

Click Through vs. Scroll Through: How Layout Influences Survey Completion

5243 Project 3

Group 5

Julieta Caroppo, Wenbo Liu, Xiaoying Wang, Fatih Uysal

Abstract

This project explores how survey layout formats impact user behavior through the design and deployment of an A/B testing framework centered on a smart watch customization survey. Two distinct versions of the survey were developed: Version A presented all questions and options on a single, scrollable page, enabling free navigation and faster completion, while Version B used a step-by-step structure, displaying one question per screen and requiring users to click "Next" to proceed. To objectively evaluate differences in user engagement, attention span, and completion behavior, we tracked detailed interaction metrics including button clicks, option selections, and page duration using Google Analytics 4 (GA4) and exported data to BigQuery for scalable analysis. The core objective of this study was to determine whether structuring the user experience sequentially versus simultaneously would influence engagement depth and decision-making efficiency, thereby providing practical insights for digital survey and UI/UX design optimization.

1- Introduction & Research Question

The effectiveness of a digital survey is no longer measured solely by response rate, but also by deeper metrics such as completion speed, user engagement, and interaction patterns. In this project, we investigate how the **presentation layout** of a survey influences user behavior, focusing specifically on metrics including click counts, form selections, completion behavior, and time spent within the application.

To study these behavioral effects, we designed and deployed an A/B testing framework using a smartwatch preference survey built with Shiny. Users were randomly assigned to one of two versions:

- **Version A:** A single-page format where all survey components were visible simultaneously.
- **Version B:** A step-by-step format where users navigated through one question at a time.

User interactions were systematically tracked through Google Analytics and stored in BigQuery for analysis. This setup allowed for a direct comparison of user engagement under different survey structures.

Research Question:

Does the survey layout format (single-page vs. step-by-step) significantly affect user behavior, including survey completion time and rate, interaction depth, and the likelihood of submitting preferences?

2 - Experimental Design & Methodology

Research Hypothesis:

The presentation format of a survey (single-page versus step-by-step) significantly affects user behavior, specifically influencing page duration and the likelihood of submitting preferences.

It is expected that the step-by-step format (Version B) will lead to higher submission rates, longer page duration, and greater user engagement compared to the single-page format (Version A).

Survey Design

We developed two distinct versions of a smartwatch preference survey using a Shiny web application:

- **Version A – Single-Page Format:**
All ten survey questions were displayed on a single scrollable page. Users could freely move through the form, answering questions in any order without mandatory navigation steps.
- **Version B – Step-by-Step Format:**
The survey was structured across multiple screens, presenting one question at a time. Users had to click "Next" to proceed to the subsequent question and "Back" to revise previous answers.

Randomization Strategy

Users were randomly assigned to either Version A or Version B upon landing on the website. Randomization was implemented through session-based logic within the application, ensuring approximately equal exposure to both variants without user influence.

Behavioral Metrics (Captured for Analysis)

Although detailed in the Data Collection section, our experiment focused on tracking behavioral outcomes such as:

- Survey completion rate (via "Submit Preferences" button clicks)
- Average time spent on the survey page (page duration)
- Number of interaction events (button clicks and option selections)

3 - App Structure

- Our web application was designed to simulate a smartwatch product preference survey and serve as a platform for A/B testing user behavior. Upon visiting the site, users are randomly assigned to one of two layout variants—Version A or Version B—with an equal probability. Both versions ask the same ten questions related to product configuration and personal preferences, but they differ significantly in layout and interaction style. This design isolates layout as the sole independent variable affecting user behavior.
- Version A displays all questions and product options on a single, vertically scrollable page. Users can freely navigate between sections to select watch features such as color, band, size, extra features, charger, and bundle preferences. The top of the page features a “Subscribe for Updates” button, while the bottom includes a “Submit Preferences” button and a review submission section, where users can leave written feedback along with their name. This layout mirrors a typical e-commerce experience and emphasizes convenience, enabling quicker completion with fewer interactions.
- Version B presents the exact same content, but distributes the ten questions across separate screens using a multi-step form. Each screen includes one question with selectable options, a “Next” button to proceed, and a “← Back” button to revisit previous responses. Unlike Version A, this version only includes a “Subscribe for Updates” button on the landing page and omits the review submission section entirely. This linear structure encourages a more deliberate pace, with additional user effort required to navigate each step.

Element	Version A	Version B
Structure Layout	All 10 questions on a single scrollable page	One question per screen with "Next" navigation
Interaction Flow	Freely scroll and answer in any order	Linear Progression, step by step

Element	Version A	Version B
Structure Layout	All 10 questions on a single scrollable page	One question per screen with "Next" navigation
Interactive Features	“Subscribe for Updates” at top, “Submit Preferences” at bottom	“Subscribe for Updates” on landing page only
Review Section	Includes review form (Name + Review field)	Not included
Expected Completion Time	Quicker due to less friction	Slower due to multiple page transitions
User Experience Goal	Fast, ecommerce-style decision making	Guided, deliberate user input

4 - Data Collection

User behavioral data was collected through the integration of **Google Analytics 4 (GA4)** across both survey versions. Custom event tracking was implemented to capture user interactions at a granular level. The primary tracked events were:

- **page_duration** — Measures the total time (in seconds) from page load until the user exits or minimizes the page (captured via visibility change events).
- **button_click** — Records every button interaction, including crucial actions such as "Submit Preferences" and “Subscribe for Updates”.
- **option_selected** — Captures responses to survey questions, including:
 - Band material selection (Silicone, Leather, Stainless Steel)
 - Watch size selection (Small, Medium, Large)
 - Feature add-ons (e.g., Cellular connectivity, Extended Warranty)
 - Accessory choices (e.g., Charger, Bundle)

- User insights (e.g., phone type, smartwatch usage habits, price sensitivity)
- **color_selected** — Captures user interactions with color swatches representing different watch colors (tracked via custom **gtag** events in both versions).
- **material_selected** — Captures band material selection via swatch clicks in **Version B** only. In **Version A**, band material was tracked under the **option_selected** event as part of the form responses.

All events were automatically exported to **Google BigQuery** using GA4's native linking functionality. Data was refreshed on a daily basis, allowing for scalable SQL-based analysis across both A/B variants. Each event was stored with associated metadata including user pseudo-ID, event timestamp, and relevant custom parameters (e.g., event labels for detailed selection tracking).

The main behavioral metrics derived from the dataset included:

- Event counts per user
- Page engagement and abandonment rates
- Average interaction time per survey step
- Distribution patterns of option selections across users
- Submission rate comparison between Version A and Version B

5 - Statistical Analysis & Results

To evaluate whether the survey presentation format (single_page vs. step-by-step page) significantly influences user behavior, we conducted a series of statistical and behavioral analyses based on the tracked events in Google Analytics. The primary metrics accessed include from submission rate, page duration, interaction behavior, and drop-off patterns.

5.1 Submission rate Analysis

For Version A, we have a total of 62 users participated in and 24 of them submitted the survey. The completion rate for version A is 38.7%.

For Version B, we have a total of 56 users participated in and 29 of them submitted the survey. The completion rate for version B is 51.8%.

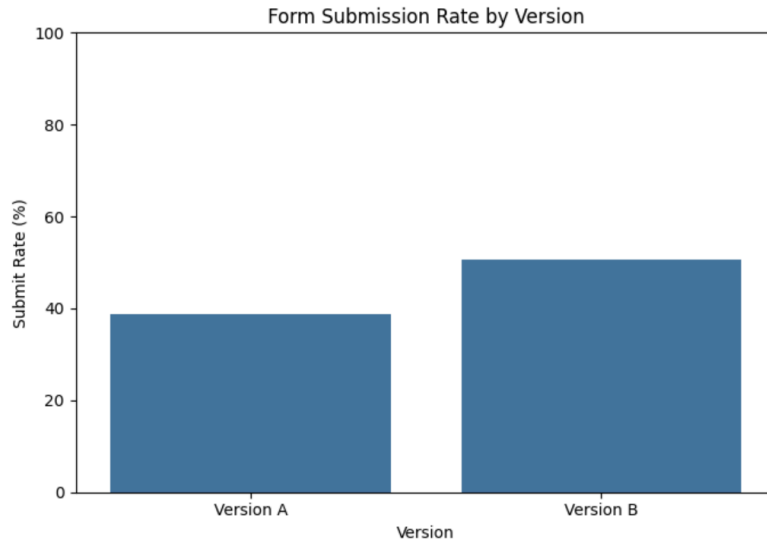


Figure 1: Submission Rate of Version A vs. B

We applied a chi-square test of independence since both the survey layout and the submission status are categorical variables. We want to assess whether there is a significant association between the two variables by comparing the observed frequencies with the expected frequencies under the assumption of independence.

The test produced a chi-square statistic of 2.22 ($p = 0.136$), indicating that there is no statistically significant association between layout format and submission rate at the 5% significance level. However, Version B showed a higher observed submission rate than expected, suggesting a potential trend worth further investigation with a larger sample size.

5.2 Page Duration Analysis

To understand the user engagement across different versions. We calculated the average time users spent interacting with the page.

Version	Avg. Time Spent on Survey (seconds)
A	57.80
B	61.46

Figure 2: Average time Spent on Survey of Version A vs. B

We use `page_duration` metric to collect the total time (in seconds) from the page. We performed a t-test for the page duration and we got a T-test statistic: -0.2373 and P-value: 0.8129. Since p-value is higher than 0.05, the result indicates no statistically significant difference in average time spent between the two survey versions. Although Version B users spent slightly more time on the survey, this difference was not sufficient to support the hypothesis that layout format significantly affects page duration.

5.3 Interaction behavior

We compared the frequency of user selections across various survey options to explore how the layout structure may influence user interaction patterns and exploration behavior. As visualized in the figure, Version A users were more likely to interact with immediately visible options such as color, whereas Version B users explored a wider range of options, including watch features and accessories, likely due to the guided step-by-step format that exposed them to each category individually.

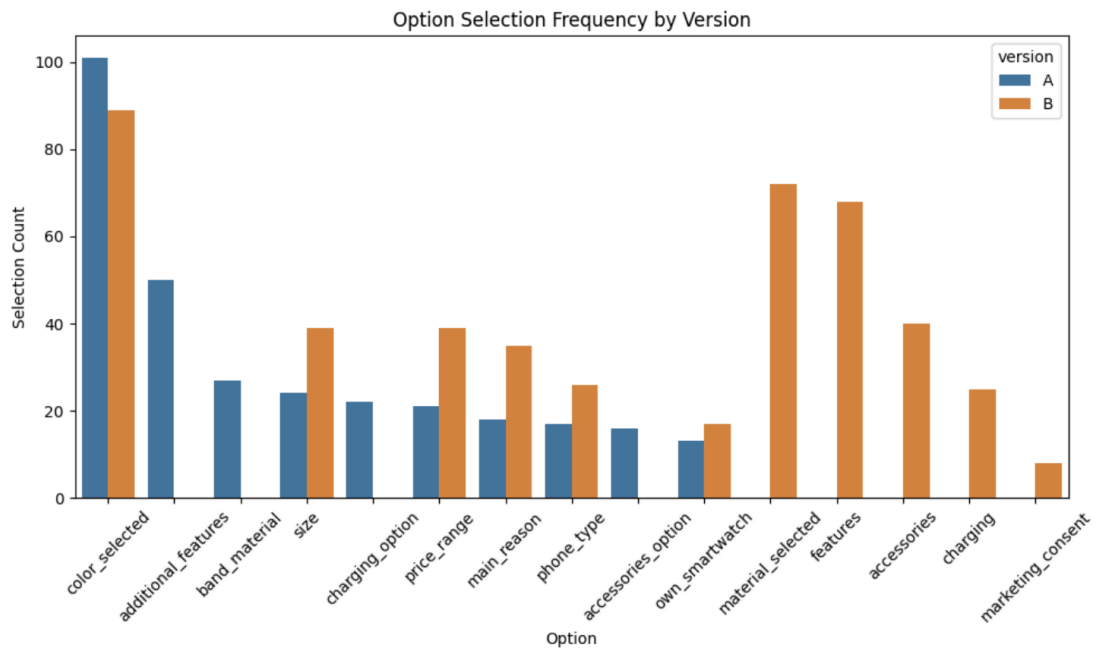


Figure 3: Metric Selections Frequency

To statistically assess whether the distribution of selected options differed significantly between the two layouts, we conducted a Chi-square test of independence. The test yielded a chi-square statistic of 329.07 with 14 degrees of freedom and a p-value near zero ($p < 0.00001$), indicating a statistically significant difference in selection patterns between Version A and Version B. This suggests that the presentation structure strongly influenced which customization categories users engaged with.

5.4 Drop-off Analysis

In addition to analyzing submission behavior and option interaction patterns, we examined the drop-off behavior to see who interacted with the survey interface but did not submit their preferences. The drop-off rate for version A is 25.5% and drop-off rate for version B is 21.66%.

From the data, we figure out that version A exhibited a higher drop-off rate relative to the number of users who interacted with the survey.

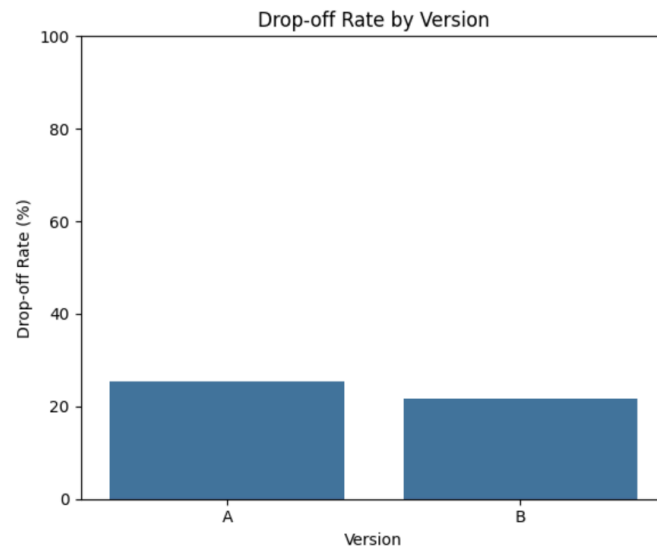


Figure 4: Drop-off Rate of Version A vs. B

The result shows that the version A (single-page layout) may induce decision fatigue, leading to a greater proportion of users abandoning the survey before submission. Conversely, version B (step-by-step layout) appears to have better supported users by navigating them through the survey with a lower relative drop-off rate.

5.5 “Subscribe for Updates” Click Rate Analysis

In an A/B test of two sign-up form designs, Version A recorded 45 “Subscribe for Updates” events out of 62 views (72.6 %), while Version B saw 38 events out of 56 views (67.9 %). Under the null hypothesis of equal true subscription rates, the pooled success rate was 70.3 %, yielding a z-statistic of +0.56 and a two-sided p-value of 0.575. The 95 % confidence interval for the difference in rates (A – B) spans from –21.2 % to +11.8 %. Because the p-value exceeds 0.05 and the interval includes zero, we conclude there is no statistically significant difference between the two designs. Although Version A’s raw rate is slightly higher, this result suggests that further design iterations or a larger sample size would be needed before confidently favoring one form layout over the other.

6 - Interpretation & Conclusion

- The result of our statistical analysis provides how survey presentation format impacts user behavior. Although version B demonstrated higher submission rates, longer page durations, and broader interaction across survey options compared to version A, the differences are not always statistically significant.
- For submission rate, while version B has a higher observed completion rate(51.8% vs. 38.7%), the chi-square test indicated no statistically significant differences between layout and submission.
- For page duration, although people spent more time on version B, we got a p-value of 0.8129 which means that the difference was not statistically significant.
- In terms of interaction behavior, there was a statistically significant difference in option selection patterns between Version A and Version B. This suggests that the step-by-step layout meaningfully influenced user engagement with the survey content.
- For drop-off rate, version A exhibited a higher drop-off rate, implying that the single-page layout in version A might cause decision fatigue.
- For “Subscribe for Updates” click rate, although Version A’s raw rate is slightly higher, we conclude there is no statistically significant difference between the two designs.
- Overall, the behavioral data suggest that while the layout change encouraged broader exploration and possibly improved user experience, it did not significantly impact core engagement metrics such as submission rates or time spent within the current sample.
- In conclusion, these findings partially support our original hypothesis that presentation format significantly affects user behavior. Moreover, we observed version B demonstrated higher submission rates, longer page durations, and broader interaction across survey options compared to version A, which aligns with our initial hypothesis. However, while the layout structure clearly influenced how users interacted with survey options, it did not significantly impact overall survey completion rates or the amount of time users spent completing the survey.
- Obviously, future research with larger sample sizes and more diverse participant demographics is needed to better confirm these trends and further explore how layout design can optimize user engagement and survey effectiveness.

7 - Further Discussion

- Our A/B Testing demonstrates how survey interface design meaningfully influences user interaction patterns, decision-making behavior, and task completion likelihood. By comparing a single-page layout (Version A) with a step-by-step layout (Version B), we observed distinct trade-offs between exploratory engagement and guided task progression.
- The step-by-step design (Version B) prompted significantly more interactions with more button click events. While this structure introduced more interface friction on the surface, it did not lead to longer completion times. In fact, users in Version B spent slightly less time on average to complete the survey, suggesting that the segmented flow allowed users to maintain focus and avoid cognitive overload. This supports prior usability research indicating that breaking down complex tasks into smaller steps can improve user flow and reduce abandonment.
- Moreover, users in Version B also made more average option selections, indicating sustained attention and potentially deeper engagement with each decision point.
- The single-page layout (Version A) encouraged a higher rate of optional interaction, particularly in the area of visual customization. Users were more likely to interact with color swatches, possibly due to the simultaneous visibility of all customization elements. This suggests that such a layout supports freeform exploration, which could be beneficial in contexts prioritizing browsing, discovery, or creative input. Since our version B introduced a new tracked event `material_selected`, which was only triggered in this layout due to the use of interactive swatches. The recorded 40 interactions across 26 users highlights that users were not discouraged by the additional step. Rather, when customization features were clearly supported within the structured flow, they remained effective and engaging.

8 - Challenges & Limitations

Several challenges were encountered throughout the development, deployment, and data collection processes:

- **Deployment Issues:**
During application deployment to Shinyapps.io, conflicts in task dispatch occasionally caused delays. In particular, concurrent deployment tasks led to temporary errors requiring manual redeployment efforts.
- **Event Tracking Debugging:**
Google Analytics 4 introduces a processing delay between event capture and dashboard visibility, complicating real-time debugging and validation of custom events. This latency

made it difficult to immediately verify the correctness of new tracking implementations.

- **GA4 Limitations:**

GA4 does not retroactively populate newly defined custom dimensions or metrics for historical data. Consequently, tracking adjustments made mid-experiment could only apply to newly incoming data, limiting the ability to perform unified analysis across the full dataset.

- **Data Sampling Risks:**

Although BigQuery exports mitigated much of the sampling risk inherent in GA4 reporting interfaces, data delays and schema updates still required careful management to ensure consistency across daily tables.

- **Outlier Problems:**

We faced an outlier problem when collecting for To account for outliers since we have the average duration 2656.43 second which is not reasonable. The problem may be due to people forgetting to close the page. We applied the interquartile range (IQR) method to filter extreme values. The cleaned duration values were log-transformed to address right-skewed distribution before conducting the t-test.

9 - Member's Contribution and GitHub Repository

- Julieta was responsible for the technical development and deployment of the web application used in this A/B testing project. Based on the survey design provided by another teammate, she built both Version A (single-page layout) and Version B (multi-step layout), ensuring that both variants were integrated under a single randomized-access link. She implemented the user flow logic, visual styling, and layout distinctions between the two formats. In addition, Julieta conducted extensive debugging to guarantee smooth functionality and consistent data tracking across both versions. She also managed the successful deployment of the application on Shinyapps.io, enabling live user participation and analytics monitoring.
- Fatih was responsible for creating the initial website, which served as the base version for testing. He later updated it to Version B for comparison purposes. He also contributed to idea generation and helped set up the initial concept for the A/B testing. Additionally, he created different surveys tailored to the different website structures. He also did statistical analysis on some of the metrics. He also contributed to the final report.
- Xiaoying Wang is responsible for the data collection, doing the statistical analysis and conducting the result analysis from Google Analytics.

- Wenbo was responsible for integrating the website with Google Analytics (GA4), enabling comprehensive event tracking such as button clicks, option selections, page engagement time, and material/color selections. Wenbo also connected GA4 with BigQuery, conducted user-level data extraction and cleaning, standardized event formats, and prepared datasets for behavioral analysis. In addition, Wenbo deployed the updated version of the Shiny App to support event tracking and data collection improvements. Beyond technical contributions, Wenbo independently completed the writing of the Introduction, Experimental Design & Methodology, Data Collection, and Challenges & Limitations sections of the report. Wenbo also formulated experimental hypotheses, defined key tracking events, designed corresponding A/B tests, and proposed appropriate analytical methods. Furthermore, Wenbo optimized the overall report structure and formatting to enhance clarity, logical flow, and readability, ensuring a professional and cohesive final deliverable.

Team:

Julieta Caroppo (jc6158)

Fatih Emre Uysal (fu2137)

Xiaoying Wang (xw2993)

Wenbo Liu (wl2939)

GitHub Repository: https://github.com/fatihEuysal/5243Team5_Project3

Application Link: <https://5243project3.shinyapps.io/myapp/>

10 - Appendix (with Figures on GA)

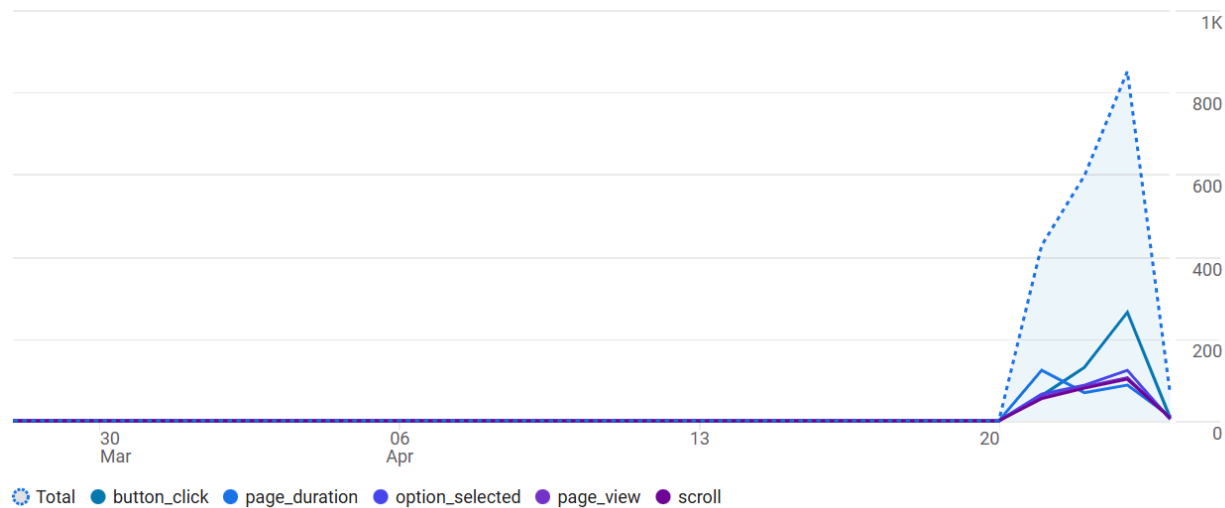


Figure 1: Version A's Daily Trend of Metrics

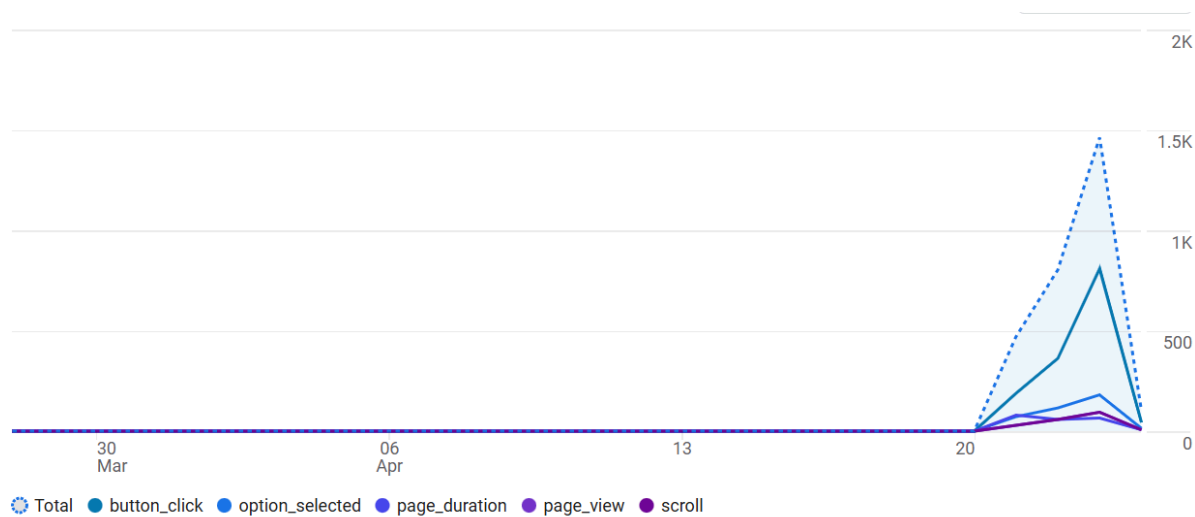


Figure 2: Version B's Daily Trend of Metrics