Exploring the Importance of Notification Design on User Engagement

Hans Matthew Abello, Maxine Beatriz Badiola, Mark John Custer, Lorane Bernadeth Fausto, Patrick Josh Leonida, Denzel Bryan Yongco and Jordan Aiko Deja

1De La Salle University, Manila, Philippines

Abstract

Push notifications are brief messages that users frequently encounter in their daily lives. However, the volume of notifications can lead to information overload, making it challenging for users to engage effectively. This study investigates how notification behavior and color influence user interaction and perception. To explore this, we developed an app prototype that tracks user interactions with notifications, categorizing them as accepted, dismissed, or ignored. After each interaction, users were asked to complete a survey regarding their perception of the notifications. The study focused on how different notification colors might affect the likelihood of acceptance and perceived importance. The results reveal that certain colors were more likely to be accepted and were perceived as more important compared to others, suggesting that both color and behavior play significant roles in shaping user engagement with notifications.

Keywords

design, notification, nudges, design patterns

1. Introduction and Background

Push notifications have become an integral part of modern communication, reshaping how we receive messages across various contexts, including social, work, and emergency notifications. These notifications deliver concise, relevant information in a short amount of time, enabling users to quickly assess whether they should engage further or ignore them.

Apple has introduced Focus Mode to experiment with notification behavior, building on its "Do Not Disturb Mode" [1]. Focus Mode allows users to tailor notification delivery by limiting the types of notifications they receive and assigning them varying levels of priority. In contrast, Android users benefit from more customization options, including a snooze feature that lets users temporarily silence notifications while ensuring they do not miss important messages [2].

In today's digital landscape, notifications are essential for keeping users connected to both their virtual and real-world obligations. However, with the increasing volume of notifications,

managing them efficiently has become a critical design challenge. A study by Pielot et al. 2014 found that users receive an average of 63.5 notifications per day, and many report feelings of stress or overwhelm due to the constant influx of alerts.

To address this challenge, streamlining push notifications for easier assessment, readability, and accessibility is crucial. Given the customization options available, it is important to systematically understand how users interact with notifications, considering factors such as content and the time it takes to respond. Additionally, this study investigates how notification color influences the likelihood of user interaction or dismissal. Previous research has shown that colors can significantly affect cognitive, psychological, and emotional responses, influencing user behavior [4], [5], [6], [7]. In this study, notifications are characterized by two variables: color and category. Three colors and three categories were tested, each with its own heading and body text, as detailed in Table 1.

Table 1
Category descriptions

Category	Header	Color	Body
Emergency	Earthquake Alert: Seek Shelter	red	`NDRRMC: (1:25PM,13Mar24) Isang Magnitude 7.3 na lindol ang naganap kaninang 1:05PM. Aftershocks ay inaasahan.`
Social	Facebook Messenger	green	`John Doe sent you a friend request.`
Work	Canvas Student	blue	`New Assignment: {[}Problem Set{]} Cache Memory Analysis: {[}1232_CSARCH2_S11{]}`

Notifications also allow for user interaction, with three primary actions: accepting, dismissing, or ignoring. **Accept** indicates that the user intends to view the content of the notification. **Dismiss** means the user has decided not to engage with the notification content. **Ignore** is used when a notification is left unattended for at least 10 seconds, marking it as ignored. Additionally, users can configure notification behaviors, which can influence the perceived urgency of the notifications, as outlined in Table 2.

Table 2
Notification behaviors descriptions

Behavior	Description
Urgent	Notifications that were ignored would disappear.
Non-urgent	Notifications that were ignored would not disappear, allowing the user to still interact with them. These notifications can stack with one another.

Interaction time, denoted as *IT*, refers to the duration it takes for a user to interact with a notification, measured in seconds. This is calculated by subtracting the time the notification was

shown (T_s) from the time the notification was interacted with (T_r), as expressed in the following equation:

$$IT = T_r - T_s \qquad\qquad (1)$$

Where T_r is the time when the notification is interacted with, and T_s is the time when the notification first appears.

The primary contribution of this paper lies in exploring how notification colors and their format (e.g., content and urgency) influence user engagement. By examining how different color schemes and notification characteristics affect user interaction time and behavior, this study aims to provide insights into optimizing notifications for better user experience and engagement.

2. Method

2.1. Experimental design

An online prototype was developed in the form of a game inspired by "Simon Says," featuring a wheel with four colors. The game begins with a randomly flashing color on the wheel, and the user's objective is to remember and reproduce the randomly generated sequence of colors. Each time the user correctly selects the right color, additional colors are added to the sequence, thereby increasing the user's score. The game ends once the user reaches a score of 8. However, if the user makes an error, the sequence restarts. During gameplay, notifications were displayed at random score intervals in the upper-right corner of the screen.

Ten university students from De La Salle University - Manila participated in this experiment. Each participant completed the experiment twice, once with urgent notifications and once with non-urgent notifications. The experiment began once the user selected a notification behavior, and it concluded when they had viewed all possible notification combinations. The results were automatically generated at the end of the experiment.

2.2. Post-experiment survey design

Upon completing the experiment, participants were asked to fill out a post-experiment survey. The survey included questions with a 5-point Likert scale, categorical questions, and open-ended questions for qualitative responses. The objective of the survey was to gather participants' perceptions of the importance of each notification, their color preferences for the notifications, and their overall feedback regarding the experiment.

3. Results

3.1. Application results

A total of 180 data points were collected for each behavior in the experiment. The rows where the status was labeled as 'ignored' were excluded from the dataset, as these observations did not

provide interaction times. This resulted in 136 entries for the urgent behavior and 144 entries for the non-urgent behavior.

3.1.1. How does interaction time vary across different colors and statuses in each dataset?

Table 3
Descriptive statistics for color in both behaviors

behavior	color	count	mean	std	min	max
Urgent	Blue	46.0	2.310	1.679	0.704	8.280
Urgent	Green	43.0	1.847	1.281	0.821	5.924
Urgent	Red	47.0	2.225	1.768	0.689	6.741
Non-urgent	Blue	48.0	2.780	3.70	0.694	18.692
Non-urgent	Green	48.0	6.315	13.579	0.641	67.916
Non-urgent	Red	48.0	3.357	8.152	0.754	53.551

For urgent notifications, as shown in Table 3, the color *blue* has a mean interaction time of 2.310 seconds, with times ranging from 0.704s to 8.280s. The color *green* has a mean time of 1.847s, with a range from 0.821s to 5.924s. The color *red* has a mean time of 2.225s, with a range from 0.689s to 6.741s. Among these colors, *blue* has the highest mean interaction time, while *green* has the lowest.

In contrast, for non-urgent notifications, as shown in Table 3, the color *blue* has a mean time of 2.781s, with a range from 0.694s to 18.692s. The color *green* has a mean time of 6.315s, ranging from 0.641s to 67.916s. The color *red* has a mean time of 2.225s, with a range from 0.689s to 6.741s. Among these colors, *green* has the highest mean interaction time, while *blue* has the lowest.

3.1.2. How does interaction time vary across different categories and statuses in each dataset?

Table 4
Descriptive statistics for category in urgent behavior

behavior	category	count	mean	std	min	max
Urgent	Emergency	44.0	2.146	1.543	0.689	6.741
Urgent	Social	47.0	2.263	1.821	0.747	8.280
Urgent	Work	45.0	1.987	1.414	0.694	6.582
Non-urgent	Emergency	48.0	4.778	12.179	0.641	67.916
Non-urgent	Social	48.0	2.756	5.387	0.787	33.455
Non-urgent	Work	48.0	4.920	9.569	0.694	53.551

For urgent notifications, as shown in Table 4, the *Emergency* category had interaction times ranging from 0.689s to 6.74s, with a mean time of 2.146s. The *Social* category had times ranging from 0.747s to 8.280s, with a mean time of 2.263s. The *Work* category had times ranging from

0.694s to 6.582s, with a mean time of 1.987s. Among these categories, *Work* had the smallest mean interaction time, while *Social* had the largest.

In contrast, for non-urgent notifications, as shown in Table 4, the *Emergency* category had a mean interaction time of 4.778s, with times ranging from 0.641s to 67.916s. The *Social* category had a mean time of 2.755s, with times ranging from 0.787s to 33.455s. The *Work* category had times ranging from 0.694s to 53.551s, with a mean time of 4.919s. Contrary to the urgent dataset, the *Social* category had the smallest mean interaction time for non-urgent behavior, while *Work* had the largest.

3.1.3. Is there any difference in the relationship between color and status between the two datasets?
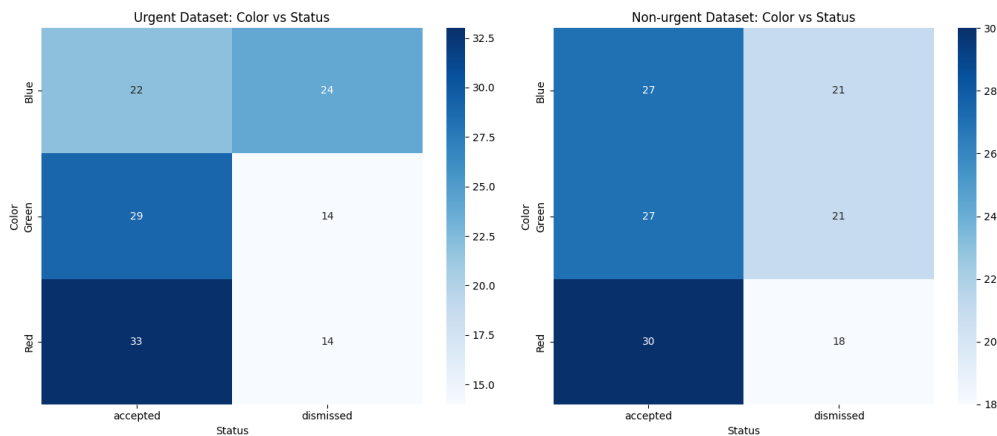

Figure 1: Color vs Status Heatmap

The heatmaps represented in Figure 2 show the number of occurrences of specific color-status combinations within the urgent dataset. The color intensity in each cell reflects the count of observations: darker shades indicate higher counts, while lighter shades represent lower counts.

In the urgent dataset, for the color Red, there are 33 accepted interactions, 14 dismissed interactions, and 3 ignored interactions. These are represented by darker shades in the respective cells. For the color Green, 29 interactions were accepted, 14 were dismissed, and 3 were ignored. Similarly, for the color Blue, there were 22 accepted interactions, 24 dismissed, and 3 ignored.

In the non-urgent dataset, for the color Red, there are 30 accepted interactions, 18 dismissed interactions, and 3 ignored interactions. For the color Green, 27 interactions were accepted, 21 were dismissed, and 3 were ignored. Finally, for the color Blue, 27 interactions were accepted, 21 were dismissed, and 3 were ignored.

Overall, the behavior shown by participants for the color Red was consistent across both scenarios, with Red having the highest number of accepted interactions in the dataset. Additionally, it is notable that there were generally more dismissed interactions in the non-urgent dataset, which is expected, as the level of urgency is lower compared to the urgent dataset.

3.1.4. Are there any differences in the average interaction times between the urgent and non-urgent datasets?

The mean times for urgent and non-urgent behaviors are **2.134s** and **4.151s**, respectively. Given the mean times for the two behaviors in Table **??**, the t-test statistic was used to determine the significance of the difference in both means. With $\alpha = 0.05$, the computed t-value and p-value is -2.455 and 0.014, respectively. This states that there is a significant difference between the two means shown in Figure **??**.

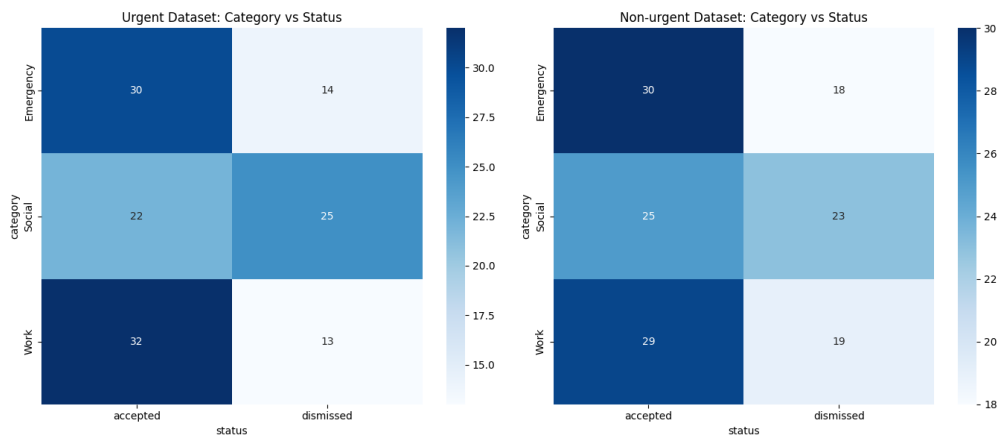### 3.1.5. Is there any correlation between category and status in each dataset?

Figure 2: Category vs Status Heatmap

The heatmaps display the contingency tables for both datasets, with rows representing categories (Emergency, Work, Social) and columns representing statuses (Accepted, Dismissed, Ignored). A chi-square test was applied to the contingency tables of both datasets to assess whether the two variables are independent of each other.

For the urgent dataset, the calculated p-value is approximately 0.03. With $\alpha = 0.05$ and a chi-square statistic of $\chi^2 = 6.883$, this indicates that the variables are dependent on each other. In contrast, for the non-urgent dataset, the chi-square test yields a p-value of 0.548 and a chi-square statistic of $\chi^2 = 1.200$, suggesting that the variables are independent.

3.1.6. Do notification interactions differ between urgent and non-urgent contexts?

After conducting our exploratory data analysis, we observed significant differences in interaction behavior between the urgent and non-urgent datasets.

The visual graphs display similar, if not identical, trends, suggesting that users interact differently depending on the level of urgency of a given notification. However, further analysis revealed that factors such as category, status, and interaction time do not vary significantly and do not exhibit a strong correlation in non-urgent notifications. In contrast, there is a significant correlation between category and status in urgent notifications.

3.2. Survey results

3.2.1. Perceived importance of color based on category

After the tests, the participants were surveyed to assess their perceived importance of the different categories and colors. The first part of the survey used a 5-point Likert scale to measure the perceived importance of each category and color, as shown in Table 5.

For both the Work and Emergency categories, participants indicated that the color Red conveyed a higher sense of importance for these categories. This finding is further supported by the follow-up question, where participants identified their most preferred color for each category, as illustrated in Figure 3.

Table 5
5-point Likert scale of perceived importance

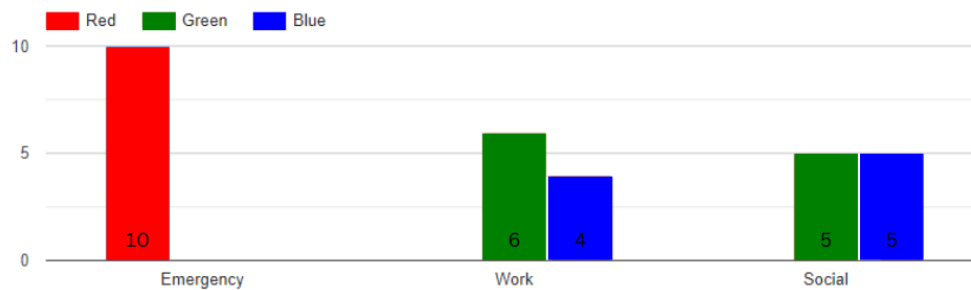| Category : Color	1	2	3	4	5
Emergency : Red	0	0	1	0	9
Emergency : Green	2	2	3	0	3
Emergency : Blue	1	2	3	2	2
Work : Red	1	0	1	0	8
Work : Green	1	0	2	3	4
Work : Blue	1	4	2	1	2
Social : Red	2	1	2	2	3
Social : Green	2	1	2	2	3
Social : Blue	1	3	4	0	2

Figure 3: Most preferred color per category

4. Design implications and limitations

The conducted experiment provided valuable insights for improving notification system design. Color coding emerged as a key factor in enhancing user interaction, enabling users to quickly assess a notification's importance without needing to read it in detail. This is particularly beneficial for urgent alerts or during tasks requiring deep concentration. However, individual
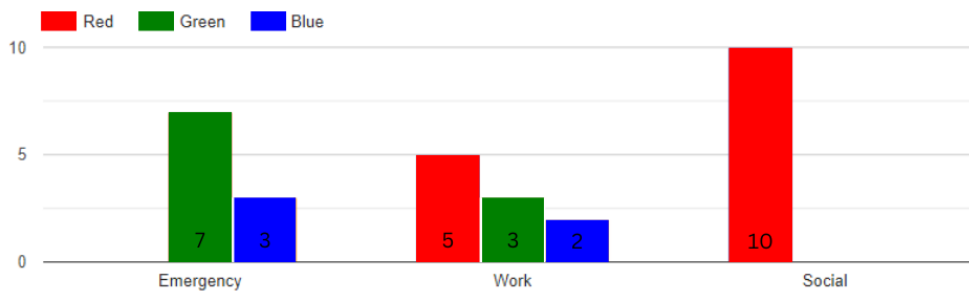
Figure 4: Least preferred color per category

differences in color perception and associations may influence its effectiveness. Therefore, user customization of color schemes is recommended to personalize the experience and improve efficiency. Additionally, integrating AI to adapt notification delivery based on user behavior could optimize timing and presentation, reducing disruptions while ensuring that essential information is delivered promptly.

Despite these insights, the study had several limitations. The sample consisted of a small, homogeneous group of De La Salle University students, which limits the generalizability of the findings. A more diverse participant pool is needed to better understand how different demographics interact with notifications. The experiment's focus on a narrow set of notification categories and reliance on visual cues alone does not fully address the complexities of real-world notification management, which also involves varying types, priorities, and auditory or haptic feedback—elements particularly important for users with visual impairments. Furthermore, the controlled experimental setup may not accurately reflect real-life notification interactions and overlooks how users may adapt to notification cues over time, potentially affecting the long-term impact of design strategies on user attention and engagement.

5. Conclusion and Recommendations

In this experiment, we aimed to investigate how factors such as notification behavior and color influence user interaction and perception. Our goal was to determine whether these factors significantly impact how users interact with different notifications and if they should be incorporated into existing notification customization systems, such as the Focus Mode on Android and iOS.

The results of the experiment show that red notifications were more likely to be accepted than other colors, and notifications with non-urgent behavior were dismissed more frequently. Additionally, notifications with urgent behavior had faster average reaction times. Survey results revealed that all participants preferred the color red for emergency notifications, although it was less favored in other categories. Green and blue were equally preferred for work and social notifications, with the least preferred color for work notifications showing the most variability. Interestingly, we observed similar preferences for green and blue in both the experiment and the survey.

Based on these findings, we recommend incorporating the variables explored in this study as additional customization options for notifications, integrating them with existing features like Focus Mode. This would allow users to tailor notification behaviors and colors to their personal preferences, enhancing the overall notification experience.

References

[1] J. Evans, Apple's focus: What does it do and how does it work?, 2021. URL: https://www.computerworld.com/article/1614283/apples-focus-what-does-it-do-and-how-does-it-work.html.

[2] J. Gilbert, How to snooze notifications on your android phone, 2024. URL: https://www.androidpolice.com/snooze-android-notifications-how-to/, section: Operating Systems.

[3] M. Pielot, K. Church, R. de Oliveira, An in-situ study of mobile phone notifications, in: Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services, MobileHCI '14, Association for Computing Machinery, New York, NY, USA, 2014, p. 233–242. URL: https://doi.org/10.1145/2628363.2628364. doi:10.1145/2628363.2628364.

[4] A. J. Elliot, M. A. Maier, Color and psychological functioning 16 (2007) 250–254. URL: http://journals.sagepub.com/doi/10.1111/j.1467-8721.2007.00514.x. doi:10.1111/j.1467-8721.2007.00514.x.

[5] Achint Kaur, MCM DAV college for women sector 36, Panjab University, Chandigarh, India, A link between colors and emotions; a study of undergraduate females V9 (2020) IJERTV9IS090319. URL: https://www.ijert.org/a-link-between-colors-and-emotions-a-study-of-undergraduate-females. doi:10.17577/IJERTV9IS090319.

[6] P. Valdez, A. Mehrabian, Effects of color on emotions. 123 (1994) 394–409. URL: https://doi.apa.org/doi/10.1037/0096-3445.123.4.394. doi:10.1037/0096-3445.123.4.394.

[7] M. Silic, Understanding colour impact on warning messages: Evidence from us and india, in: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '16, Association for Computing Machinery, New York, NY, USA, 2016, p. 2954–2960. URL: https://doi.org/10.1145/2851581.2892276. doi:10.1145/2851581.2892276.