

Pilot Experiment Using Lo-Fi Prototype

1 Study Design

1a Goals of the Experiment

The main goal of the experiment was initially to improve the general UI's usability, and more specifically, to clarify the difference between setting a timer and a stopwatch. This goal emerged from confusion we ourselves faced when implementing the timer and stopwatch features in our prototype. After inspecting our prototype to find something UI-related worth conducting the user study on, we identified the lack of clarity that the separation between both timed events had. With this in mind, we started brainstorming a new UI ("UI 2") that had a clearer separation. This led us to pose the following question: how do different configurations of the timer/stopwatch interface affect how intuitive it is to use, measured by the total task completion time and the total clicks used? By measuring both of these metrics on different UIs, we would get a better sense of which UI simplifies the use of the timed features in our application and thus, achieve our goal of designing an application with a more intuitive UI. Both timed features are accessed when the "play" icon is pressed on the calendar page, but how can we make the distinction between both clear while maintaining a non-cluttered interface?

1b Study Protocol

The goal of the experiment is to understand which of two UI options (see Appendix i) is clearer and easier for users to set timed events (timer or stopwatch). Additionally, more general feedback regarding the UI is gathered for future improvement. Participants for this experiment were gathered on a convenience basis given time limitations; all participants are second or third-year Computer Science students.

i Interview Protocol

The following outlines the step-by-step process used for each participant.

1. The participant is welcomed and debriefed with a short introduction: "This study aims to test which of our low-fi prototypes for a calendar app with timer functionality is most convenient and clear to use. We will present two different prototype UIs and ask you to perform the same two tasks on each. Lastly, we will ask you some general questions about our UI and ask for any additional feedback for improvements. Please also note that this experiment is performed on a low-fi prototype, so not all functionality is implemented; not all buttons are functional, so only try to interact with elements that seem relevant to the task at hand."
2. The researcher asks whether the participant has any further questions, answers these and ensures that they are comfortable participating in this experiment.
3. The participant takes a seat in front of the provided computer.
4. The researcher loads the first UI.
5. The first task is read out loud to the participant: "The first task is to start a 2-hour timer using the UI".
6. The researcher starts the stopwatch and initializes the click counter once the participant is ready, and the participant proceeds to complete the task.
7. Once the participant has completed the task, the stopwatch is stopped and the resulting time and number of clicks are recorded by the researcher.
8. The second task is read out loud to the participant: "The second task is to start a stopwatch using the UI".
9. Steps 6 and 7 are repeated.
10. The researcher loads the second UI.
11. Steps 5 to 9 are repeated.
12. The participant is asked a series of qualitative questions (see: Appendix ii) about the UI. Lastly, the researcher asks for any final remarks, potential improvements, and other general feedback.
13. The participant is thanked for their participation and the experiment is concluded.

1c Independent and Dependent Variables

i Independent variable

The independent variable is the configuration of the timer/stopwatch UI (either UI 1 or UI 2). Each participant will perform two identical tasks on both configurations. Due to the different layouts of both UIs, we are interested in how they will influence the outcomes of the measured dependent variables.

ii Dependent variables

The dependent variables are the total time taken (in seconds) and the total clicks performed to complete a given task; these both depend on the type of interface that the user interacts with. To measure both dependent variables on the low-fi prototype, interactive elements were implemented using Figma; basic functionality to move between screens was supported (clicking a certain element would bring the user to another screen view).

Time taken This was measured manually using a stopwatch. The stopwatch was started as soon as the participant declared "go" and stopped as soon as the task was completed. After each task, the researcher took note of the measured time.

Number of clicks This was measured by the use of the [Mac Mouse Click Counter](#) application. The counter was initialized to 0 after the task was read out. As soon as the participant completed the task, the researcher took note of the number of clicks before moving on.

2 Study Result

2a Narrative Report on the User Study

i Setup

As mentioned in Section 1a, our experiment aimed to investigate user preferences between two different UI layouts. The main difference between both layouts is the way that the timer and stopwatch functionalities are separated. Our motivation for this user experiment stemmed from dissatisfaction with our initial UI design and doubts about how to best provide both functionalities while maintaining a clean and minimalist interface.

The first UI that was presented to participants was our initial prototype and layout, where a timed event is by default a stopwatch, and a timer is only initialized by explicitly setting a duration. The second UI features a toggle functionality to create visual separation between a timer or stopwatch event, without the same implicit expectation that only setting a duration for a timed event changes it from a stopwatch to a timer. Our hypothesis was that users will find the distinction between both types of timed events clearer in the second UI since the first UI does not give any visual hints for the distinction. This hypothesis was tested by considering both dependent variables of the experiment (see Section 1c) as well as gathering general feedback from participants.

Ideally, we would have wanted to test on a minimum of 12 participants, though due to time constraints, the user experiment was conducted on only 4 different participants, all of whom are Computer Science students at the VU and selected semi-randomly; all participants are acquaintances that were present at the university on the day of experimentation. As this study does not require any group division, all participants conducted identical experiments. Step-by-step details outlining the general study protocol are described in Section 1b.

The two main tasks that we asked participants to execute were to start a 2-hour timer and start a stopwatch on each of the UIs. For each of the tasks, we measured time and number of clicks to give us a representative metric of the UI's ease of use; we expected that tasks completed on the clearer UI would result in less time and clicks.

Given that both tasks require some interactive elements in the prototype, we used Figma to implement the necessary user interactions. For instance, when the user clicks the play button, Figma brings the user to the resulting timer/stopwatch screen. It should be noted that the app functionality was not at all implemented for these experiments, user interactions were a simple switching between static screens. Example snippets demonstrating how the interactive “flow” was implemented for Figma screens (or “states”) are shown in Appendix iii.

In an attempt to maintain consistency among different participant results, we considered that the type of laptop should be a controlled variable. This way, we minimized the effects that different devices and their input types would have on the experiment's outcome. All participants conducted the experiment using a MacBook Pro and the trackpad native to it for cursor input. Additionally, we ensured that all participants were seated in front of the device in a quiet environment with minimal noise to reduce the effect that distractions would have. None of the participants showed any signs of distractions.

ii Outcome

When analyzing the outcomes of our experiment, we found the following patterns among participant behavior:

- It took the most amount of time for participants to understand the first UI initially when performing the first task. Participants clicked many different icons, trying to navigate the UI, however, because the UI was only a low-fi prototype, not all functionality existed. Participants 1 and 2 tried to enter labels for the timed events using keyboard input but soon realized that this was unresponsive, and all participants tried to drag the duration bar to set a 2 hour timer which was not implemented due to Figma's limited functionality. Nevertheless, in general, participants quickly adapted to the UI and performed the following tasks faster and using less clicks (there was less “exploration” involved in measurements).
- UI 1's lack of visual separation between timer and stopwatch was emphasized by the experiments. When asked to start a timer using UI 1, three of the four participants tried to engage with other icons; they all assumed that the location of a stopwatch would be separate from the timer. Due to their confusion, they were given a hint by the researcher that the stopwatch could be initialized from the same place as the timer. In all cases, we received feedback targeting exactly this confusion about separation, as well as clear preference of UI 2 to UI 1. “It's way more explicit” was a comment made by Participant 4 when asked about their preference of UI 2.
- The use of the bookmarks icon, used to navigate to presets for timers, was unclear to all participants. This information is not so much related to the main focus of our study to improve the separation between timer

and stopwatch, but was notable enough of a feedback point to include as a UI-related remark. Participant 2 additionally remarked that it was unclear if the listed items indicated how much time activities took in general, if they were recent events, and why the dates of each item was missing; this feedback point gave us insight as to an additional improvement needed in the application, and a potential case for a future user study.

- While “exploring” the UI’s general layout, two participants discovered the toggle feature and were able to predict its use. This was an unexpected outcome of our user study, as we did not plan for users to interact with this element. Nevertheless, it provides us valuable feedback for positive and well-implemented aspects of our UI.

In conclusion, the user experiments supported our hypothesis; UI 2 was more clear to all participants. A more in-depth analysis with respect to the gathered metrics is discussed in Section 2b.

2b Structured Result

The raw results gathered from the user experiment are provided below in Table 1:

Participant	UI	Task	Time taken (s)	Number of clicks
1	1	1	33.83	25
		2	20.35	8
	2	1	12.6	12
		2	8.31	7
2	1	1	37.82	18
		2	13.4	6
	2	1	15.58	11
		2	6.58	3
3	1	1	37.61	14
		2	18.66	9
	2	1	12.89	13
		2	8.32	6
4	1	1	34.82	20
		2	17.28	7
	2	1	13.68	10
		2	7.33	5

Table 1: Experiment Results

We decided to plot the gathered results as histograms to get a more visual distinction between different user performances and possible trends. The histograms feature four distinct colors to separate different participants’ performances. Furthermore, the UI used for each task is represented by different patterns in the bars; the X pattern denotes UI 1 and O denotes UI 2. Note that all provided histograms specify the type of task performed; task 1 corresponds to the “Timer Task” and task 2 corresponds to the “Stopwatch Task”.

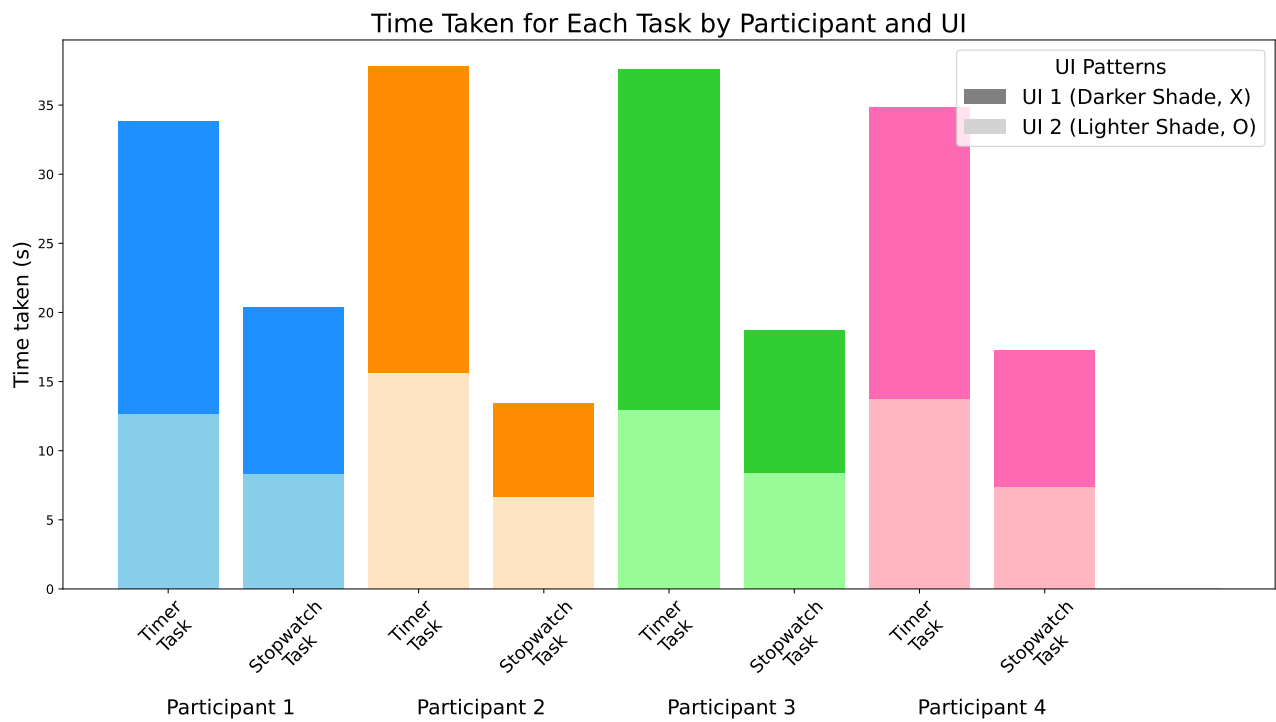


Figure 1

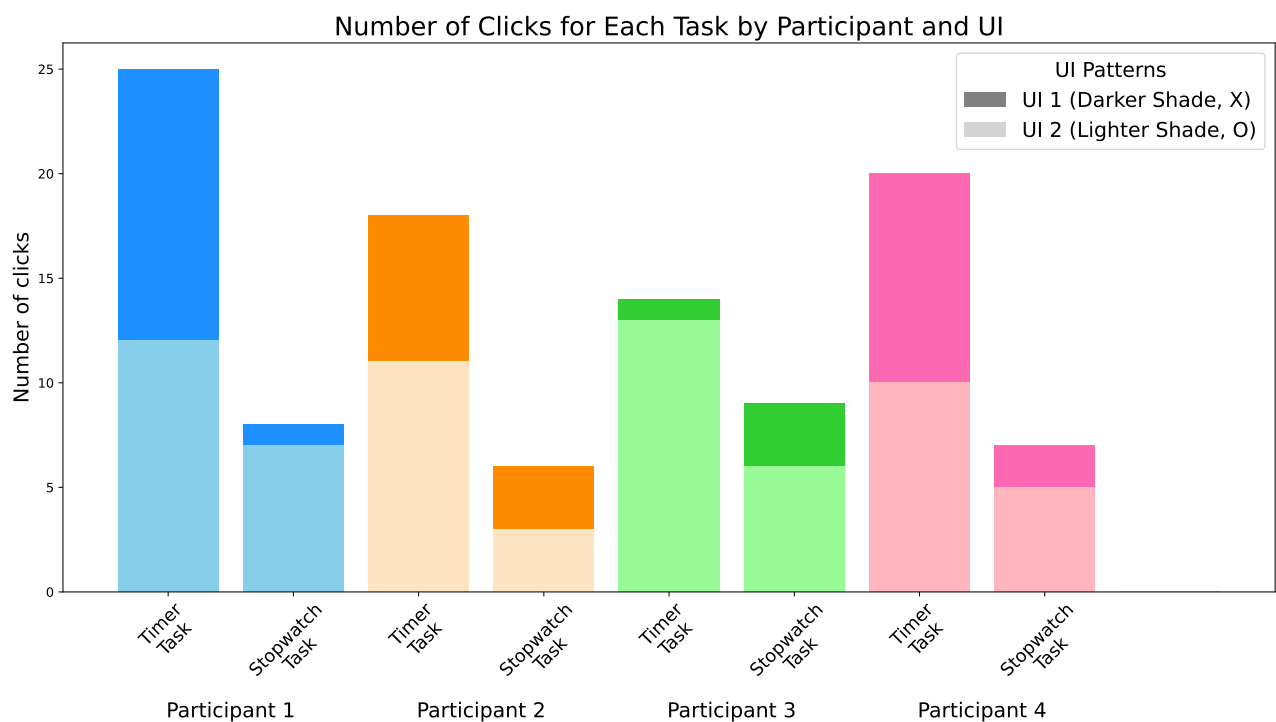


Figure 2

In both histograms (Figures 1 and 2), we see clear similarities emerge in participant performance:

Task related observations: Task 1, the timer task, consistently took more time and more clicks than task 2, the stopwatch task. This can be explained by the order in which the tasks were provided. Users were first asked to complete the timer task, so initial confusion about the UI likely impacted the total duration and number of clicks used. After completing the first task, users were more familiar when asked to complete the second stopwatch task, thus, reducing the total time and clicks needed.

UI-dependent results: Among all participants we see a significant decrease in the competition times and

number of clicks when asked to perform any task on UI 2. This can be attributed to the increase in clarity of UI 2's distinction between timer and stopwatch.

To get a more general overview of our results, we computed the average time and number of clicks it took for each task to be completed using the two different UIs. The results are demonstrated in Table 2:

UI	Task	Average time taken (s)	Average number of clicks
1	1	36.02	19.25
	2	13.69	11.5
2	1	17.42	7.5
	2	7.64	5.25

Table 2: Average Results

Lastly, we plotted the data from Table 2 as two separate histograms, to emphasize the trends observed from the histograms in Figures 1 and 2. Each bar represents the relevant dependent variable being measured where the bar's color indicates whether the task was performed with UI 1 or UI 2. Additionally, error bars are included in the figures to indicate the absolute minimum and maximum measurements obtained in relation to the calculated mean.

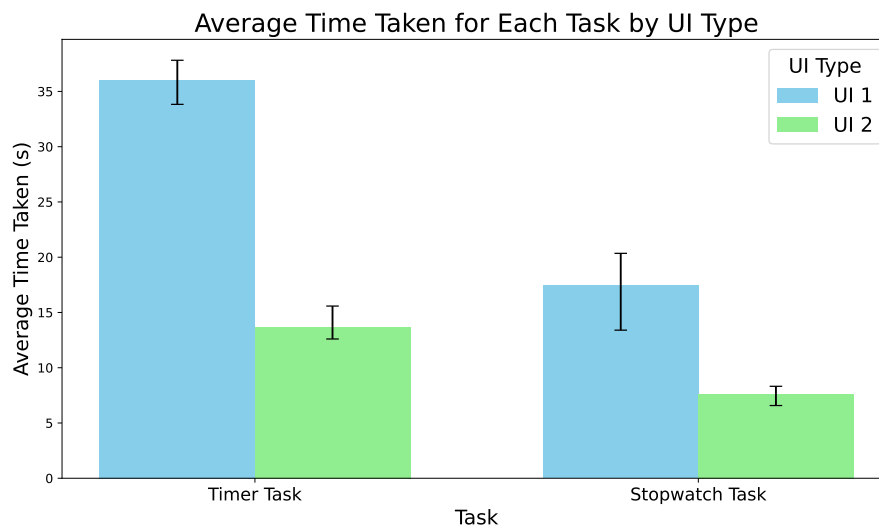


Figure 3

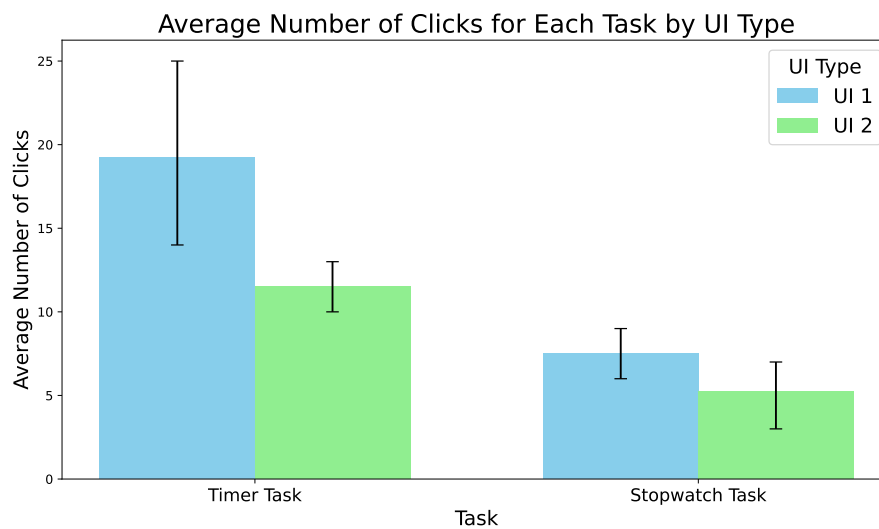


Figure 4

From the above Figures 3 and 4, we again see that UI 2 consistently outperforms UI 1 in all metrics. Furthermore, it is notable that the error bar on the UI 1's timer task's average number of clicks is significantly larger than other error bars observed. This is likely due to the initial unfamiliarity that participants experienced when first presented with the UI; being the very first task using this interface, some participants tried clicking every possible button to “look around” while others navigated more directly to the play button to initialize the timer. This observation is consistent with the average time it took for participants to complete the timer task on UI 1.

Another notable observation is the small size of the error bars for UI 2's average time on the stopwatch task. This was the last task that users were presented with, and thus, they were a lot more used to the UI and how to navigate it. Additionally, the distinction between timer and stopwatch is more explicit in UI 2, having already primed participants on what steps had to be taken to complete this second task during the competition of the first.

The reduction in average time and number of clicks is further evidence to support our initial hypothesis: UI 2 is clearer, easier to use, and preferable to UI 1. The unclear separation between the types of timed events in UI 1 seemed to cause similar confusion in participants as it did to us; the study was successful!

3 Reflection

3a Validity of Experiment

Our experiment was very insightful given the fact that we have found a clear winner between our two prototypes and have gained both valuable feedback and data that will help us improve the next iteration of the app. The winner was easy to infer from the shorter time and the lower amount of clicks required to do a task with the second UI configuration.

The lo-fi prototype ended up being a decent tool despite needing a lot of adjustments. Given the nature of the app, having a timer and a calendar in one UI, there was some confusion among the users about how to approach it at first. However, this combination of features is representative of our final version, and the statistics we gathered along with their feedback proved to be very important in making the app easier to use without removing functionality. We also implemented some primitive features that are adjacent to the test (i.e. interacting with the cog wheel brings the user to the settings screen). Nevertheless, the partial functionality offered by our low-fi posed some limitations to our measurements and, hence, our experiment's validity. The lack of fully functional interactive elements resulted in a longer measured task completion time and more clicks. One such feature that was not implemented but commonly attempted was for participants to attempt dragging the slider to set a 2 hour timer.

Another limitation of our study was the setting's unfamiliarity in regards to the unknown UI to our participants. One reason which may have contributed to this is that we did not provide time before the tasks for participants to navigate and explore the UI and its features, potentially having inflated the actual clicks and time it would have taken to complete the first tasks presented on a new UI. However, we believe that the initial time it took for participants to get familiar with the UI did not significantly impact the trends we observed nor the aim of our study; we were still able to conclude that UI 2 is more user-friendly.

Our test users are comprised exclusively of computer science students and as such the feedback and the statistics that we gathered may be heavily biased. That is, the time and clicks measured from our tech-savvy participants may not be fully representative of our more general target user. Hence, we risk following feedback that might guide us to make an app focused on students. Moreover, the number of participants (4) does not provide enough data points to warrant stable enough results to be reproduced accurately.

3b Prototype Improvements

As aforementioned, the experiment showed us that the second UI configuration was superior to our previous UI 1, providing us a starting point for an improved prototype. The added stopwatch/timer toggle along with the greyed-out time bar appeared to make the difference between the two modes much more explicit without adding too much visual complexity, resulting in much shorter times and fewer clicks. As such, we will further develop this configuration by making the toggle button and the time adjustment bigger, using a darker gray color for the time adjustment bar, and increasing the hitbox of the toggle button. Moreover, we will add the ability to type in the time limit by clicking the elapsing time, since we observed that this was an expected feature by our users. We will also add the ability to cancel a timer/stopwatch.

From our study, we also received very valuable feedback related to other features. We learned that the bookmarks section is confusing, the current time is not represented in the calendar, our statistics view is not explicit enough, and our navigation bar is not clear. To improve the bookmark functionality we will add a short description and change the name to presets. The calendar view will have a red line to point at the current time to give a better sense of the current "position". Percentages will be added to the pie chart to make the data more readable, and the navigation icons will be better highlighted with appropriate contrast. This way our app will provide an experience with much less friction and confusion for the user.

Appendix

i UI Types

Below, we demonstrate the different UI's used in the experiment. The first UI's layout is identical to that of our initial prototype. The second UI demonstrates an adaptation made to the timed event functionality of our initial prototype.

UI Type 1

The flow for UI 1 can be found at the following [link](#).

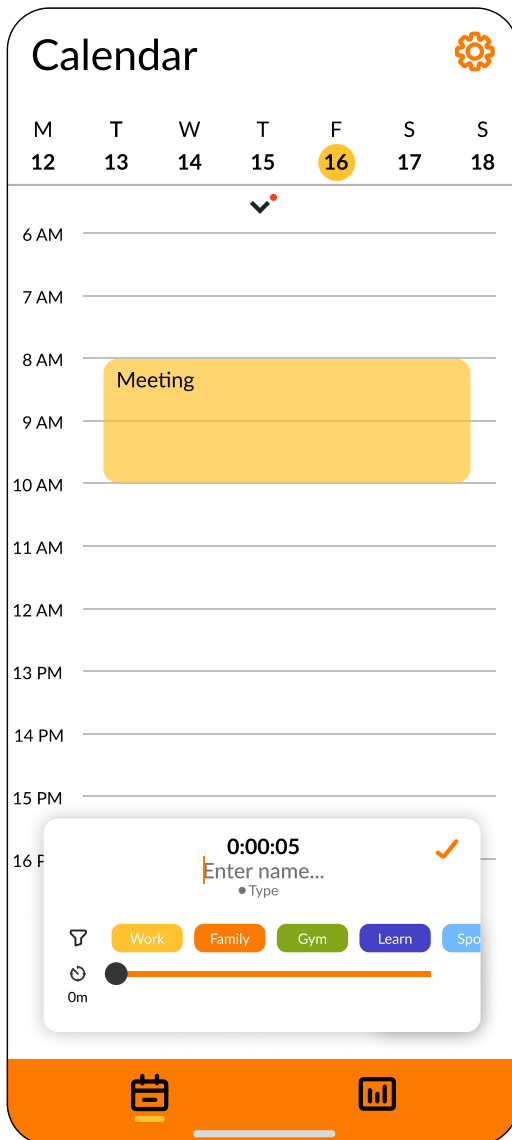


Figure 5: Stopwatch UI 1

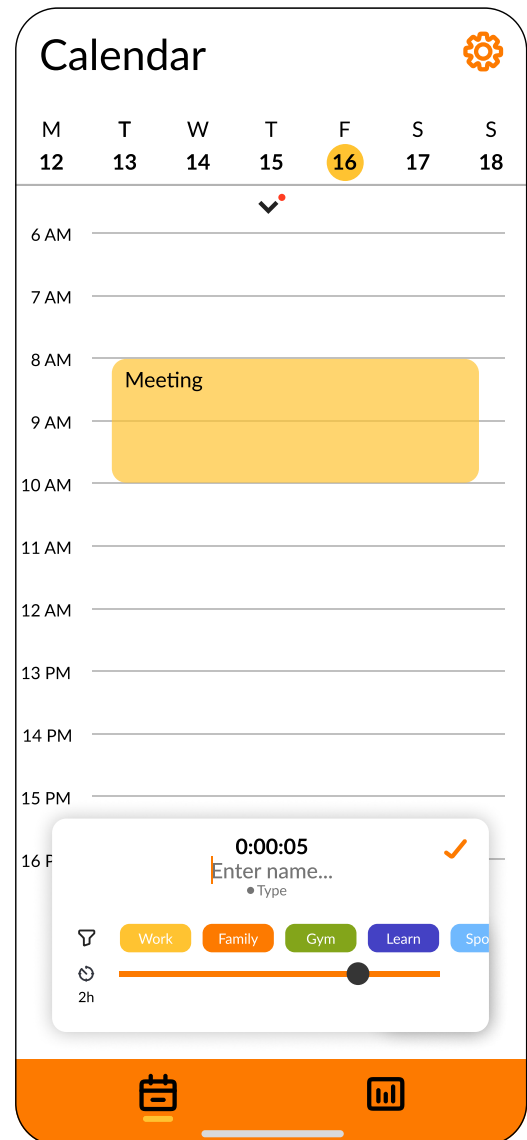


Figure 6: Timer UI 1

UI Type 2

The flow for UI 2 can be found at the following [link](#).

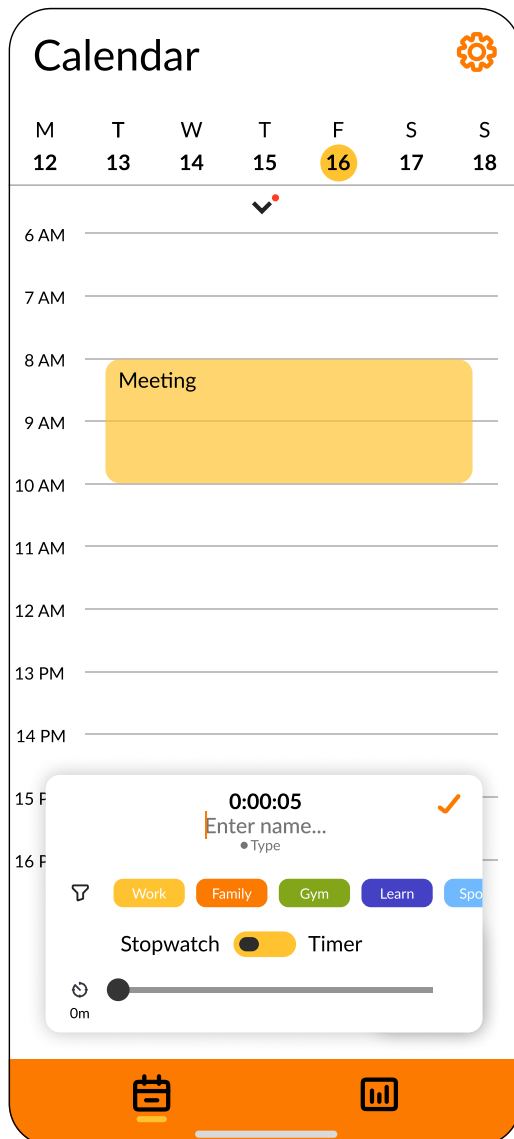


Figure 7: Stopwatch UI 2

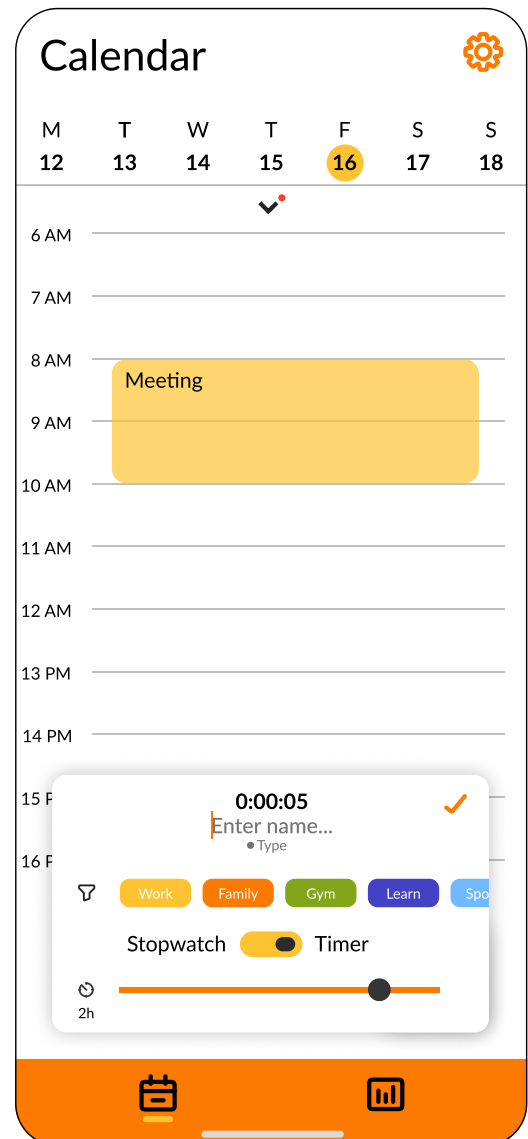


Figure 8: Timer UI 2

ii Qualitative Questions

The following questions were asked to extract qualitative feedback from participants of the user experiment. These questions were posed at the end of the experiment, after participants completed all tasks on both UIs.

UI Comparison: Which of the two UIs did you prefer and why? Do you have any suggestions to make the distinction between timer and stopwatch clearer?

Saved Events / Bookmarks: Do you think the edition of a bookmarks view is helpful? Would you use it?

Toggle: Is the toggle feature clear and easy to use? Would you use it?

Statistical overview: Are the visualizations used in the statistical overview clear and easy to understand? Are there any other kinds of statistical elements that would make it easier to understand how you use your time?

Layout: Are the calendar and statistical overview appropriately separated? Would you prefer if they were on a single screen?

Icons: Does the use of icons help you understand the functionality of the buttons? Can you guess what each of the icons represents?

Color: Do you think the use of color is appropriate? Is there something you would do differently?

Usage: Can you see yourself using this app? What advantages and disadvantages does this app have over other calendar apps?

iii Figma Flow Snippets

Here are some example snippets from our Figma Flow implementation. Both snippets feature the design and interaction in UI 2. The first snippet is more zoomed in and focused on the timer/stopwatch screens and the second snippet demonstrates a more zoomed-out view of the general landscape.



