HCI Evaluation Report for Quizzzz! Application Design Draft

Jocelyn Woods, Joris Belder, Lucian Toşa, Okan Demir Baykal, Xingyu Han, Yoan Naydenov

March 2022

Introduction

The 'Quizzzz' program draft design is the subject of this human-computer interaction evaluation. It consists of the splash, waiting room and game screens and the halftime and final leaderboard. The transitions are expected to be a sliding animation.

In order to evaluate our design, we recruited another group as evaluation experts. For better reflection and data collection, a form, including questions about the different HCI evaluation aspects, has been provided to the group. We will find out what should be revised in the current design from this (Baykal et al., 2022).

Methods

Experts

We recruited 6 experts, with different cultures but similar educational backgrounds. According to the study of Heuristic Evaluation by Jakob Nielsen, this number sits in the range that makes for an optimal "Ratio of Benefits to Costs" (Nielsen, 1995). The group consists of men between the ages of 18 and 19, all computer science undergraduates.

Procedure

In order to get valuable feedback from our experts it is imperative that we provide detailed instructions on what we wish them to do. Firstly, they reviewed the design of our prototype with the 10 usability heuristics in mind (Nielsen, 1994). Next, they should each fill in a google form containing heuristics questions individually and submit it for evaluation. The form had a few main types of questions such as rating handling of heuristics from 1-5, asking for direct feedback on what could be improved, and presenting users with a specific screen and questioning them based on this. The combination of all this data should help us to discern our biggest heuristic flaws. We would like to mention that some of the problems the experts noticed can be influenced by the fact that the images presented are significantly smaller in the form that they would be if the game was played full-screen.

Measures (Data collection)

This heuristics report aims to measure design user-friendliness. We believe that a Google Forms questionnaire targeted at evaluating each heuristic will provide an understanding of areas requiring improvement. Most questions simulate actual game stages, and ask our experts to evaluate based on possible issues. If their intuitive answer differs from the expected result we can reevaluate how user friendly that screen or function is. We also asked them to rank the quality of each heuristic. We will then evaluate the form's results to assess if the evaluators understand our design and how well they believe we achieved our goals. We will also make use of impact-frequency matrices to assess the severity of the problems pointed out by our evaluators (Kovačević et al., 2019).

Results

Visibility of System Status

The overall rating the evaluators gave us for the visibility of system status heuristic was 4.2 which does leave room for improvement (Baykal et al., 2022). We did see that given a screen all of our evaluators were able to accurately surmise what screen they were on, additionally some added what track they believed had been taken to get there. We also asked them if they could list any specific issues and multiple members suggested that we reevaluate the random positions of the names and disordered appearance of emojis (Baykal et al., 2022). Following these responses, we will attempt to create a new design with stronger consistency in both these regards in hope of a more user friendly interface.
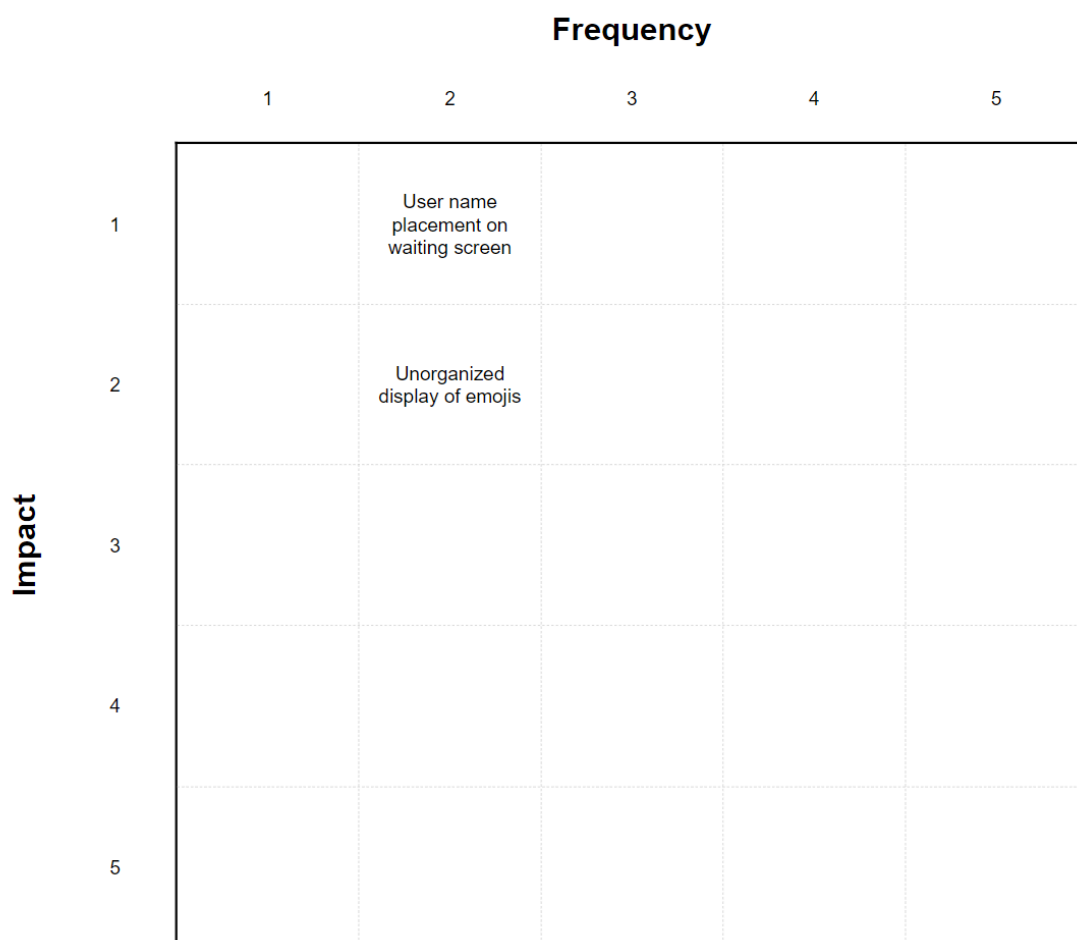
Frequency

	1	2	3	4	5
1		User name placement on waiting screen			
2		Unorganized display of emojis			
3					
4					
5					

Impact (vertical axis label)

Figure 1. Frequency-impact plot for issues related to the "visibility of system status" heuristic.

Match between system and the real world

For this heuristic the rating we received from our experts was a 4.8 (Baykal et al., 2022) . This is a rather positive response to our questionnaire, but of course we can still find room to improve. Looking at our direct feedback the only words our experts found to be unclear were due to the colouring of our slides. On some pages the high contrast between the background colours caused some readability issues (Baykal et al., 2022).
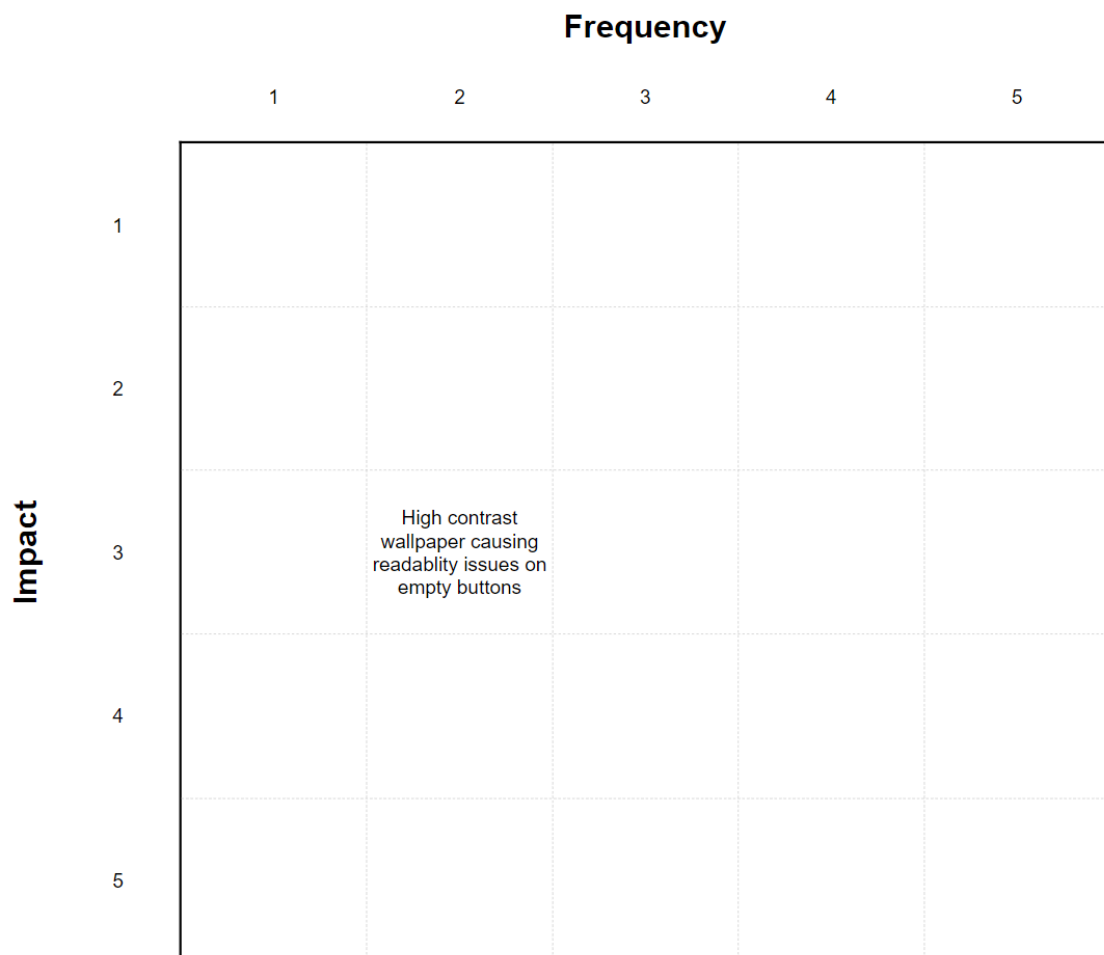
Frequency

	1	2	3	4	5
1					
2					
3		High contrast wallpaper causing readablity issues on empty buttons			
4					
5					

Impact

Figure 2. Frequency-impact plot for issues related to the "match between system and the real world" heuristic.

Flexibility and efficiency of use

The design was evaluated on its flexibility and efficiency of use with a score of 4,4 out of 5 (Baykal et al., 2022). This suggests our design is easy to use. According to the results, the design offers enough flexibility. The only issue found is that a player can't change his nickname while waiting in the room for others (Baykal et al., 2022). We qualify this as a low severity problem, because it doesn't have a significant impact on the game and a player can return to the main menu to change name.
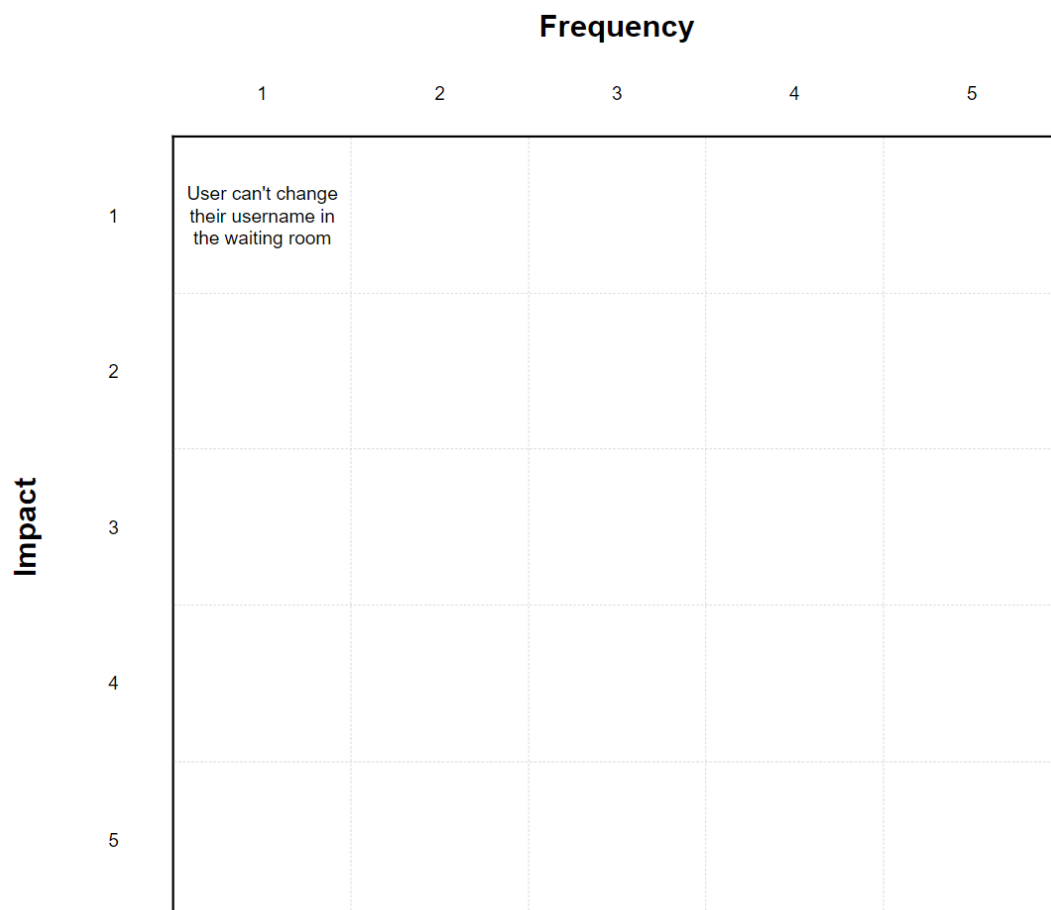
Frequency

	1	2	3	4	5
1	User can't change their username in the waiting room				
2					
3					
4					
5					

Impact

Figure 3. Frequency-impact plot for issues related to the "flexibility and efficiency of use" heuristic.

Aesthetic and minimalist design

For this heuristic the overall score of the design is 4,2/5. Despite the average result, it was stated that the background makes people nauseous or is distracting from the questions and making them harder to read (Baykal et al., 2022). We categorise this problem as very high impact, but low frequency as perhaps not everyone will notice this. Another issue about the design happens when a large number of emojis are used at the same time, because they cover parts of the answers and players can't focus on answering the question (Baykal et al., 2022).
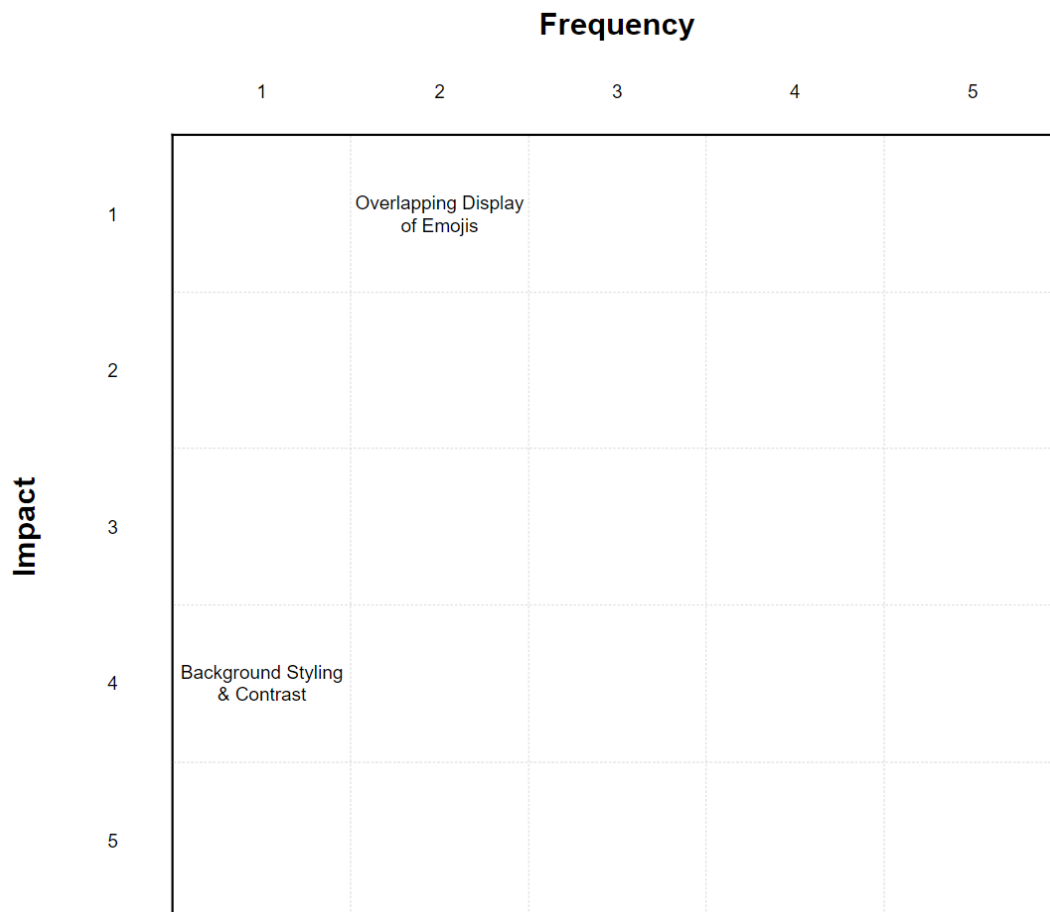
Frequency

Figure 4. Frequency-impact plot for issues related to the aesthetic and minimalist design heuristic.

Error prevention

For this heuristic, we asked the experts whether there were any error-prone entities in the design (Baykal et al., 2022). Two were noted by the experts. Firstly, the jokers and exit button were too close together, causing a player to press exit instead of playing the joker (Nielsen, 1994). We have assigned an impact of 5 and a frequency of 3 for this problem. As for the other error, if there are too many players in the waiting room the names were expected to overlap. This is something that was kept in mind during development, and the name positions are generated such that this would not be the case. We also asked the experts whether there were any places in the design where an undo button would be helpful (Baykal et al., 2022). The experts suggested two places where this functionality would be appreciated. The first one was the ability to change your answer. This function was planned to be implemented, but not demonstrated in the given design. The other suggestion was the ability to change your name in the waiting room[5]. This functionality was missing as the usernames were shared between singleplayer and multiplayer modes, but is now considered. Just like the other heuristics we also asked the experts to grade our error prevention on a scale of 1 to 5. The average rating of the experts came out to a 4.
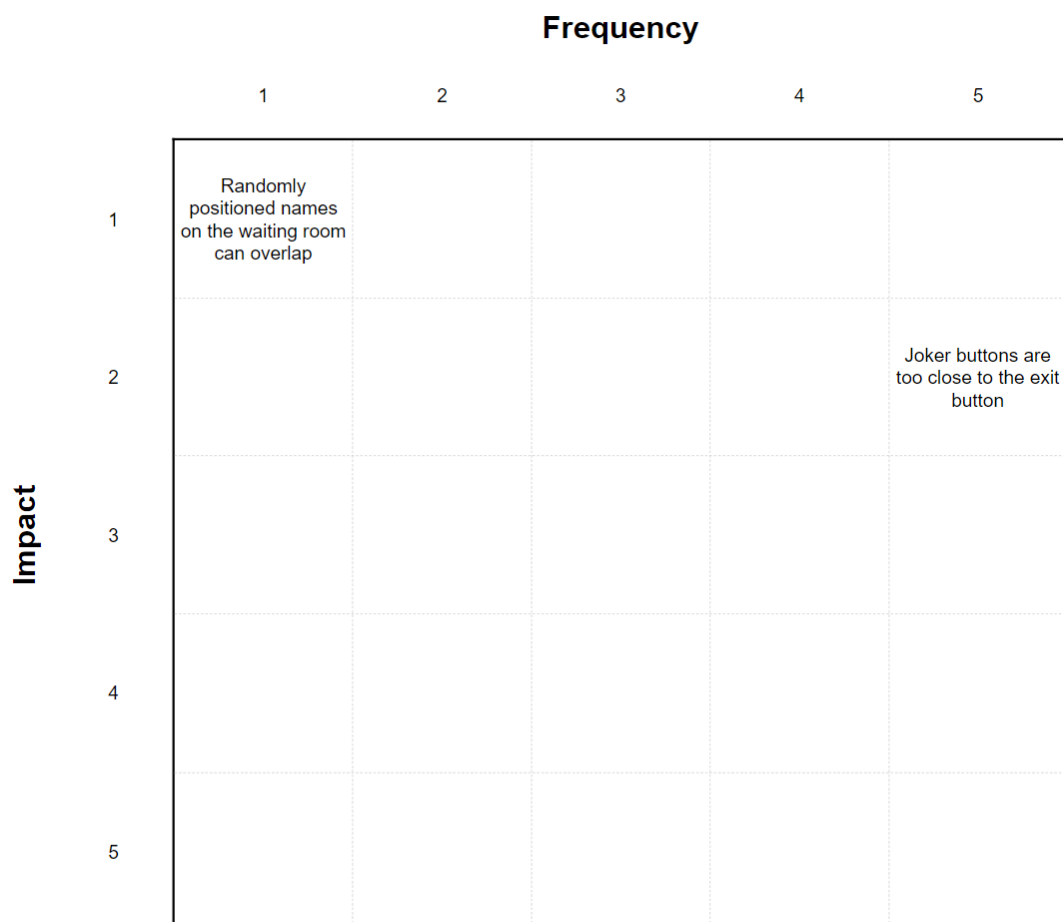

Figure 5. Frequency-impact plot for issues related to the "error prevention" heuristic.

Recognition rather than recall

For this heuristic, the experts were asked two questions (Baykal et al., 2022). The first question was whether there was any information that the experts needed to remember across views, which got all negative responses. Nextly, they were asked whether the help options were in the right views, which got only one negative response. The evaluator that did not agree, could not find any help options in the game screen, where there are indeed no help options. We believe that the answer of the aforementioned evaluator can be disregarded as other experts could easily find the buttons and this evaluator only mentions the game screen. The average ranking for this heuristic was 4.8 out of 5 (Baykal et al., 2022). This is extremely good and implies that there is no need to change the design to improve this heuristics score.
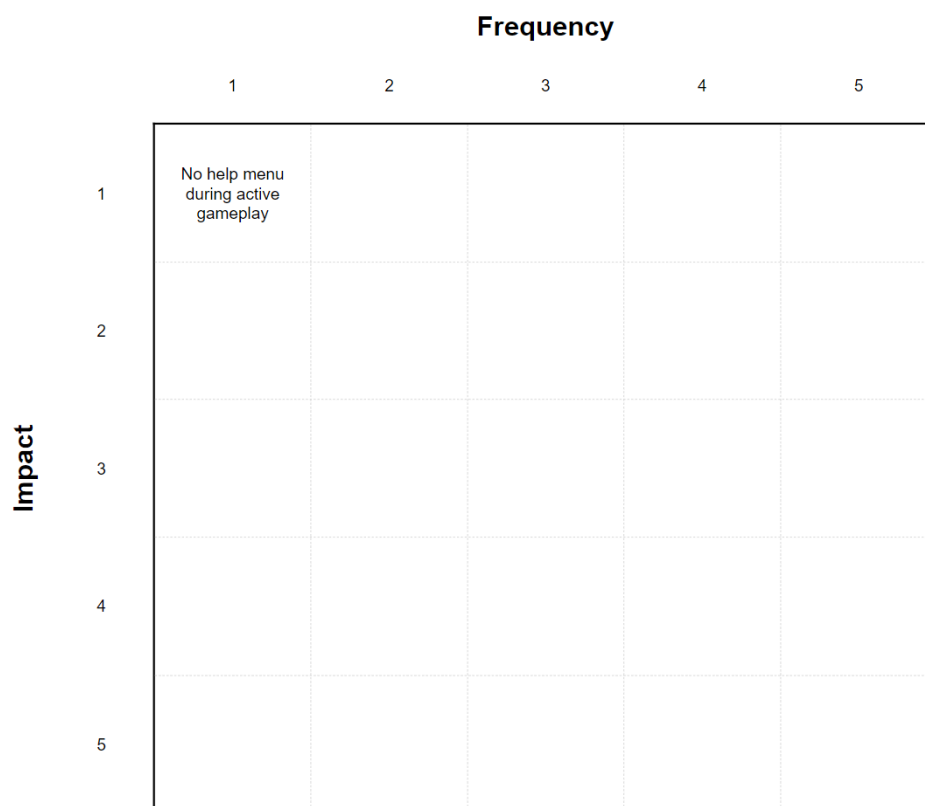
Frequency

	1	2	3	4	5
1	No help menu during active gameplay				
2					
3					
4					
5					

(Impact on vertical axis)

Figure 6. Frequency-impact plot for issues related to the "recognition rather than recall" heuristic.

Help users recognize, diagnose, and recover from errors

For this section, the evaluators were provided with an image from our design displaying the main page which showed an error saying that the chosen username is already being used (Baykal et al., 2022). One of the evaluators considered that it is not clear that there is an error. Afterwards, the evaluators were asked if they knew the cause of the error and a possible solution. Some of the evaluators had different opinions about this, there was also a notable feedback that our design makes the error unreadable (Baykal et al., 2022). From this we concluded that readability of errors from our design should be considered. We also should include more information with our errors so that the users don't get confused. Overall, for this section, the evaluators' ratings averaged at 4.2 out of 5 (Baykal et al., 2022).
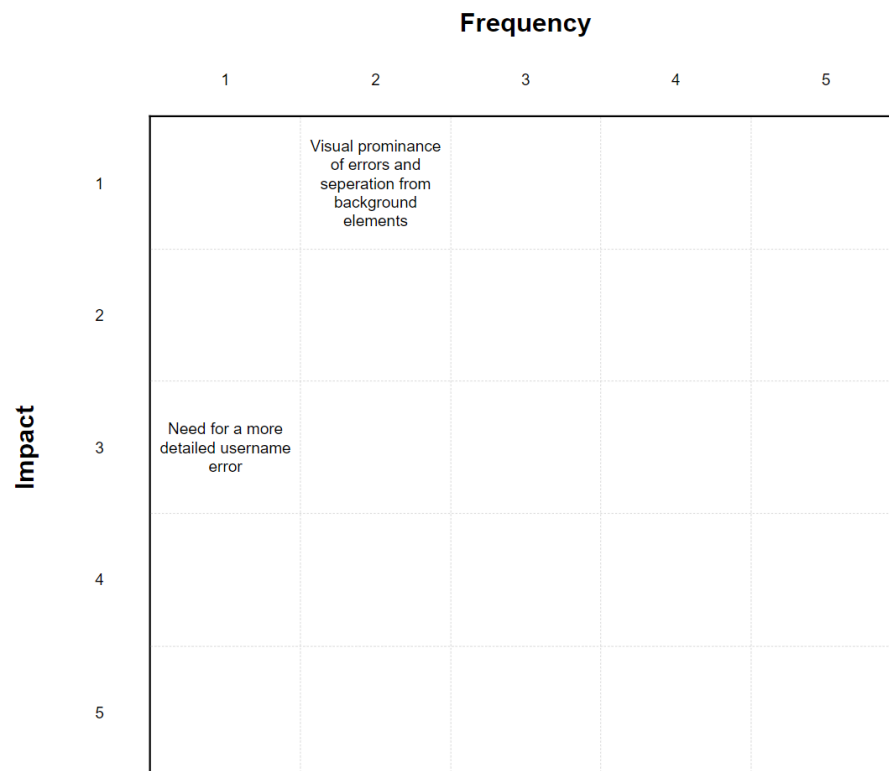
Frequency

	1	2	3	4	5
1		Visual prominance of errors and seperation from background elements			
2					
3	Need for a more detailed username error				
4					
5					

Impact

Figure 7. Frequency-impact plot for issues related to the "help users recognize, diagnose, and recover from errors" heuristic.

Help and documentation

For this section, our evaluators were asked if they can easily find a help page, which got positive responses from everyone (Baykal et al., 2022). From this we can deduce that it is intuitive to find help. Considering the feedback from the evaluators, there is no issue regarding this heuristic. For this section, the evaluators' ratings stand at a 5 out of 5 (Baykal et al., 2022).

User control and freedom

According to the result of the survey, the evaluators gave the 'User control and freedom' part a score of 4.6/5 (Baykal et al., 2022). The questions are mainly about whether the user can decide their game status anytime. The result shows that our design does give the freedom to users to make any choice. Though our design gets a relatively high score, we believe there is still more freedom we can offer to users.

Consistency and standards

For this section, our question provided to the evaluators is about whether users can understand the the same effect in different parts of the game screen structure (Baykal et al., 2022). Though our design got a score of 4/5, evaluators showed different opinions on this part. The majority of evaluators clearly understood and described the different meanings of the red answer button and the joker button while some evaluators stated that there may still exist uncertainty as they can't find out what the joker button is and why the button is red (Baykal et al., 2022). As this question rarely influences the game process, we give it an impact score of 2/5 and a frequency score 2/5. In further improvement, the effect and explanation of the joker button will be included in the help documentation and additional constraints may be added.
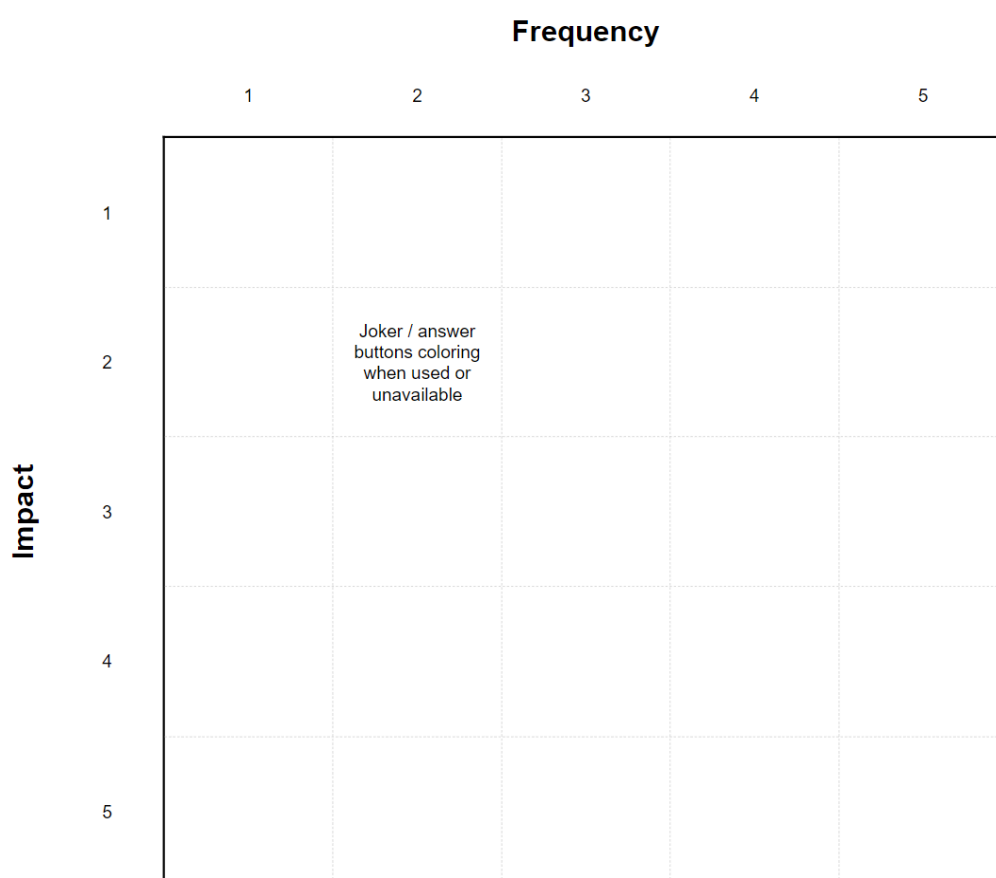
Frequency

Figure 8. Frequency-impact plot for issues related to the "consistency and standards" heuristic.

Conclusions & Improvements

What are your main conclusions from these results?

In conclusion, our evaluators, most importantly, complained about mainly three aspects of our design:

- Visibility
- Aesthetic organisation
- Accessible button positions

We have seen that our use of high-contrast colours on the background graphic has caused high impact low frequency visibility problems. Along with that, our use of randomised locations for user names on the waiting room view was received negatively, which was a low impact, low frequency problem. We have also received mixed feedback about the locations and colours of various buttons and controls. As this issue was observed by everyone, we viewed it as a medium impact, high frequency problem.
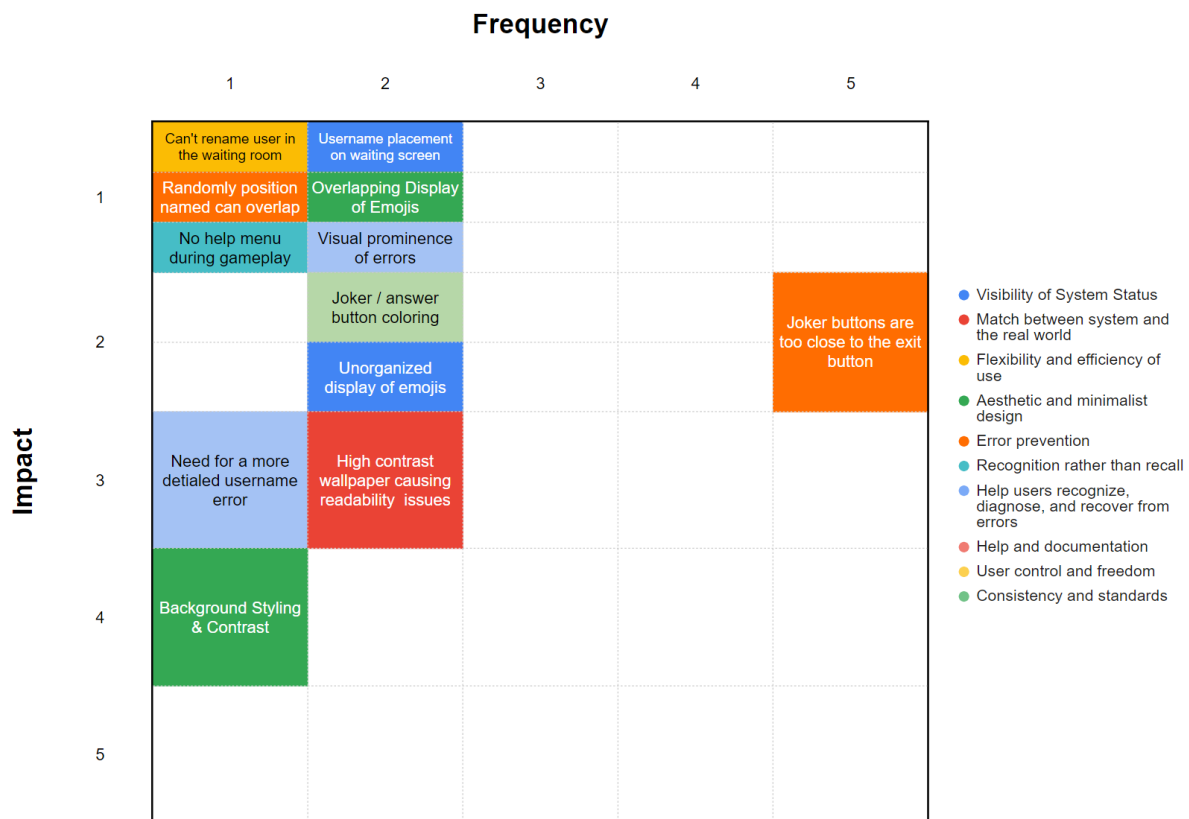
Frequency

	1	2	3	4	5
1	Can't rename user in the waiting room / Randomly position named can overlap / No help menu during gameplay	Username placement on waiting screen / Overlapping Display of Emojis / Visual prominence of errors			
2		Joker / answer button coloring / Unorganized display of emojis			Joker buttons are too close to the exit button
3	Need for a more detialed username error	High contrast wallpaper causing readability issues			
4	Background Styling & Contrast				
5					

Impact

- ● Visibility of System Status
- ● Match between system and the real world
- ● Flexibility and efficiency of use
- ● Aesthetic and minimalist design
- ● Error prevention
- ● Recognition rather than recall
- ● Help users recognize, diagnose, and recover from errors
- ● Help and documentation
- ● User control and freedom
- ● Consistency and standards

Figure 9. Frequency-impact plot of identified problems of all chosen heuristics.

Improvements

To address the problems our evaluators pointed out. We decided to make the following changes to our final design:

- We will fill the backgrounds of the buttons and change the colour of the text inside, to make the text inside the buttons more readable and provide better contrast with the wallpaper.
- We will relocate / resize some text fields and buttons for better accessibility.
- We will make sure that the emojis do not cover important parts of the game screen by slowly fading them away after they're submitted.

However, we have also received some feedback that was influenced by the personal creative preferences of our evaluators, such as our approach to emoji-based communication and display. We have decided to stick with our implementation, as the rather "random" appearance of the emojis were intended to keep the users submitting them anonymous.

Final GUI Design

Note that this report is only about the design, so it is not necessary to already show the finished improved implementation.

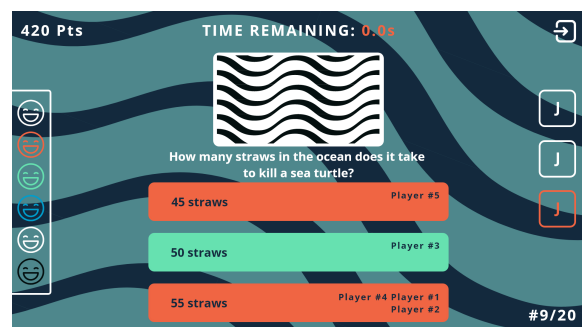*Figure 10.* Previous design for round ending.

Figure 11. Improved design for round ending.

A lot of changes were made to the screen which shows who answered what. First of all the jokers have been shifted downwards so that people would not accidentally click on the exit button instead of the joker. The answer buttons have been given a solid background colour and the colour of the text inside has been changed for readability purposes. Also the emojis have been given colours to show that they are all different.
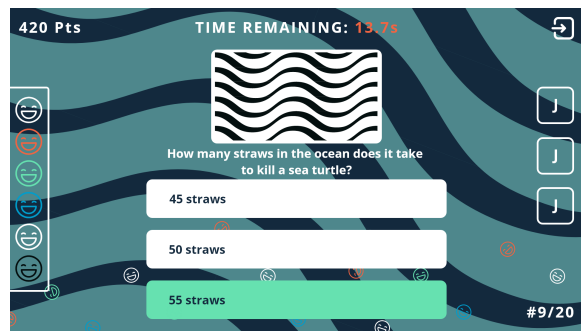
Figure 12. Old design for submitting an answer

Figure 13. Improved design for submitting an answer

Similar changes have been made to this view as to the previous view. For example the jokers have been shifted down and the buttons have been given a solid background colour. Something else that has been changed is that the emojis are now behind all the texts and buttons. This is to make sure that the text is readable at all times
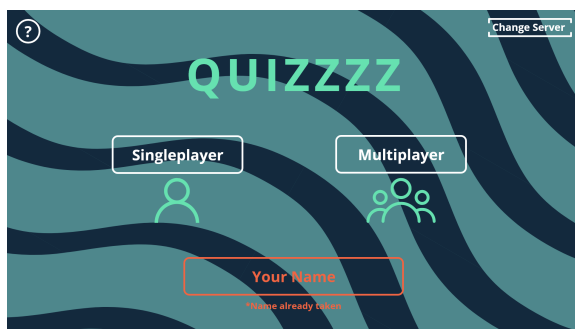

Figure 14. Previous splash screen design.

Figure 15. Improved splash screen design.

Just as with the previous two views the buttons have been filled in and the colour of the text has been changed. The size of the error message has been increased to make the text more readable. A submit button has also been added. This makes it more clear as to how a player should set the name they chose. This section also makes use of filled and hollow buttons in hierarchical order to direct the user towards higher-importance elements (Yalanska, 2021).

We also added three new views to the design as they were absent from the first iteration of the design.
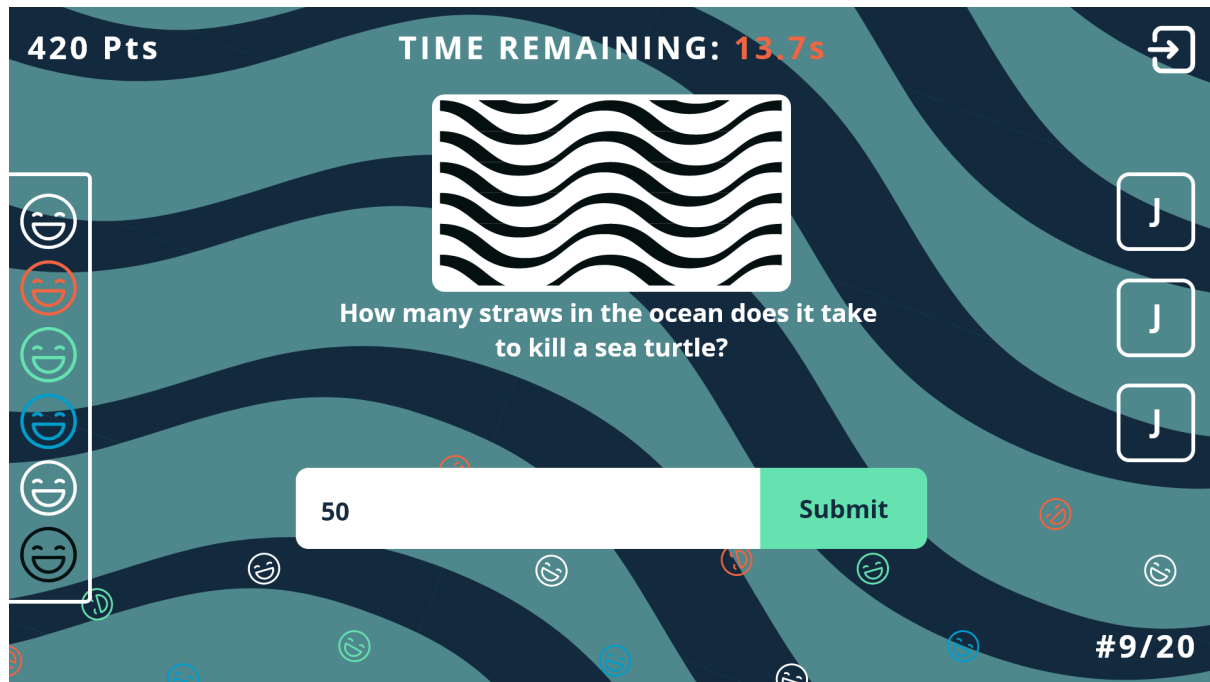

Figure 16. Open-ended question screen design.

This view shows what the question screen for an open question looks like.
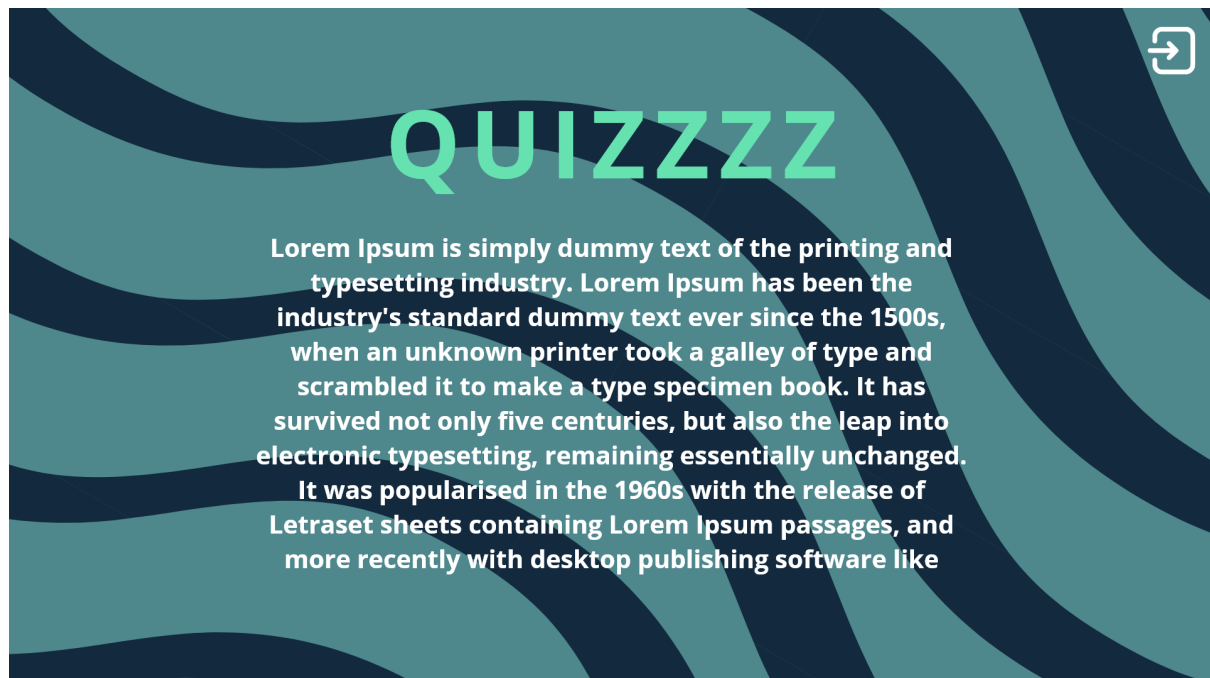

Figure 17. Tutorial page design.

This view shows the help screen which you can access by pressing the encircled question mark icon. Dummy text is used as a placeholder for the help text.
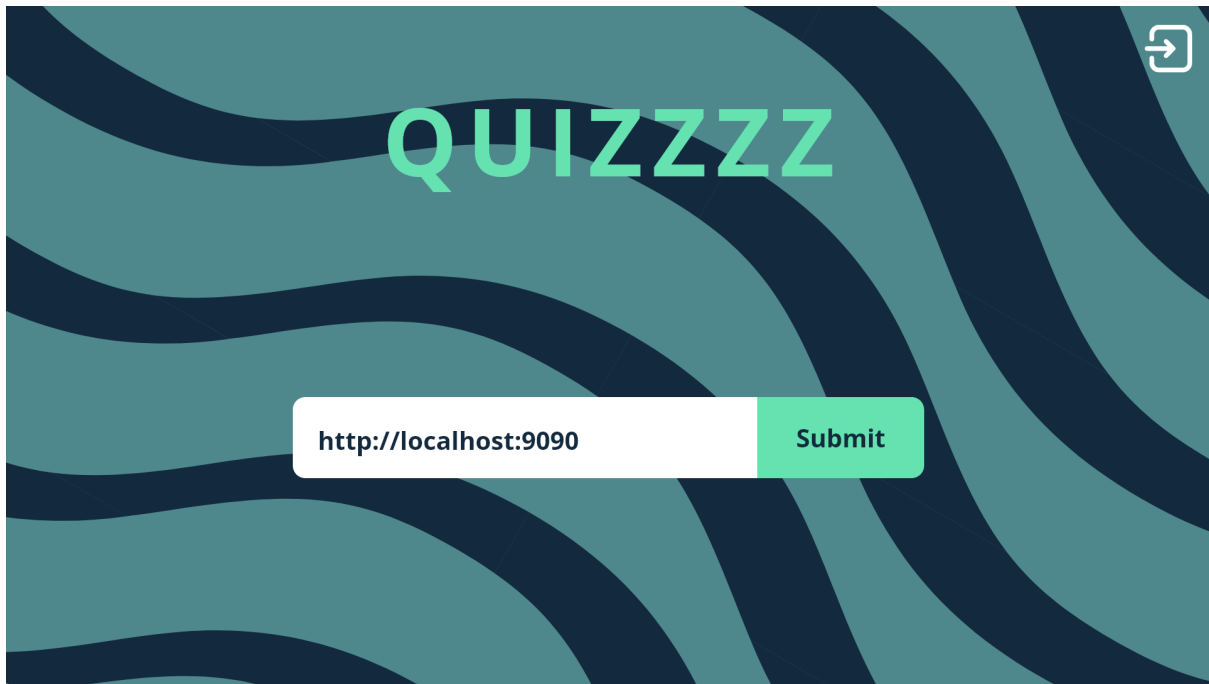
Figure 18. Server connection menu design.

This view shows the screen where you can change the server URL to connect to. This view can be accessed by pressing the "change server" button on the splash screen.

Bibliography

1. Baykal, O., Belder J., Han, X., Naydenov, Y., Toşa, L., Woods, J. (2022, February 24). *QUIZZZZ Heuristics report.* Google Forms. https://docs.google.com/forms/d/1P1ffj9HUGFbHnh6R-XLKbAyb-YQKzcIej_Aj9LYYgnI/
2. Nielsen, J. (1995). How to conduct a heuristic evaluation. Nielsen Norman Group, 1(1), 8.
3. Nielsen, J. (1994a). Enhancing the explanatory power of usability heuristics. Proc. ACM CHI'94 Conf. (Boston, MA, April 24-28), 152-158. https://doi.org/10.1145/191666.191729
4. Kovačević, N., Stojiljković, A., & Kovač, M. (2019). Application of the matrix approach in risk assessment. Operational Research in Engineering Sciences: Theory and Applications, 2(3), 55-64. https://doi.org/10.31181/oresta1903055k
5. Baykal, O., Belder J., Han, X., Naydenov, Y., Toşa, L., Woods, J. (2022, February 24). *QUIZZZZ Heuristics report - Form answers* [Data set]. https://docs.google.com/forms/d/1P1ffj9HUGFbHnh6R-XLKbAyb-YQKzcIej_Aj9LYYgnI/edit#responses
6. Yalanska, M. (2021, October 7). UI Design: Basic Types of Buttons in User Interfaces. Tubik Blog: Articles About Design. Retrieved March 2, 2022, from https://blog.tubikstudio.com/ui-design-basic-types-of-buttons-in-user-interfaces/