

Usability Test Results

The results from the tests, based on metrics, compared to goals and how severe each issue is.

Times

Note: Times for User Test 1 for Tasks 2 to 5 weren't filled in. This was due to timekeeping errors for the first user test.

Tasks	Time (seconds)				
	User Test 1	User Test 2	User Test 3	User Test 4	User Test 5
Task 1	30	15	8.84	8.28	15
Task 2	–	40	40	24.18	31.24
Task 3	–	15	19	14.26	17.28
Task 4	–	16	7.54	8.2	19.57
Task 5	–	40	13.34	6.27	11.72

NASA TLX Survey

NASA TLX Survey Results

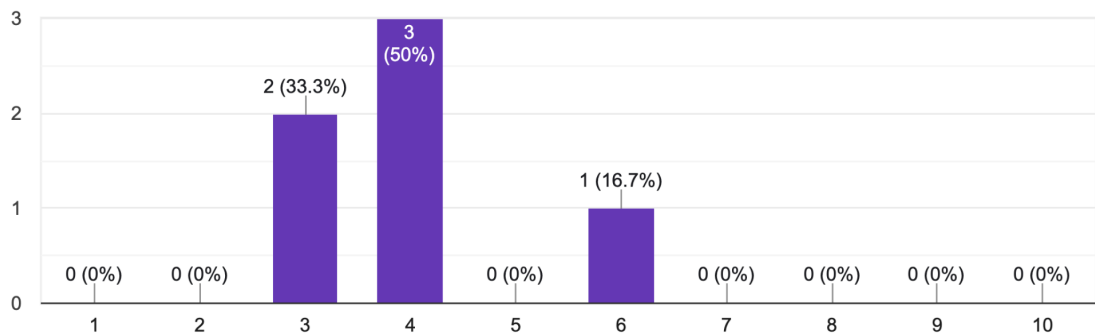
User	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
User 1	3 ▾	1 ▾	1 ▾	8 ▾	4 ▾	1 ▾
User 2	4 ▾	2 ▾	2 ▾	7 ▾	7 ▾	1 ▾
User 3	6 ▾	7 ▾	5 ▾	7 ▾	7 ▾	6 ▾
User 4	4 ▾	1 ▾	2 ▾	6 ▾	5 ▾	3 ▾
User 5	3 ▾	1 ▾	1 ▾	9 ▾	3 ▾	1 ▾
User 6	4 ▾	6 ▾	5 ▾	8 ▾	4 ▾	6 ▾

Scaled NASA TLX Survey Results

User	Mental	Physical	Temporal	Performance	Effort	Frustration
User 1	6 ▾	2 ▾	2 ▾	16 ▾	8 ▾	2 ▾
User 2	8 ▾	4 ▾	4 ▾	14 ▾	14 ▾	2 ▾
User 3	12 ▾	14 ▾	10 ▾	14 ▾	14 ▾	12 ▾
User 4	8 ▾	2 ▾	4 ▾	12 ▾	10 ▾	6 ▾
User 5	6 ▾	2 ▾	2 ▾	18 ▾	6 ▾	2 ▾
User 6	8 ▾	12 ▾	10 ▾	16 ▾	8 ▾	12 ▾

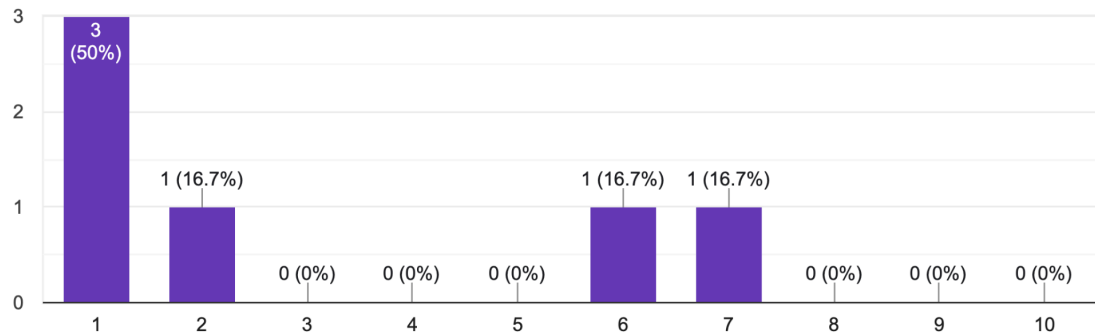
Mental Demand How mentally demanding were the tasks?

6 responses



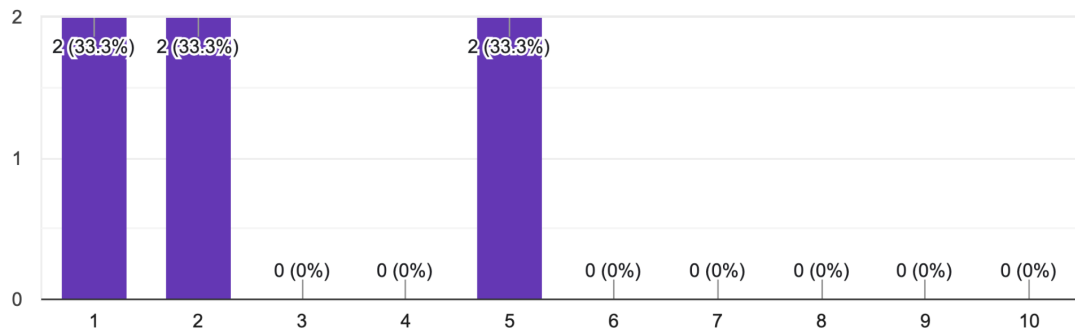
Physical Demand How physically demanding were the tasks?

6 responses



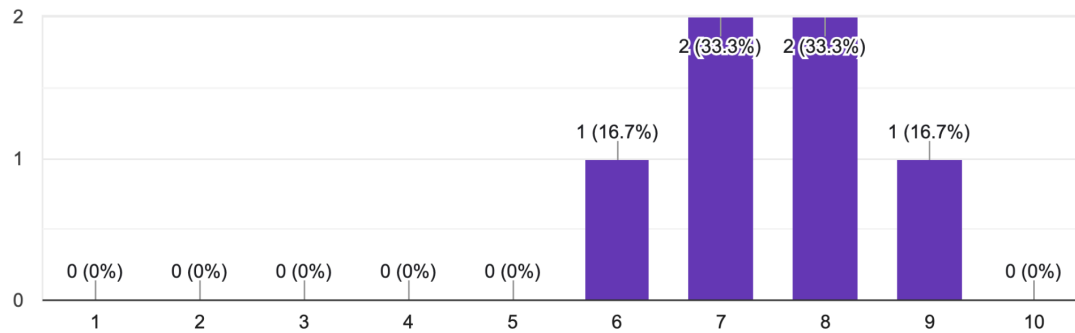
Temporal Demand How hurried or rushed was the pace of the tasks?

6 responses



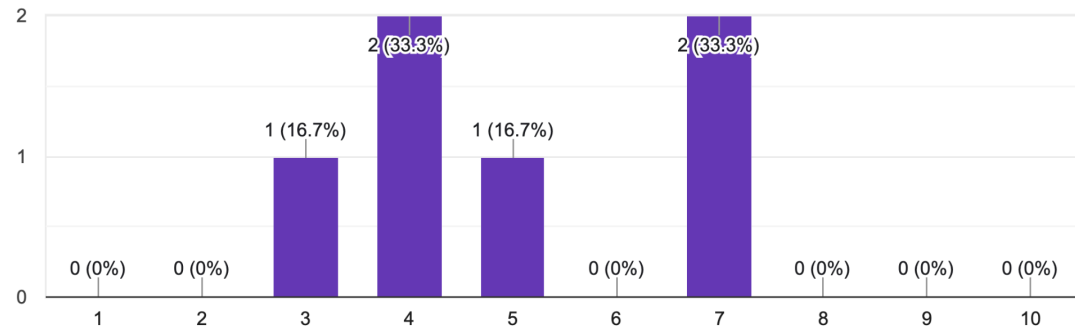
Performance How successful were you in accomplishing what you were asked to do?

6 responses



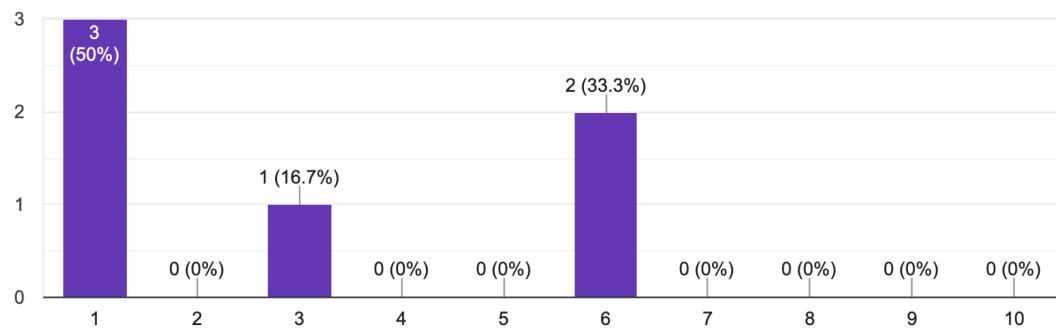
Effort How hard did you have to work to accomplish your level of performance?

6 responses



Frustration How insecure, discouraged, irritated, stressed, and annoyed were you?

6 responses



Usability Goal Calculations: TLX Scores

Usability goals defined in our testing plan gave us an average score to aim for per category in the NASA TLX template.

The following are the average results based on the survey conducted:

Mental Demand: $(6 + 8 + 12 + 8 + 6 + 8) / 6 = 8.00$

Physical Demand: $(2 + 4 + 14 + 2 + 2 + 12) / 6 = 6.00$

Temporal Demand: $(2 + 4 + 10 + 4 + 2 + 10) / 6 = 5.33$

Performance: $(16 + 14 + 14 + 12 + 18 + 16) / 6 = 15.00$

Effort: $(8 + 14 + 14 + 10 + 6 + 8) / 6 = 10.00$

Frustration: $(2 + 2 + 12 + 6 + 2 + 12) / 6 = 6.00$

Compared to our goals, most of the categories scored within the expected range with 1 outlier and 2 notable deviations.

Massive outlier was the performance which was a lot higher than expected. Users felt like they were not that successful in completing the task. This may have been an issue with our testing feedback process rather than the prototype but could also be an issue with our design.

Additionally, Physical Demand and Effort deviated slightly beyond the ± 3 threshold with a difference of 3.2. +0.2 is that a big deal but it's notable that they are the highest values aside from performance. This means that our prototype was more mentally demanding and took more effort. Although this is still very close to what we had expected.

Category	Target	Actual (Avg)	Difference
Mental Demand	7.6	8.00	+0.4
Physical Demand	2.8	6.00	+3.2
Temporal Demand	5.8	5.33	-0.47
Performance	5.2	15.00	+9.8
Effort	6.8	10.00	+3.2
Frustration	5.8	6.00	+0.2

Usability Goal Calculations: Times

Along with measuring TLX scores given by testers from surveys, we additionally employed a more objective measurement which was the time taken for each tester to perform a task. Usability goals are met if the average times taken by testers for each task align with the ranges provided for each task.

Task	Target Time	Average Time	Difference	Goal Meet
Task 1	40s ± 15s (25-55s) ▾	15.82s	-9.18s	Yes ▾
Task 2	20s ± 10s (10-30s) ▾	33.11s (3 users)	+3.11s	No ▾
Task 3	60s ± 30s (30-90s) ▾	16.89s (4 users)	-13.11s	Yes ▾
Task 4	30s ± 15s (15-45s) ▾	12.83s (4 users)	-2.17s	Yes ▾
Task 5	60s ± 30s (30-90s) ▾	17.08s (4 users)	-12.92s	Yes ▾

Results of the actual times recorded vs target times show that all tasks 1,3,4,5 were completed faster than their target time apart from task 2 which was slower. This shows that we mostly meet our goal except for task 2. This shows making a route took marginally longer amount of time to complete than anticipated and could be considered a usability issue.

Usability Test Discussion

What were the most interesting things you found from user testing?

One of the most interesting findings when user testing was how many users struggled with interface elements that we assumed were obvious. For example, participants often mistook the bottom knob on the routes menu for an iPhone swipe gesture. The participants thought it was just there for aesthetic vibes rather than actually having a functionality. When they were told there was actually some functionality to it they tried to drag the knob instead of clicking it. It was also revealing to see how redundant navigation options like the hamburger menu and the bottom navigation bar caused confusion because users hesitated over where to find key features when those features appeared in more than one location. We also observed a wide range of task completion times such as Task 1 ranging from 8.28 seconds to 30 seconds and task 5 ranging from 6.27 to 40 seconds highlighting how differently people approached the same flows and showing why both performance metrics and qualitative feedback are important.

What did you learn from conducting the user testing?

Conducting the tests taught us that no matter how polished a design appears on paper, real users will reveal mismatches between their mental models and our intended workflows. We learned that even simple features such as reading Notices or clicking the Saved Routes button can become stumbling blocks if users are not primed with consistent terminology or immediate functional feedback. In particular, data from the NASA TLX survey showed that physical demand (average 6.00 versus target 2.8) and perceived performance (average 15.00 versus target 5.2) were much worse than expected, so it became clear that our layouts required clear visual hierarchy and more accessible touch targets. Finally, watching users hesitate when they encountered unimplemented features like a non-working start

button made us realize how important it is to have every interactive element functional before the usability testing begins as it would have made for a smoother experience for the participants and users in general.

Another thing that we learnt was that we needed to consider the possibility that incompleteness in the prototype could cause UX problems when completing the tasks. The purpose of the prototype isn't to be fully functional, we aren't at that stage of development, but artefacts in the prototype shouldn't disrupt tasks from being completed. One example of an incomplete part of the prototype causing problems was the first task (finding the next departure of the HVL line) during the first session. Here, the participant completed the task but was unaware of that fact because all times were "00:00" to indicate a placeholder value, and they weren't sure which was the current line. Because of this, we changed the first from finding the next departure to instead just finding the timetable.

Another example of this was experienced in the second session in which the participant in the fourth task (opening a saved route) tried going to Settings → Saved Routes to access the saved routes, which failed because that button wasn't implemented. Most other parts of the settings also were incomplete, but this was especially problematic because that part of the UI was supposed to be related to the task. We could've put a label that stands out explicitly mentioning that it's not implemented, or we could've completely removed it from the prototype. One final example was that in many parts of the UI, such as the routes presented when planning a trip, there were multiple buttons and only one of them worked. This was demonstrated as a problem because of the need to guide the user to click a specific button, rather than letting the user figure things out themselves.

What were the most significant results you found from user testing?

The most significant quantitative result was the gap between our usability goals and actual perceptions on the NASA TLX survey. Performance was nearly three times worse than the target and effort demand was substantially higher for example an average effort of 10.00 versus a target of 6.8. Qualitatively, three issues stood out as major. First, a double click problem for route selection frustrated several users. Second, confusion around the Saved Routes button in settings distracted participants from completing tasks smoothly because of it not functioning. Third, the redundant navigation between the hamburger menu and the bottom bar led to hesitation and wasted time. Each of these issues consistently produced frustration or hesitation indicating that fixing or removing these elements should be a top priority. Lastly, inconsistent timetable data that showed 00:00 in all cells created doubt about whether users had reached the correct screen which is a cosmetics issue that had an outsized negative impact on perceived reliability.

All Issues Classified

“Bottom Knob” on routes menu was unclear (misunderstood as iphone swipe gesture)

Description: To progress task 4, testers had to open a bottom drawer as the last step. However once they got to the map page to do this, they thought the task was finished. All testers were confused when prompted to open this bottom drawer as the UI to open it looked very similar to the iPhone home button and didn't think it was apart of the design.

Classification: Catastrophic

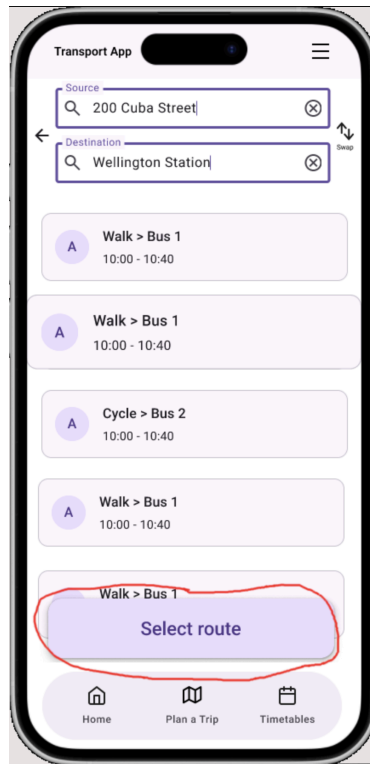
Route selection required double click (bug/unintended)

Description: This bug was reported by the participant in the second session during the second task of planning a trip. This problem occurs when the user is presented with a list of routes from their source to their destination after clicking

the Go button on the map. When the user clicks the second route to select it, they need to double-click the "Select route" button to go to the next stage of seeing their trip on the map.

This problem resulted in the participant being confused why initially clicking the button didn't do anything, and so the remaining participants were guided what to do in this step to complete the task.

Here's a screenshot of this UI with the "Select route" button that needs to be clicked twice circled:



Additionally, this problem doesn't exist when the user doesn't select the second route. When they click "Select route" without this selected, they go immediately to the map. After investigating this in Figma, I found that the problem was an overlay was being opened whenever the second route button was pressed, and this overlay was configured to close whenever the user clicked off it. This meant that the first click on the "Select route" button was "absorbed" by closing the overlay, with the second click only actually showing the trip on the map.

Classification: Major. This issue is major because it is a significant barrier in preventing users from completing their task. This is evidenced by the participant in the second session noting that it made the UI feel unresponsive, and the fact that all participants after were guided what to do.

Unclear on where to find notices (didn't understand concept of notices)

Description: One of the tasks was to get the participant to read a notice about the Hutt Valley line. The user did not understand the concept of notices and where it would be. I do think this is more of a user issue than our issue as it's very obvious to see on the home page. It is also on the Timetables page when you click the Hutt valley Line train service.

Classification: Minor

Settings saved routes button should go to the homepage where the saved routes are.

Description: One of the users was tested to go into settings and make it dark-mode. They were then told to open a saved route in the homepage. From the main settings screen they would have seen the subtext headings which one of

them read as Saved Routes. Which they would have assumed took them to a new screen or something on the lines of that. However they were put to confusion because this Saved Routes button wasn't functionally working and did nothing. Leaving them to ask where to go.

Classification: Major

Hamburger menu and navigation bar are somewhat redundant

Description: Most features are already in the bottom-nav bar. The hamburger menu doesn't provide any more features aside from the settings page. We can remove the search feature entirely as the timetables page already has a search feature. Remove the hamburger menu entirely and replace it with the settings button.

Classification: Minor

Make interface more colorful to distinguish sections in the homepage

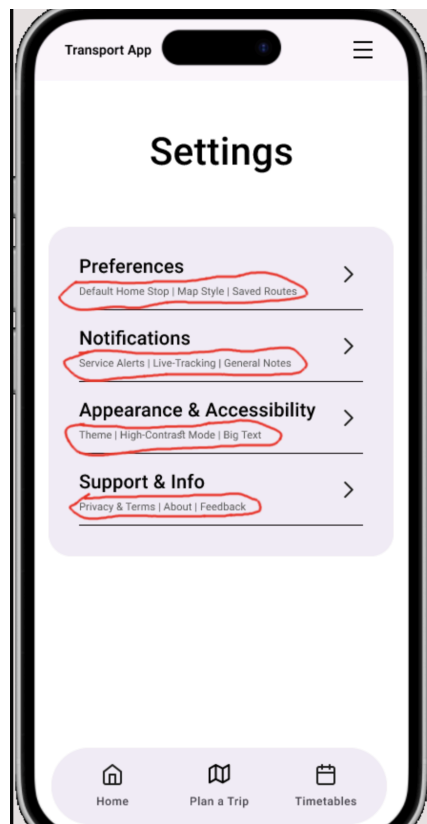
Description: One user found it difficult to find notices on the homepage due to each "section" of the homepage (favorites, saved and relevant notices) not being distinguished enough. This made each section get blurred together and interpreted as one big section rather than 3 smaller section

Classification: Minor

Font size of the small subtext in the settings should be larger

Description: This issue was reported by the participant in the third session when they were completing the fifth task of changing to dark mode. The first part of this task is to reach the settings page from the hamburger menu, which shows four buttons for each category of settings. Under the main label for each button is some smaller text listing each setting in each of these pages, which the participant described as being too small and hard to read.

Here's the image with the text that was described as too small circled:



Looking at the prototype in Figma, the font size of that text is 11. Additionally the colour contrast between the text and background might've contributed to being difficult to read. The hex colours of the text and background are #585858 and #f3edf7 respectively.

Classification: Minor. This isn't a higher classification because even under the assumption that the user cannot read the text at all, they can still manually enter and exit each settings page to search for what they want, which is only a small inconvenience. This also isn't a Severity 1 (Cosmetic/Minimal) issue because this isn't *just* a stylistic issue, users being unable to read that text does slightly affect their ability to perform tasks.

Fix start button which doesn't work in saved routes

Description: The participant was asked to go from saved routes and then pick a suggested route. After this they were then taken to a map with a start button. They clicked it assuming that was going to work but the button wasn't functionally implemented yet. This caused brief confusion until clarified by the team.

Classification: Minor

The HVL timetable should be completed, not every cell should be "00:00" as that was confusing.

Description: When participants opened the timetable page, all cells in the timetable showed "00:00." This made it unclear whether the timetable was functioning correctly or if they were even on the correct page. Several participants had to double-check to confirm they were viewing the timetable at all.

Classification: Cosmetic/Minimal