A/B Testing

By Jack Billings & Kennedy Kaufman

Our group is testing a planner app called ToDoist. The app allows for the input of plans from a digital device. It then compiles events into a mutable list with the option to view in a calendar format. Through our competitive research, we found that other popular planner apps have a ‘reminder feature’ when inputting, encouraging the use of digital reminders to assist with managing time. However, ToDoist lacks an effective reminder button. During our interviews for personas, we learned that setting digital reminders substantially assisted our interviewee’s in their time management, describing it as an effective feature for remembering tasks. In ToDoist, users need to click an obscure button, a square with an arrow in the corner, that takes them to a different page where users must scroll down to set a reminder. This is not an efficient or user-friendly design. We assume that the placement and symbol are not effective for better time management.

Our two hypotheses are as follows:

* if the reminder button is redesigned, then it will make the user’s experience more satisfactory.
* if the reminder button is redesigned, then it will facilitate better time management by encouraging the use of the reminding feature.

Based on these hypotheses, we came up with two independent variables: The position of reminder feature and the symbol of reminder feature. The dependent variables were if users understand where to set a reminder and how confident they are in finding the reminder feature. Overall, our redesign should promote a more satisfactory experience.

Next, we came up with an A/B test design based on our hypotheses, independent variables, and dependent variables. We started our test with demographics questions such as age, whether they are a student, whether they use a planner, and what methods they use to remind themselves of plans. We chose these questions to get more background information about the subjects and how they manage their time. We used Qualtrics heat map questions to ask users where they would click on the ToDoist home screen to set a reminder. Our control was the actual ToDoist homescreen. To test our independent variables we changed the location of the original reminder button to the lower bar with the other buttons. Next, we changed the image of the reminder button to a gray folder with a bell in front of it and put that button at the original spot of the reminder feature then at the lower bar with the other buttons. We decided to test within subjects so we randomized these four images and their corresponding questions. Whenever we showed each of these four images, we asked the subjects to rate how confident they were in their selection, rate how difficult it was to locate the button, whether they would change the button, and whether the current location of the reminder button effectively encourage them to set reminders. We asked these questions because they help answer our hypotheses. How confident, how difficult, and whether users would change the reminder button answered our hypothesis that if the reminder button is redesigned, then it will make the user’s experience more satisfactory. Asking if the current location of the reminder button effectively encourages users to set reminders answers our hypothesis that  if the reminder button is redesigned, then it will facilitate better time management by encouraging the use of the reminding feature.

To acquire participants for our A/B test, we posted our Qualtrics link onto the SI422 Discussion section on Wednesday. In addition, we posted links on our Facebook pages and reached out to friends, club members, roommates and family members. Knowing that most of our respondents would be students and peers, we structured our demographic questions to reflect it. For example, when asking the user their age, our options included 18 and younger, 19, 20, and 21 or older. In addition, we inquired about their planner usage, time management methodology, and occupation. This helped narrow our target population and learn more about this demographic of users.

**Analysis**

We received a total of 61 responses to our survey. Of these 75.93% of respondents were ages 19 to 21, covering a majority of our demographic. Over 80% were students, with about 40% of respondents not using any sort of planner.

Our data aligned with our insight that ToDoist digital reminder button is hidden and the symbol is not intuitive. Currently, the button is removed from other common buttons and it is depicted by a box with an arrow. We believe using a bell would help users recognize this. From our responses, 70% used digital reminders, so this could be an important feature to our users.

The following figures all depict the heat maps we received (they were randomized for the actual test). Figure one has no treatment. Although most clicks were correct, half of respondents click the wrong button. Of that, only 20 respondents were “somewhat confident”, 14 were “somewhat unconfident”, and the rest fell between. The best insight came from the fourth questions, when we asked user’s how they would change it if they could. Most responses asked for the button to be more intuitive and obvious. The results were similar when we changed the location of the button. Still, user’s struggled.

We noticed an immediate positive reaction to creating a new symbol for reminding. Almost every person clicked the correct button with 77% of respondents being confident in their choice. That same percentage found it easy to locate and half of responses would not change it if they could. More people also found this button more effective. Results were about the same when changing both location and symbol.

For the future, we learned some important things that we would change. First, we struggled for a bit to get our full responses. This made it difficult to pull meaningful insights from our data. Having an extra team member would help spread the survey to more people. Also, we only received around 18 responses from kids in SI422.

Next, we would ask a few more demographic questions to learn more about where the data is coming from. For example, some good supplementary questions would be to ask how comfortable users are with technology and their mobile phone usage. We would also pilot our survey with more detail. It took longer than we thought to craft a well designed test. Even into our third pilot, we found spelling errors or other mistakes. The little things can slip by easily.

**Appendix**

Treatments

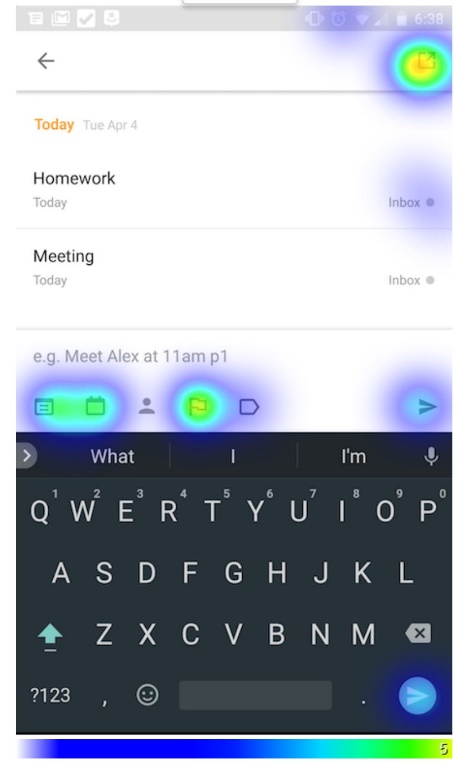


Figure 1 No Change

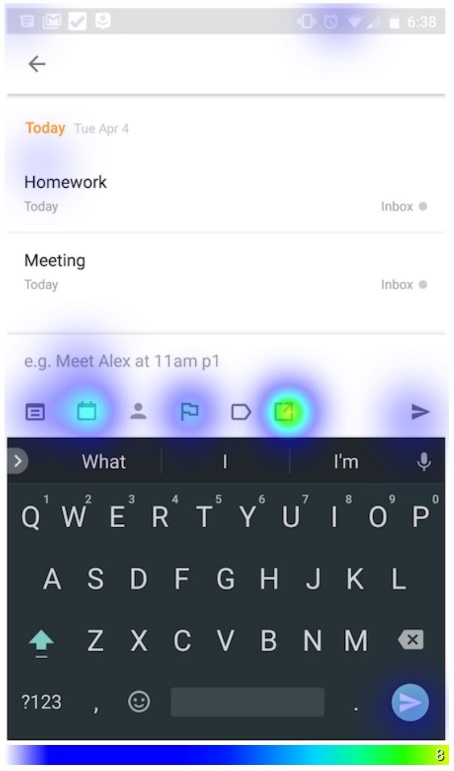


Figure 2 Changed location of digital reminder

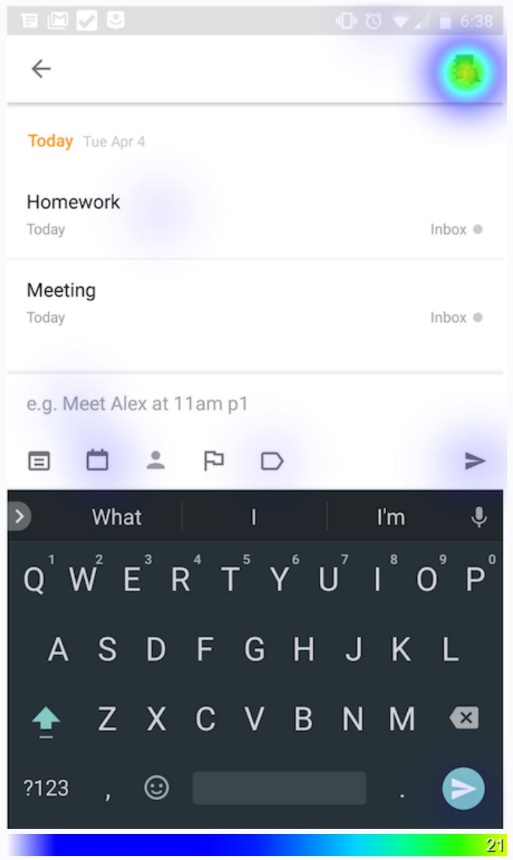


Figure 3 Change Symbol of digital reminder

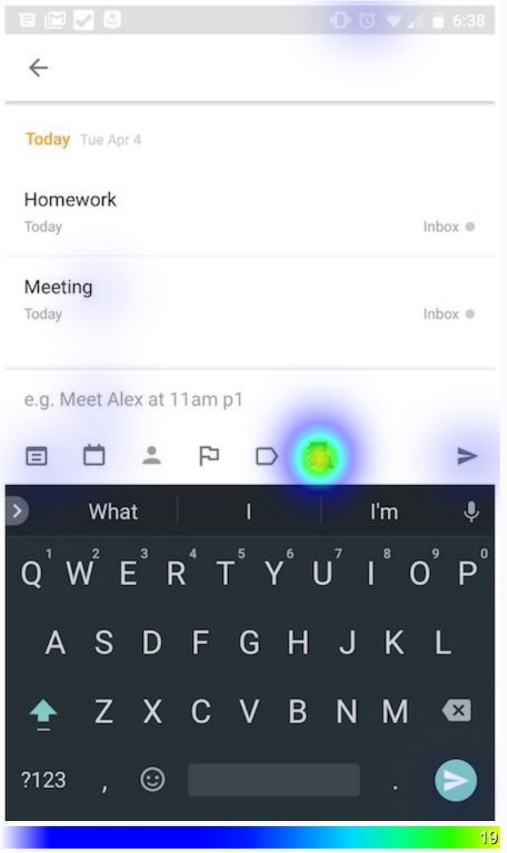


Figure 4 Change location and symbol of digital reminder

**Raw Data** (See external attached file)