designs. The mean rating for the 7-point Likert scale question "I like this design very much" was 5.72 for drop-down design, 4.94 for side-scrolling design and 3.94 for pop-up design. The drop-down design was rated 45.18% higher than pop-up design and 15.79% higher than side-scrolling design. There was a statistically significant difference in preference depending on which design type was used, $\chi^2(2) = 10.554$, $p = 0.005$. Post hoc analysis with Wilcoxon signed-rank tests resulted in a significance level set at $p < 0.017$. There were no significant differences between the drop-down and side-scrolling designs ($Z = -1.282$, $p = 0.200$) or between the pop-up and side-scrolling designs ($Z = -1.659$, $p = 0.097$). However, there was a statistically significant difference in perceived usefulness between the pop-up and drop-down designs ($Z = -3.134$, $p = 0.002$), indicating that drop-down was perceived to be better than pop-up design.

Answers to other questions related to preference in the post-test survey also demonstrated that drop-down design was the winning design. For the question "Which of the three designs do you like best?", it was determined that drop-down design was the most preferred (11 of 18, $p < .05$), using binomial test with confidence intervals. No participant preferred pop-up design. The 95% confidence interval was 39% to 80%. The lower boundary of the confidence interval indicated preference was unlikely to fall below 39%, which was still above the chance threshold of 33.3%.

According to answers to the second close-ended question, the drop-down design was chosen as the option that participants would change to if given the opportunity to do so. This difference in response towards drop-down design was statistically significant (12 of 18, $p < .05$) compared to the side-scrolling design.

Based on the responses to the question concerning to the efficient way of putting emails into folders, it was determined that drop-down was perceived to be the most efficient way to put the email into a folder (11 of 18, $p < .05$). No participant thought the pop-up design was the most efficient.

According to the responses to the question related to the ease of operation, the mean rating was 6.50 for drop-down design, 5.33 for side-scrolling design and 5.22 for pop-up design. The drop-down design was rated 42.5% higher than pop-up design and 14% higher than side-scrolling design in this aspect. There was a statistically

CLIENT

significant difference in perceived ease of operation depending on which design type was used, $\chi^2(2) = 11.902$, $p = 0.003$. Post hoc analysis with Wilcoxon signed-rank tests resulted in a significance level set at $p < 0.017$. There was no significant difference between the pop-up and side-scrolling designs ($Z = -0.731$, $p = 0.465$). However, there was a statistically significant difference between the pop-up and drop-down designs ($Z = -2.850$, $p = 0.004$) or between the side-scrolling and drop-down designs ($Z = -2.007$, $p = 0.045$) indicating that drop-down design was perceived to be easier to operate than side-scrolling and pop-up.

As revealed by the responses to the second close-ended question, the drop-down design was statistically significantly perceived to be the easiest way to put an email into a folder (13 of 18, $p < .05$) compared to the other two designs. No participant found that pop-up design was easy to operate.

From the responses for the question pertaining to the aesthetically pleasing design, it was determined by using binomial test with confidence intervals that the side-scrolling design was perceived to be the design (12 of 18, $p < .05$). No participants thought the pop-up design was aesthetically pleasing. The 95% confidence interval was 44% to 84%. The lower boundary of the confidence interval indicates preference is unlikely to fall below 44%, which is still above the chance threshold of 33.3%. Side-scrolling was statistically significantly perceived to be the most aesthetically pleasing.

4.2 Experiment 2

4.2.1 Efficiency

A two-way ANOVA was conducted that examined the effect of design and number of folders on task completion time. There was a significant main effect of design on time to complete tasks ($F(1, 53) = 9.793$, $p = 0.002$). There was also a significant main effect of number of folders on task completion time ($F(1.755, 93.001) = 14.021$, $p = 0.000$). Pairwise comparison with 95% confidence intervals reveals that there was a statistically significant difference between 5 folders and 10 folders ($p=0.000$). Similarly, there was a statistically significant difference between 5 number of folders and 15 folders ($p=0.000$). But there was no statistically significant difference between 10 number of folders and 15 folders ($p=1.000$).

It was necessary to identify simple main effects because there was a statistically significant interaction between the effects of design and number of folders on task completion time, $F(1.687, 89.386) = 6.077$, $p = .005$. The results of the interaction effect analysis are below.

- The within-subjects contrasts showed that there was a statistically significant interaction between 5 folders and 15 folders on the designs ($F(1,53) = 14.081$, $p=0.000$).
- Similarly, there was a statistically significant interaction between 10 folders and 15 folders on the designs $F(1,53) = 4.974$, $p=0.03$).
- However, there was no statistically significant interaction effect between the 5 folders and 10 pairs on the two designs.

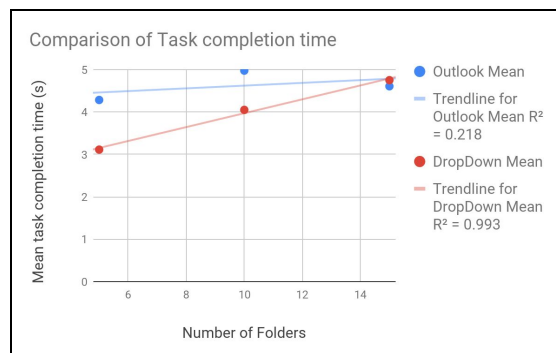


Figure 5: Task completion time (Experiment 2)

CLIENT

From the data, it was identified that the average task completion time of the Outlook design was 4.62 seconds which was higher than that of the drop-down design for which the mean task completion time was 3.97 seconds. It was also identified from the means of task completion time across designs that the time taken to complete tasks for drop-down design with 5 number of folders (mean = 3.11, SD = 1.39) was lower than that of Outlook design with 5 number of folders (mean = 4.28, SD = 1.56). However, when the number of folders were 15, the task completion time for Outlook (mean = 4.60, SD = 1.74) was less than drop-down design (mean = 4.75, SD = 2.34) (figure 5). Also, the trendlines indicating the mean task completion time of the designs intersect at 15 number of folders.

4.2.2 Effectiveness

There were in total 10 errors recorded across all participants (Figure 6). Of these, 6 errors were recorded when the participants were using the Outlook with 10 folders (error rate = 11.53%).

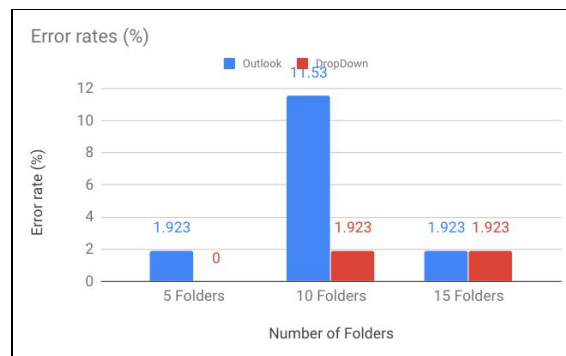


Figure 6 : Error rates (Experiment 2)

In Outlook design with 5 folders, drop-down with 10 folders and both the designs with 15 folders, one error was observed in each of these conditions. No error was recorded when the drop-down design with 5 folders was used. It was thus identified that the overall effectiveness of the drop-down design was higher than the Outlook design.

4.2.3 Satisfaction

All participants' responses for satisfaction towards Outlook's 'move-to' design and the drop-down design were recorded through surveys and were analyzed to determine which design was more preferred. Questions from all the surveys were from four categories: overall preference, preference in terms of discoverability, preference in terms of ease of operation and preference in terms of aesthetic pleasure.

Overall, the drop-down design was the most preferred one among the two designs. The mean rating for the 7-point Likert scale question "I like this design very much" was 3.93 for Outlook 'move-to' function design and 6.00 for drop-down design. The drop-down design was rated 52.67% higher than Outlook 'move-to' function design with respect to preference. A Wilcoxon signed-rank test showed that there was a significant preference for the drop-down move-to design ($Z = -3.523$, $p < 0.001$) over the Outlook design.

Answers to the other question related to preference in the post-test survey also demonstrated that drop-down design was the winning design. For the question "Which of the two designs do you like best?", it was determined that drop-down was statistically significantly preferred (17 of 18, $p < .05$), using binomial test with confidence intervals.

CLIENT

The drop-down design was considered to be more excellent in discoverability. On average, Outlook design was rated 3.39 for the 7-point Likert scale question "This 'move-to' button is very easy to find.", while the drop-down design was rated 6.00. The drop-down design was rated 76.99% higher than Outlook 'move-to' function design with respect to discoverability. A Wilcoxon signed-rank test showed that the difference was statistically significant ($Z = -3.644$, $p < 0.001$).

The mean rating for the question about perceived ease of operation was 4.50 for Outlook 'move-to' function design and 6.06 for drop-down design. The drop-down design was rated 34.67% higher than Outlook 'move-to' function design with respect to ease of operation. The difference was statistically significant according to the result of the Wilcoxon signed-rank test ($Z = -3.457$, $p < 0.001$) indicating that drop-down design was perceived to be easier to operate than Outlook design.

As stated by the answers to the second close-ended question, the drop-down design was also statistically significantly perceived to be the easiest way to put the email into a folder, using binomial test with confidence intervals (18 of 18, $p < .05$). With all participants choosing the drop-down design for this question, the drop-down design was identified as the obvious winning one in terms of ease of use.

There was no statistically significant difference between the two designs in terms of aesthetic pleasure. 13 out of 18 participants found that the drop-down design was more aesthetically pleasant compared to Outlook 'move-to' function design. However, the binomial test with confidence intervals revealed that the difference was not statistically significant (13 of 18, ns).

5. General Discussion

Results of all tests in studies were delved into in terms of the potential cause and effect of the obtained efficiency, effectiveness and satisfaction of each design in this section. Limitations of two experiments were also discussed. In the end, insights found from the results were proposed for future design.

5.1 Efficiency

Out of the three proposed designs, the drop-down design was found to be the most superior design in terms of task completion time. Participants took significantly less time to complete tasks while using the drop-down design than the pop-up design and the side-scrolling design in the exploratory study. This may be attributed to the fact that the drop-down design shared the design advantages of each of the other two designs but not the disadvantages. The drop-down design and the side-scrolling design displayed folder items on the same page with email bodies. This reduced the requirement of work memory of the participants to remember the email content. The pop-up design displayed folders on a new window and this increased the need for work memory. The second shared advantage was that the drop-down design and the pop-up design both have vertical list items, which has been proved to be easy to interpret and more efficient than horizontal list as in side-scrolling design (Bucks, Walrath & Hancock, 1987).

In the experimental study, the efficiency of the two designs were affected by factors like the position of the 'move-to' button, design type and the scalability. In Outlook design, the move-to button was hidden in a three-dot menu, requiring efforts to find the function in a list with other items. In the drop-down design, on the other hand, the availability of move-to button in plain sight reduced the information processing time of participants which ultimately contributed to the smaller task completion time. Also, the number of clicks that required to find the 'move-to' button was reduced in drop-down design, reducing the operation time while participants finishing the task. Secondly, the Outlook design followed similar design layout as the pop-up design. Thus it shared the same disadvantage as the pop-up design related to work memory. Finally, analysis of designs based on scalability revealed that the task completion time of drop-down design was lower for smaller number of folders. This is because as the number of folders increased, the amount of scrolling required in this design also increased. This was not the case in Outlook design, where the entire screen was used as a window to display the folder list - leading to smaller scrolling needs.

Though there was a non-linear increase in the task completion time for the Outlook design, the general trend reveals that, for an increase in the number of folders, the increase in the time cost of the drop-down design was

CLIENT

more than the increase in the time cost of the Outlook design. This suggests that, eventually, the higher efficiency of drop-down design holds good only for fewer number of folders and that the efficiency will be less than the Outlook if there are more than 15 folders. One possible explanation for this trend is, Outlook occupied the whole page to display folders while the drop-down design only has less than half screen to show them. Hence, in this case, the time needed for scrolling in the drop-down design compensated for the time used for more clicks in Outlook design, making task completion time when using Outlook even less than using the drop-down design.

5.2 Effectiveness

In experiment 1, all the errors were recorded in the pop-up design. This might be due to the fact that this design showed the folder list in a separate window which impacted the work memory aspect. As users opened the pop-up window which covered the entire screen, they were unable to refer the email content anymore. Four participants were identified to have forgotten the assigned priority of the email when using the pop-up design during the test. The other two designs did not have this issue as they displayed the folder list on the same window as the 'move-to' button.

In experiment 2, 80% of errors happened in the Outlook design. This again might be due to the same reason as why the pop-up design had more errors in experiment 1. From the results of both the experiments, it was identified that the drop-down design had the least error rates which indicated that this design is the most effective.

5.3 Satisfaction

Among the three new designs of 'move-to' function, side-scrolling was perceived to be the most aesthetically pleasing design. According to the feedback from participants, this might be due to less space being taken, closer position to the "move-to" button and accessibility to the email content while triaging emails when using the side-scrolling design. It took the least amount of space on the screen which made it the most compact design among all the 3 designs. The folder list in the side-scrolling design opened closely to the move-to button. Therefore, side-scrolling design required the least eye gazing movement for the users. This reduced their efforts of locating the target priority button as quickly as possible, thus lowering the level of frustration. Also, the users could view the email content simultaneously even when the scrollbar was open. This again prevented the issues pertaining to work memory.

The drop-down design was revealed to be the easiest design to operate, easiest design to find, and also the one that users most preferred compared to both the other two designs and the Outlook design. In addition to the reason that the content remaining visible after clicking the move-to button reduced users' memory load, the vertical layout of drop-down list required fewer scrollings compared to the horizontal layout. Aside from fewer clicks being required, drop-down design also provided a direct access and a clear signifier to the "move-to" function, thus making it more excellent in discoverability as well. Hence, the drop-down design was the most preferred design in the study.

5.4 Limitation

In both experiments, the small sample size was an obvious limitation. There were only 18 participants, most of them were from a same college, which led to narrow kind of email management population. Also, most of the participants were iOS users, and only a fraction of participants were Android users. Since all the tests were conducted on the iOS system, the familiarity level with the system might have had a slight influence on participants' performance. For future research, a larger sample size and a more diverse sample is recommended.

The fewer number of trials turned out to be another limitation. In the first experiment, participants only experienced five trials for each condition. In second experiment, the number of trials for each condition was three.

CLIENT

In reality, there could be much more emails in user's inbox, complicating the situation that user might face with, and adding other effects such as fatigue.

Moreover, In the second experiment, the number of folders only had three levels, 5, 10 and 15. This leads to a limitation when studying the influence of the number of folders on independent variables. The result showed that Outlook had a better performance with 15 folders than with 10 folders, which was opposite to the result of drop-down design. However, since there are only three levels, it was hard to tell how the increasing number of folders will affect the Outlook interface.

Also, when using surveys to collect participants' subjective opinions, the questionnaire didn't ask the behavioral intention, such as "If you like the drop-down design better, do you think this design can encourage you to classify emails?" In further research, such questions might address more practical significance for the design.

Two limitations of the first experiment were fixed in the second experiment. In experiment 1, there was a minor delay in the prototype of side-scrolling design, influenced the performance of that particular design. Such delay was avoided in experiment 2. Also, experiment 1 failed to consider the menu scalability. In experiment 2, scalability of the menu was the focus point of the design.

5.5 Insights

One thing revealed from the experiments is that, when information is needed for users to make decisions or achieve task, it is better to be visible for recognition rather than recalling (Norman, 2013). In other words, design should offer users access to accomplish their goals with the minimum memory load (Norman, 2013).

The other thing is that when a function is supposedly designed for users to use, it is advisable to put it somewhere that people have directly access to it. Another point worth noting is that the breadth of menu can influence its performance. Not only information processing time resulting from the number of items would increase, it would also influence users' perceptions of complexity of tasks (Campbell, 1988). In addition the drop-down list does not have good scalability when the number of items within it is large. Hence it may not be the best option for menus with comparatively more items.

6. Conclusion

Taken together, it is evident from the results that the drop-down design have a significant advantage over the the other two new designs as well as the Outlook design in terms of efficiency, effectiveness and user preference. However, it is important to note that with regard to scalability of menus, the drop-down design and Outlook design performed differently. When the number of folders in the folder list were more, the efficiency of drop-down design, in terms of mean task completion time, tended to degrade relatively more than that of Outlook design. Nevertheless, the Outlook design was found to have an inconsistent mean task completion time across different number of folders. These results suggest that for less number of folders, drop-down design can positively serve as a design alternative to the Outlook design. By implementing this drop-down design while building interfaces for email clients, it can be expected that users will deal with the problem of email management in a more efficient and effective way.

References

- Backs, R. W., Walrath, L. C., & Hancock, G. A. (1987, September). Comparison of horizontal and vertical menu formats. In *Proceedings of the Human Factors Society Annual Meeting* (Vol. 31, No. 7, pp. 715-717). Sage CA: Los Angeles, CA: SAGE Publications.
- Bao, X., Herlocker, J., & Dietterich, T. (2006). Fewer clicks and less frustration: Reducing the cost of reaching the right folder. *Proceedings of the 11th International Conference on Intelligent User Interfaces*, 2006, 178-185.
- Barley, S. R., Meyerson, D. E., & Grodal, S. (2011). E-mail as a source and symbol of stress. *Organization Science*, 22(4), 887-906.
- Brumby, D., Zhuang, S., Begole, B, Kim, J, Inkpen, K, & Woo, W. (2015). Visual grouping in menu interfaces. In: Begole, Begole, B and Kim, J and Inkpen, K and Woo, W, (eds.) CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. (pp. pp. 4203-4206). Association for Computing Machinery (ACM): New York, NY, USA. (2015).
- Corporation. (2013). *Human-computer interaction : An empirical research perspective*. Amsterdam ; Burlington: Elsevier Science.
- Campbell, D. J. (1988). Task Complexity: A Review and Analysis. *Academy of Management Review*, 13(1), 40-52.
- Carenini, G., Ng, R. T., & Zhou, X. (2005, August). Scalable discovery of hidden emails from large folders. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining* (pp. 544-549). ACM.
- Carmel, D., Halawi, G., Lewin-Eytan, L., Maarek, Y., & Raviv, A. (2015, October). Rank by time or by relevance?: Revisiting email search. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management* (pp. 283-292). ACM.
- Dabbish, L. A., & Kraut, R. E. (2006, November). Email overload at work: an analysis of factors associated with email strain. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work* (pp. 431-440). ACM.
- Foster, G., & Foxcroft, T. (2011). Barrel menu: A new mobile phone menu for feature rich devices. *Proceedings of the South African Institute of Computer Scientists and Information Technologists Conference on Knowledge, Innovation and Leadership in a Diverse, Multidisciplinary Environment*, 97-105.
- Grbovic, M., Halawi, G., Karnin, Z., & Maarek, Y. (2014, November). How many folders do you really need?: Classifying email into a handful of categories. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management* (pp. 869-878). ACM.
- Gwizdka, J. (2013, June). Effects of working memory capacity on users' search effort. In *Proceedings of the International Conference on Multimedia, Interaction, Design and Innovation*(p. 11). ACM.
- Jacko, J. A., Salvendy, G., & Koubek, R. J. (1994). An experimental study for menu design guidelines for menu selection. *Proceedings of the 12th Triennial Congress of the International Ergonomics Association*, 4, 351-353.
- Jones, W., Phuwannartnurak, A., Gill, R., & Bruce, H. (2005). Don't take my folders away: Organizing personal information to get things done. *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, 1505-1508.
- Kurtenbach, G., Fitzmaurice, G. W., Owen, R. N., & Baudel, T. (1999, May). The Hotbox: efficient access to a large number of menu-items. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 231-237). ACM.
- Lantz, A. (1998). Heavy users of electronic mail. *International Journal of Human-Computer Interaction*, 10(4), 361-379.
- MacKenzie, I., & Ebooks Corporation. (2013). *Human-computer interaction: An empirical research perspective*. Amsterdam; Burlington: Elsevier Science.
- Miller, D. P. (1981). The Depth/Breadth Tradeoff in Hierarchical Computer Menus. *Proceedings of the Human Factors Society Annual Meeting*, 25(1), 296-300.
- Neisser, U., & Beller, H. K. (1965). Searching Through Word Lists. *British Journal of Psychology*, 56(4), 349-358.
- Neustaedter, C., Brush, A. J., & Smith, M. A. (2005, April). Beyond from and received: Exploring the dynamics of email triage. In *CHI'05 extended abstracts on Human factors in computing systems* (pp. 1977-1980). ACM.

CLIENT

- Norman, D. (2013). *The design of everyday things (Revised and expanded ed.)*. New York: Basic Books a member of the Perseus Books Group.
- Savage, R. E., Habinek, J. K., & Barnhart, T. W. (1982, March). The design, simulation, and evaluation of a menu driven user interface. In *Proceedings of the 1982 Conference on Human Factors in Computing Systems* (pp. 36-40). ACM.
- Segal, R. B., & Kephart, J. O. (1999, April). MailCat: An intelligent assistant for organizing e-mail. In *Proceedings of the third annual conference on Autonomous Agents* (pp. 276-282). ACM.
- Whittaker, S., & Sidner, C. (1996). Email overload: Exploring personal information management of email. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 276-283.
- Whittaker, S., Matthews, T., Cerruti, J., Badenes, H., & Tang, J. (2011). Am I wasting my time organizing email?: A study of email refinding. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3449-3458.
- Williams, E. J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Chemistry*, 2(2), 149-168.
- Williams, P., & Hennig, C. (2015). Effect of web page menu orientation on retrieving information by people with learning disabilities. *Journal of the Association for Information Science and Technology*, 66(4), 674-683.
- Zhao, S., & Balakrishnan, R. (2004, October). Simple vs. compound mark hierarchical marking menus. In *Proceedings of the 17th annual ACM symposium on User interface software and technology* (pp. 33-42). ACM.

Appendix A

Post-design and post-test questions

Table A1: Satisfaction survey questions (for experiment 1)

Category (Experiment 1)	Corresponding questions of surveys (Experiment 1)
Overall Preference	7-point Likert scale question: <ul style="list-style-type: none"> • “I like this design very much” Close-ended question(s): <ul style="list-style-type: none"> • “Which of the three designs do you like best?” • “If you can change the move-to function in your email app to one of the designs below, which design would you like to choose?”
Efficiency	Close-ended question(s): <ul style="list-style-type: none"> • “Which of the three designs do you think is the most efficient way to put the email into a folder?”
Ease of Operation	7-point Likert Scale question(s): <ul style="list-style-type: none"> • “This ‘move-to’ button is very easy to operate.” Close-ended question(s): <ul style="list-style-type: none"> • “Which of the three designs do you think is the easiest way to put an email into a folder?”
Aesthetic Pleasure	Close-ended question(s): <ul style="list-style-type: none"> • “Which of the three designs do you think is more aesthetically pleasing?”

Table A2: Satisfaction survey questions (for experiment 2)

Category (Experiment 2)	Corresponding questions in surveys (Experiment 2)
Overall Preference	7-point Likert scale question: <ul style="list-style-type: none"> • “I like this design very much” Close-ended question(s): <ul style="list-style-type: none"> • “Which of the two designs do you like best?”
Discoverability	7-point Likert scale question: <ul style="list-style-type: none"> • “This ‘move-to’ button is very easy to find.”
Ease of Operation	7-point Likert scale question(s): <ul style="list-style-type: none"> • “This ‘move-to’ button is very easy to operate.” Close-ended question(s): <ul style="list-style-type: none"> • “Which of the two designs do you think is the easiest way to

REDESIGNING 'MOVE TO' - AN EMAIL TRIAGE FUNCTION DESIGN IN MOBILE EMAIL

CLIENT

	put an email into a folder?"
Aesthetic Pleasure	Close-ended question(s): <ul style="list-style-type: none">• "Which of the two designs do you think is more aesthetically pleasing?"

Appendix B
Figures of Satisfaction Results

Experiment 1

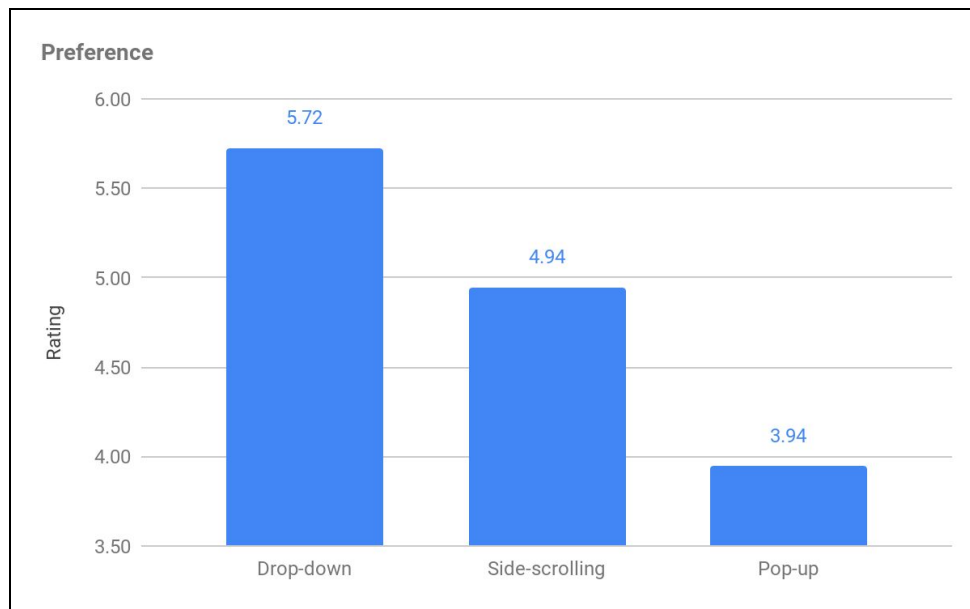


Figure B1: Preference (for experiment 1)

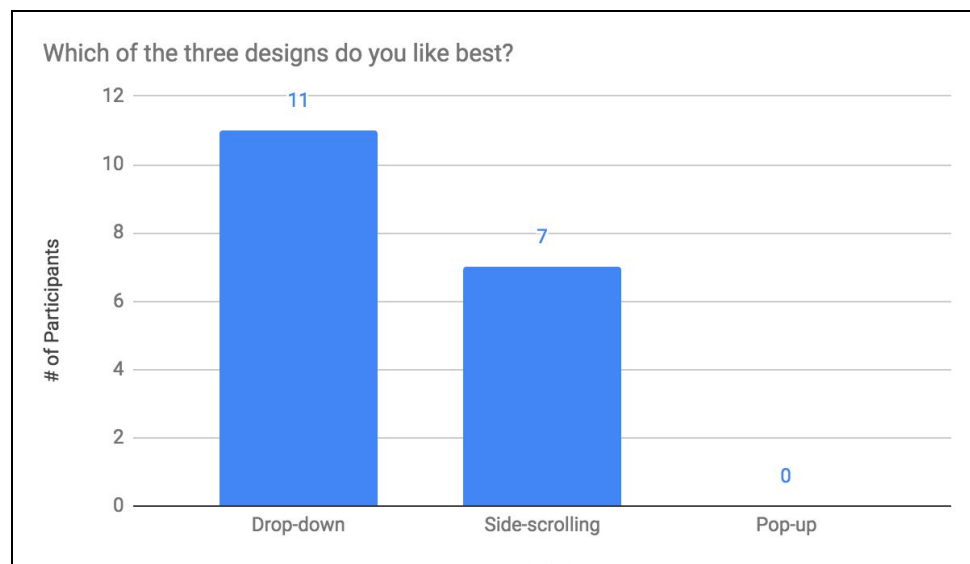


Figure B2: "Which of the three designs do you like best?" (for experiment 1)

CLIENT

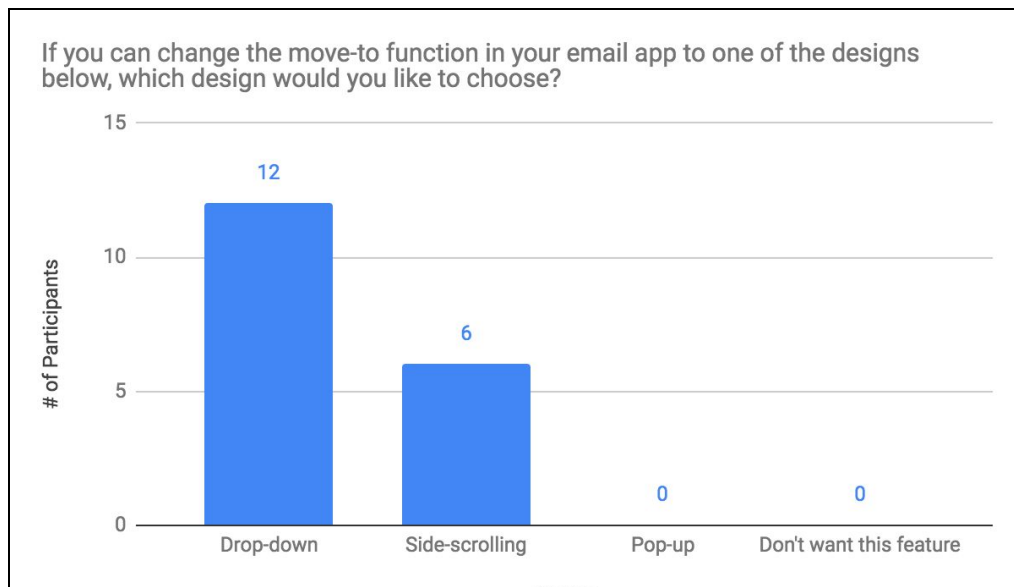


Figure B3: “If you can change the move-to function in your email app to one of the designs below, which design would you like to choose?” (for experiment 1)

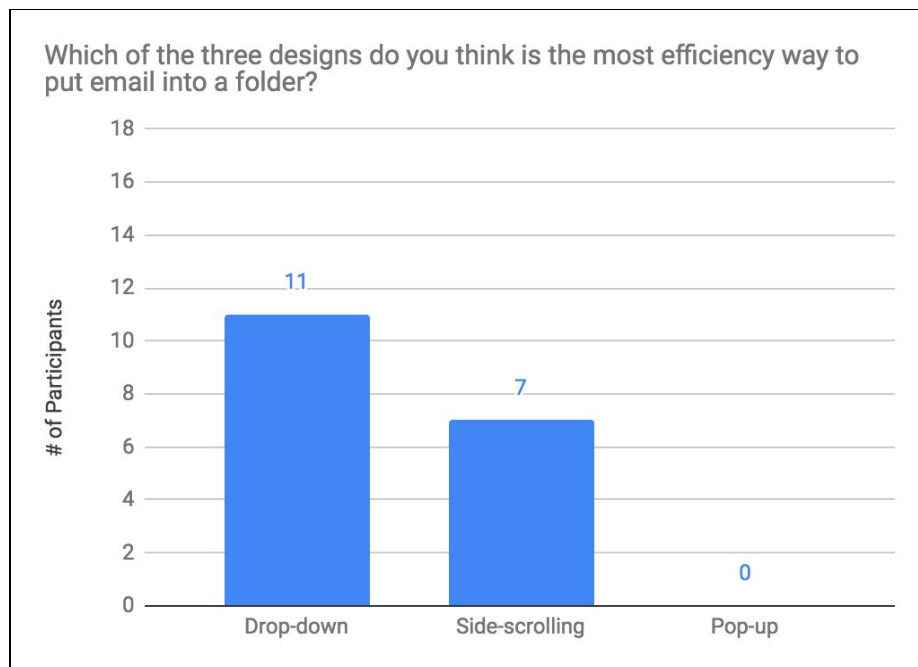


Figure B4: “Which of the three designs do you think is the most efficient way to put email” (for experiment 1)

CLIENT

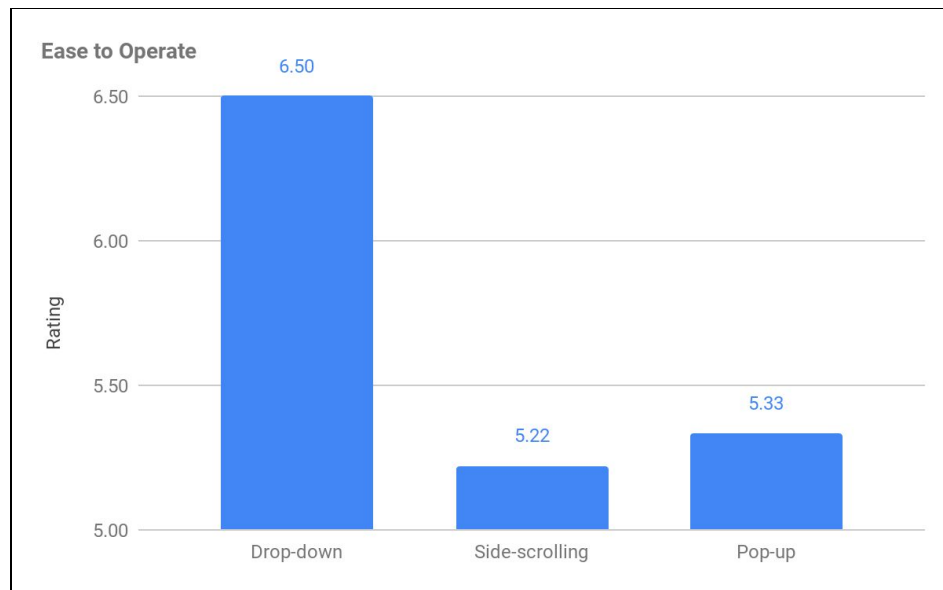


Figure B5: Ease to operate (for experiment 1)

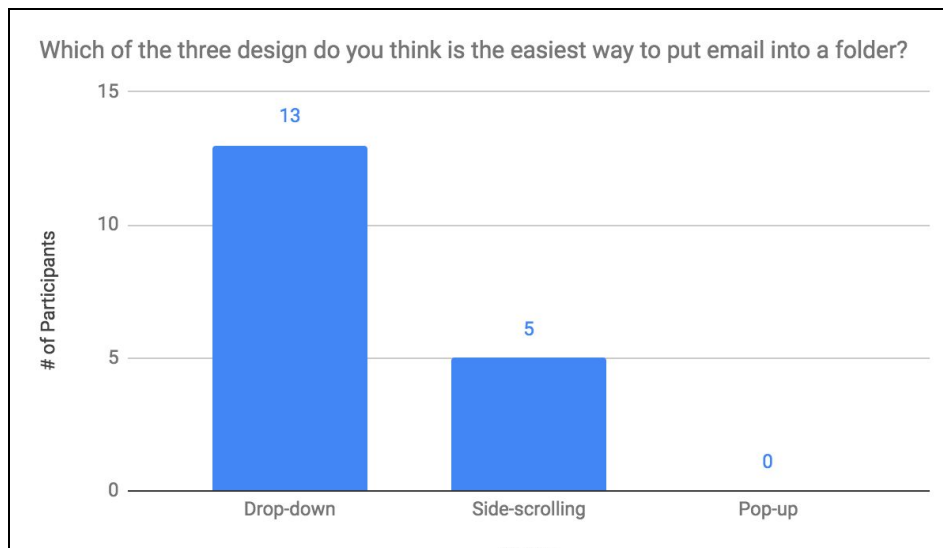


Figure B6: "Which of the three designs do you think is the easiest way to put an email into a folder?" (for experiment 1)

CLIENT

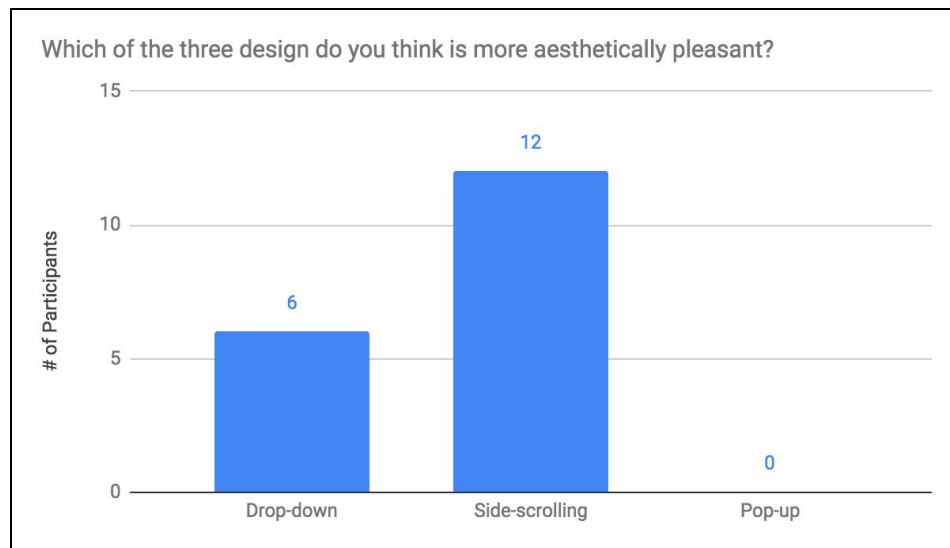


Figure B7: “Which of the three designs do you think is more aesthetically pleasing?”(for experiment 1)

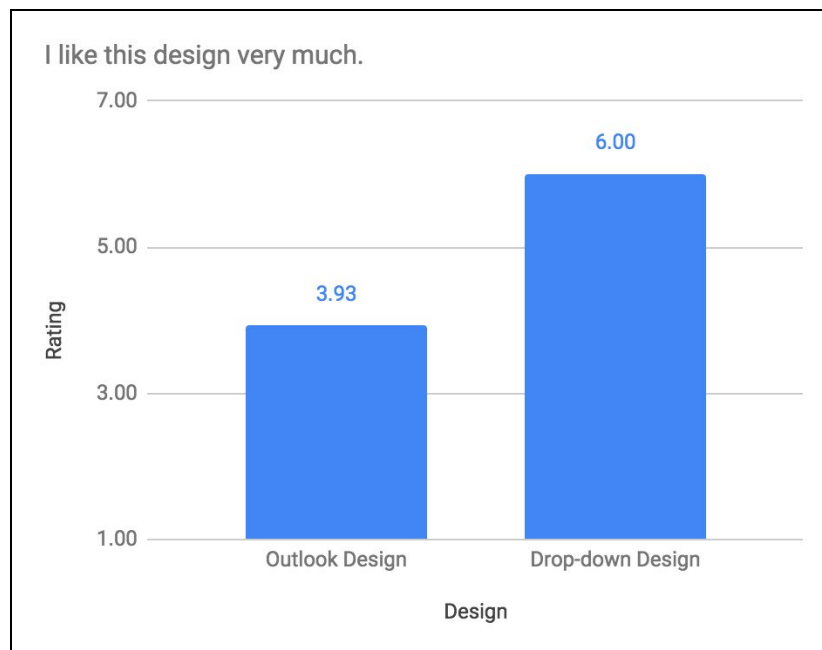
Experiment 2

Figure B8 : ” I like this design very much” (for experiment 2)

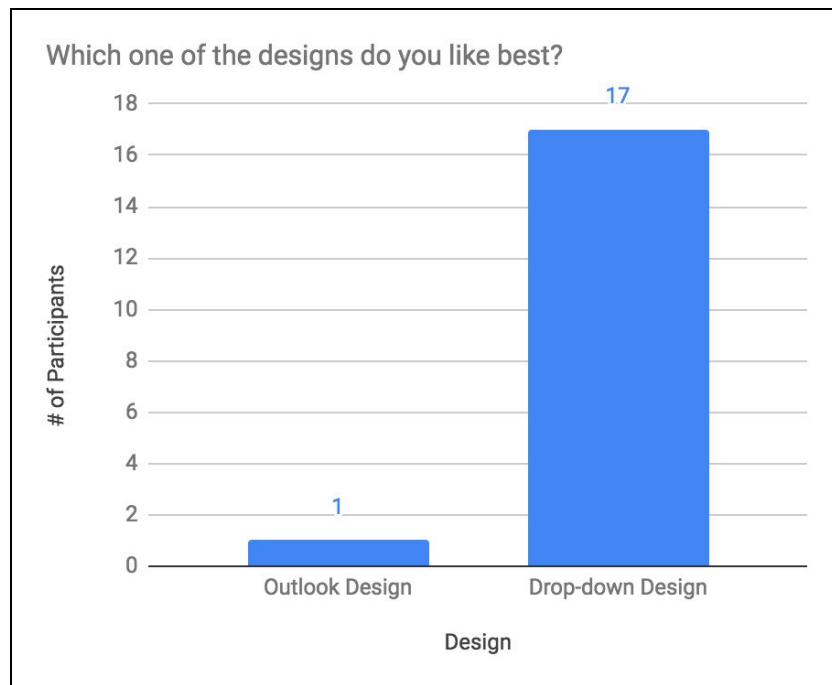


Figure B9 : "Which one of the designs do you like best?"(for experiment 2)

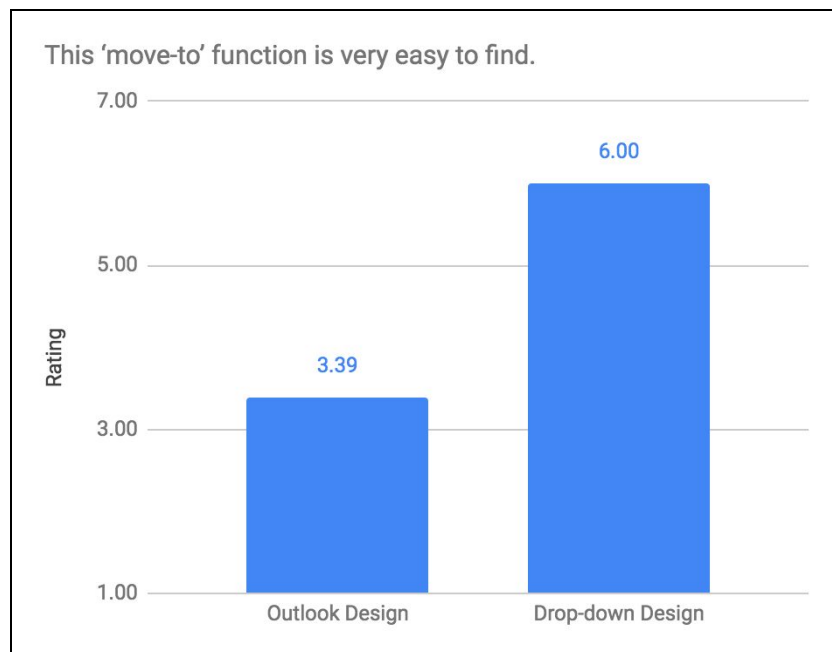


Figure B10 : "This 'move-to' function is very easy to find."(for experiment 2)

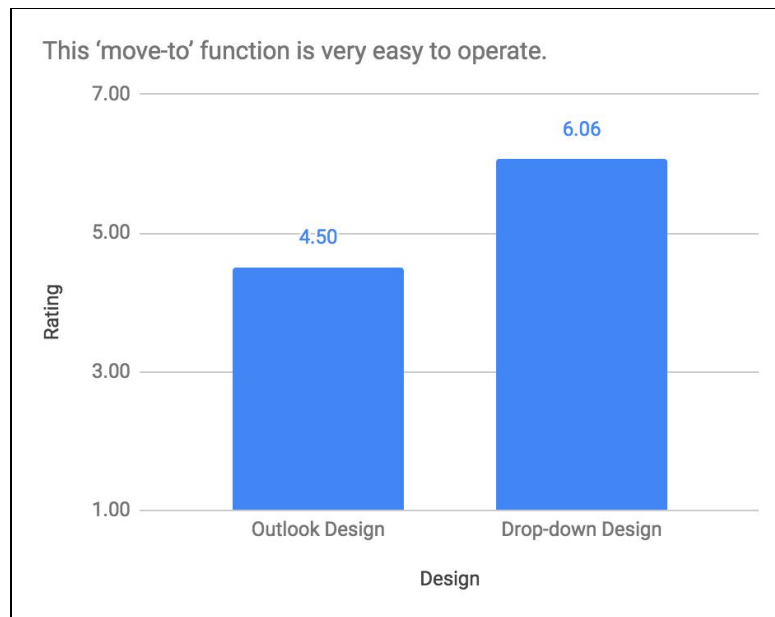


Figure B11 : "This 'move-to' function is very easy to operate"(for experiment 2)

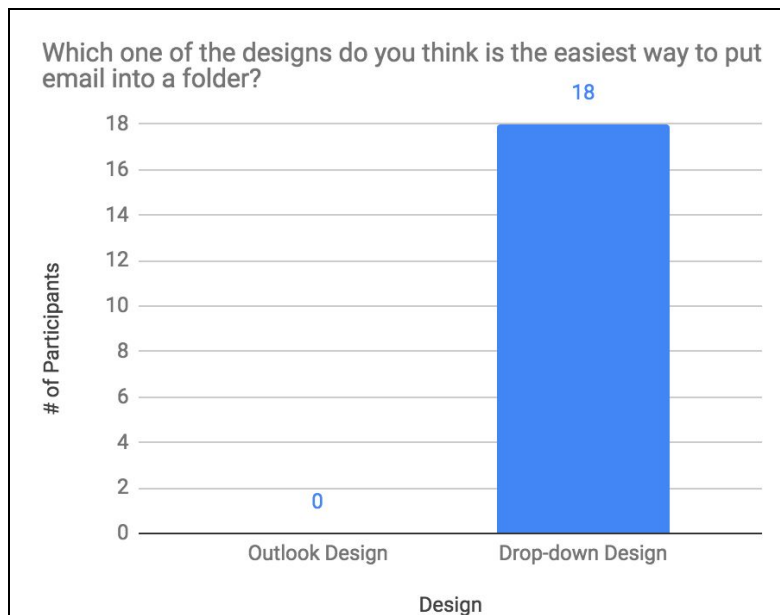


Figure B12 : "Which one of the designs do you think is the easiest way to put the email into a folder?"(for experiment 2)

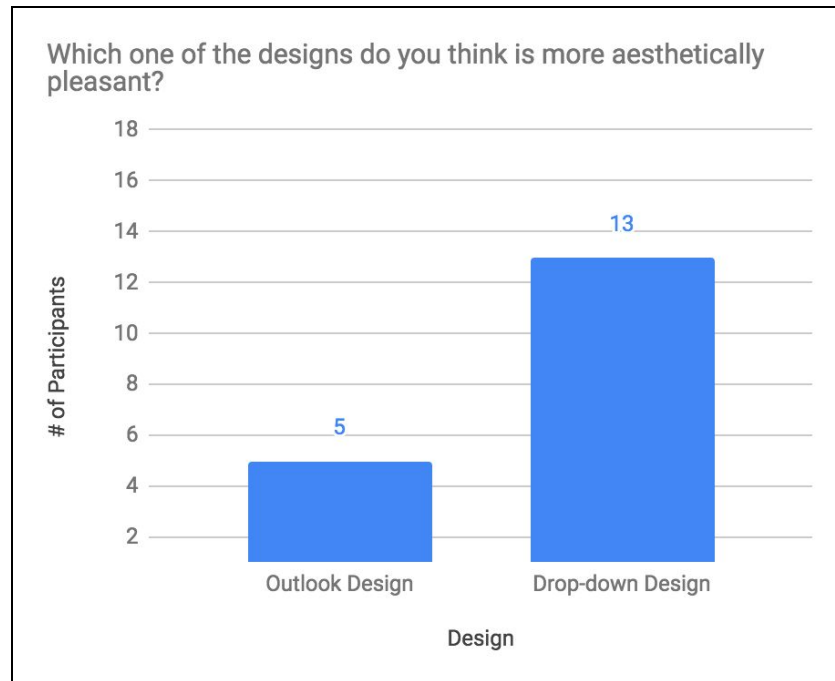


Figure B13 : "Which one of the designs do you think is more aesthetically pleasing?"(for experiment 2)