Figma vs Framer - A Between Subject Lab Study

Seyedehdelyar Tabatabai
Computer science
Oregon State University
Corvallis, OR, USA
tabatase@oregonstate.edu

Puja Agarwal
Computer science
Oregon State University
Corvallis, OR, USA
agarwalp@oregonstate.edu

Pavan Sai Nallagoni
Computer science
Oregon State University
Corvallis, OR, USA
nallagop@oregonstate.edu

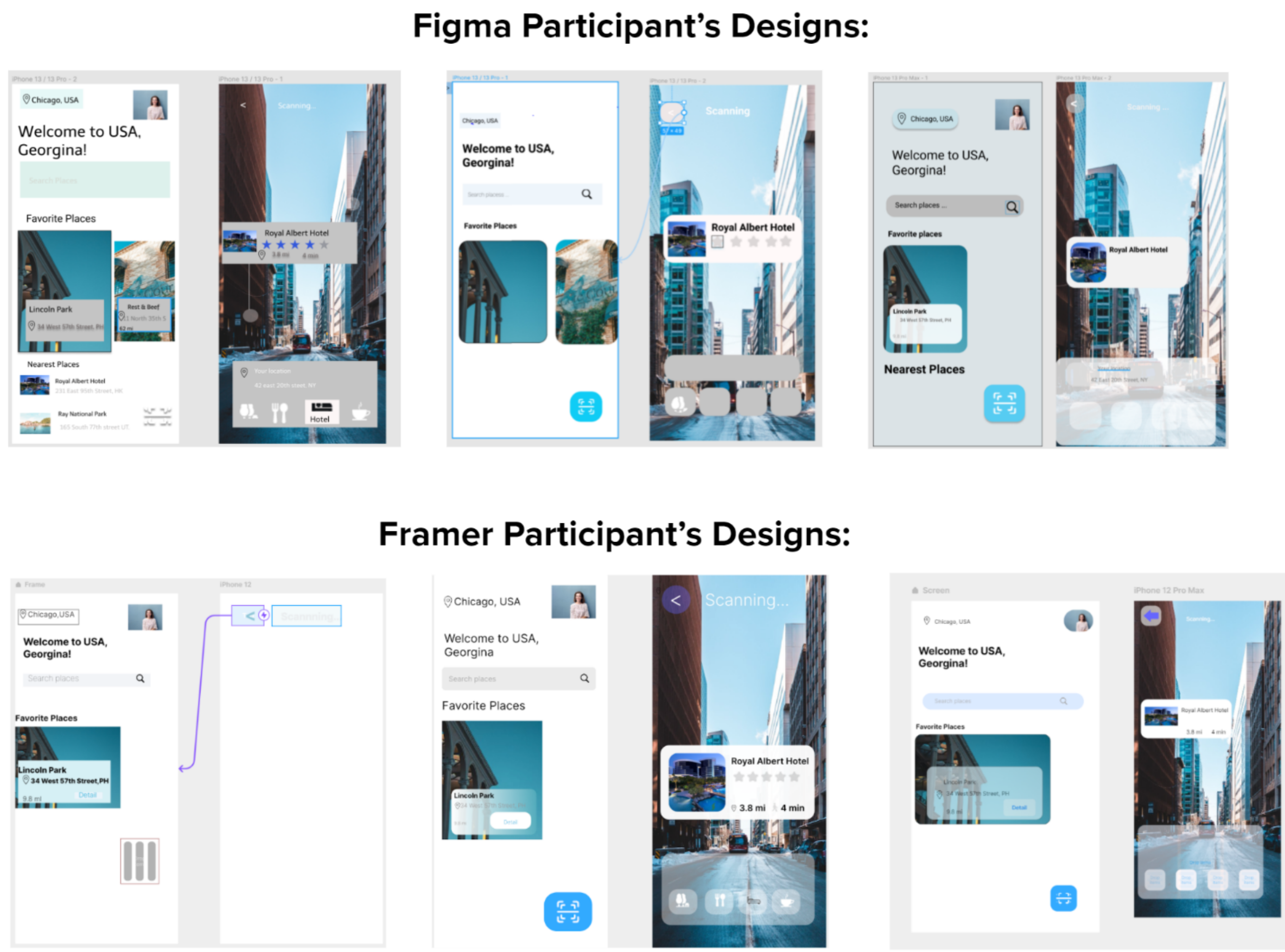

Figure 1: Screenshots of participants' designs in Figma and Framer.

ABSTRACT

Background: Figma.com[1] and Framer.com[2] are two web-based platforms and are vector graphics editors and prototyping tools. These two tools have many similarities and differences. Users who are completely new to the area of digital designing and need a starting point to enter this area struggle to evaluate both of these tools for many different features. Since, they need to choose one of them to begin their work and it is very time consuming as they do not have enough knowledge for this evaluation. **Aims:** We posit that the tool which has better effectiveness in designing the prototype, efficiency in learnability of prototype and usability should be recommended to the new users. **Method:** We recruited 30 participants and divided them into two groups equally and made them work on one tool each. We created two tasks and a post-study questionnaire, where we calculated the accuracy of their designs, time taken for the learnability of the prototype and the ease of use of tools. **Results:** We found that the accuracy of performing the design tasks is better in Framer when compared to Figma. Similarly, time taken to learn the prototyping feature is less in Framer than in Figma. **Conclusion:** Considering the results we got from the statistical analysis of two tasks and a post-study questionnaire, we concluded that the Framer has better effectiveness in design tasks, efficiency in learning the prototyping features and ease of use than Figma.

Keywords

Figma; Framer; User Interface, User Experience

1. INTRODUCTION

With the growth of the use of different software tools in our daily lives, we spend a huge amount of our time on different software tools such as mobile apps, websites, etc. With this surge of using new technologies, the User Interface (UI) and User Experience (UX) design is getting much more important to save time, increase usability, and use fewer resources. Many tools come in handy for designers to help them design the interface and experience for the software users. However, an UI/UX designer needs to decide which tool is more efficient, effective, and usable for them in this manner. In this study, we are going to perform a lab study to test how two of the design/prototyping tools differ in some features. **The goal of this study is to analyze and compare which of the two design/prototyping tools (i.e. Figma and Framer) are more efficient, effective and more learnable for users new to the digital design area.** Figma is highly preferred by both academics and industries and when hiring UI/UX designers, familiarity with Figma is mostly one of the basic requirements. It allows the creation of custom icons and illustrations necessary to communicating ideas and stories with quick action convenience. Figma has specific features that Framer does not have. Figma has features such as version controlling, a plug-in store and provided paid assets. Framer is a free, open-source software that is readily available for the study. With Framer, users can validate their designs in a real-time manner with the environment provided in Framer. It also has design features that Figma does not have like max-width and logics between designs. It helps in iterating and animating interface ideas for an app, website, or product with its convenient native tools. In addition, Framer can generate a quick QR code which helps in sharing the prototype with stakeholders/testers. Framer also has a feature to generate code for programmers based on the designs that Figma does not have. Thus, both have comparable designing features with certain distinct differences. We will focus mainly on the designing features of the two tools to analyze and compare their designing features. Towards the end of the study, if we conclude that Framer is more effective, efficient, and usable in implementing a given design, users may use Framer instead of Figma and save money.

Imagine Jane who has no experience with digital design tools; She is going to start a new project for herself and she needs to decide which of the Figma and Framer tools suit her design needs better. Also she does not have enough time and knowledge to compare both of these tools so that she can understand which one is more suitable for new users. New users to these tools have no knowledge of the tools on how to design screens and create interactive prototypes. Prototypes are a model of an application which guide how the final product will actually look like with the proposed design. Both Figma and Framer have a feature that users can implement interaction between the screens to model their designs.

In this research, in order to analyze the tools, we have three research questions to be answered as follows:

RQ1: - Effectiveness:
How accurately can users implement a design of a given application using Figma and Framer?
RQ2: - Efficiency in Learnability of Prototyping:
Which of the tools is more efficient in learning the prototyping features for new users?
RQ3: - Usability:
How is the ease of use different in these two platforms for different tasks? How satisfied are new users with working with Figma or Framer?

In this study, we recruit two groups of computer science students who do not have any previous expertise in any design/prototyping tools. The first group is going to test Figma and the second group is going to test Framer. A set of same design tasks are going to be given to both groups. These tasks include implementation of two screenshots of an existing software exactly in Figma and Framer. In this case, our ground truth is the screenshots and how accurately users are able to implement the design in the tool is going to be measured by grading them based on this ground truth [**RQ1**]. Details about grading are explained in next sections. This grading data is qualitative and is based on how similar the implemented designs are to the original screenshots regarding the use of colors, positions, proportions, sizes, etc. We are going to break the designing process down into multiple design tasks. During the lab study, we record participant's screens, audio, and also the researchers' observations. Also we are going to have a post-task questionnaire. These data are going to be analyzed later for usability analysis of Figma and Framer [**RQ3**] . Also, the time that each user spent on the interaction design task in these tools is being analyzed to assess the [**RQ2**]. After the tasks are completed, users are going to be asked about the ease of use of the tools with a questionnaire[**RQ3**]. So, the conclusions are going to be made from different sources (i.e.Observations, recordings). The variety of data sources increase the triangulation of this study.

2. METHODOLOGY

In this section we provide information about our data collection, recruitment strategy and data analysis method.

2.1 Study Design and Recruitment Strategy

In this section, specific study design is described along with justification for the selection.

This study involves analysis and comparison of common design features of Figma and Framer for which we've decided to perform a **between-subject** lab study where 30 participants will be selected and equally divided into two groups to perform the given tasks. We chose to do a between-subject study because our participants do not have any prior experience with the tools and during the lab study, they do not have enough time to learn both of the tools with the time given to them. Also, with this study design, tasks can be reused for both groups of users and also the learning effect is eliminated. Here, one group will be using Figma and the other will be using Framer. We sent out an email to Oregon State University's students mailing list and asked for volunteer participants. Thirty participants are recruited randomly based on the responses we got from the emails which we sent to the University's mailing list. The target population includes OSU students with no prior experience in design tools and who are not enrolled in CS567 for Fall term 2021.

The selected 30 participants are representative of our target population because they are students and do not have prior experience with the study tools. These 30 participants include OSU undergraduate or graduate students with no prior experience with either of the tools and exclude the students from CS 567, Fall 2021. We randomly selected 30 participants from the responses we got from the OSU mailing list. These participants were recruited based on pre-study questionnaires which we sent them using the OSU mailing list. These questionnaires include the following

1. Are you 18 years old or over? (Yes/No)
2. Do you have any prior experience with Figma? (Yes/No)
3. Do you have any prior experience with Framer? (Yes/No)
4. Do you have any experience with any other design/prototyping tools? (Yes/No)
5. Are you a student of CS 567 Fall 2021? (Yes/No)
6. Are you willing to participate in our lab study? (Yes/No)
7. Are you able to devote 2 hours based on our availability?

Later, we divided the participants equally (15 in each) into two groups. The group assignments are being done randomly and the randomization strategy is that we used a random team generator website called Commentpicker[3]. This method eliminates the bias in recruiting and group assignment.

2.2 Replicability, Study Tasks Description and Tutorial

In this section, we describe the tasks in detail so that other researchers can replicate the procedure in the future studies.

Before beginning the study, the participants sign the informed consent form which is provided to them by the researchers. The questions in the informed consent are already approved. Here are the data provided in the informed consent: (1) The purpose of the form, (2) Why the research study is being done, (3) Why participant is chosen to participate in the study, (4) What will happen if the participant takes part in the study, (5) What are the beneficence of the study, (6) Will the participants be paid or not, (7) Who have access to the information of the user studies, (8) Voluntariness of the participation in the study and (9) Who to contact in case of any questions after the study.

This consent form includes the details participants need to do in the study. Also, the goal of this study and it's beneficence is included in the informed consent form with minor details being neglected in order to avoid expectancy bias in the participants. After that, a pre-study questionnaire is given to get the participant's information and questions about their prior experience with design tools and their field of study. These questions include their age, their experience with design tools and their willingness to participate in the study. Later on, participants' names will be coded into the participant numbers for confidentiality purposes.

A brief tutorial is given to the participants of both groups on how to use the tools (Group 1 - Figma, Group 2 - Framer) to be consistent with all the participants. The tutorials are in the same level of difficulty and detail for both groups. Tutorials are not very comprehensive because the goal of the study is to analyze the learnability of the tools for new users so we want participants to explore the tools themselves and find out how features work in both tools. The tutorials for task 1, only contained general information about the tool and for task 2 included showing an example of what they need to have as a final result and what interaction task means. The tutorializer describes the tools through a tutorial script that they have written so that they are consistent with all the participants. Each participant of Group 1 will be given access to the Figma tool and similarly, each participant from Group 2 will be given access to the Framer tool. Also, a handout containing the screenshots with desired final outputs are given to each participant. Each handout is numbered with a primary key with respect to the participant and it consists of a list of tasks they need to perform.

In the tutorials, we make sure that no subject expectancy bias is affecting the participants behaviors by having a neutral language towards the tools. We are allowing the participants to ask any questions related to the studies during the study but they were told that they should not use search engines or ask questions related to the tools themselves. The study goals are going to be explained to the users and the study facilitator is going to read the study script to the users to be consistent with all the participants. Since we are recording the time for the interaction task to evaluate the learnability and the efficiency, we do not ask the participants to do think-aloud because it creates a cognitive load for them. Their responses to the tasks are recorded in the exported files and screenshots. Basic operations are being taught to them in the handouts about the tools and they are not allowed to use the internet or documentation. The facilitator guides the participants throughout the study. However, participants are not allowed to ask questions during the tasks since we want them to perform it themselves without the researcher's inference. The researcher interventions happen when the participants are running out of time for each task. Participants are free to drop out the study before, during or after the study.

In this study we are having 2 main tasks. Task 1 is replication of a

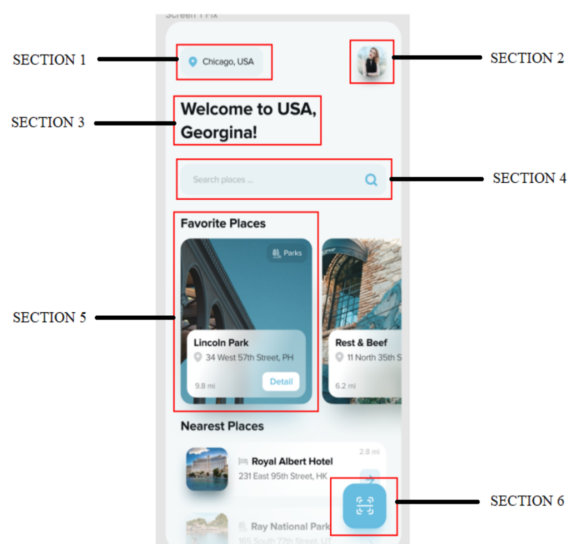

Figure 2.1 : Representation of Task 1 with different sections marked

given screenshot design the Framer/Figma which itself includes 10 sub-tasks. Task 2 includes performing two interactions with the prototyping feature of the tools which participants need to find how to do it themselves and the time taken to be done is being recorded.

Figure 2.2 : Representation of Task 1 with different sections marked

Task 1: Page Interactions

The participants were given 45 minutes to perform this task. This task involves the following sections

a. Create an iphone 13 frame as shown in Figure 2.1.
b. Design the sections which are shown in Figure 2.1 as close as possible to the given design.
c. Create a new iphone 13 frame as Screen 2, Figure 2.2.
d. Design the sections which are shown in Figure 2.2 as close as possible to the given design.

During these tasks, the participants were free to skip the sections and come back again to perform those sections. Ultimately, we're going to compare the final screenshot of the participants' design with the required design based on the grading criteria which is described in the Analysis Section. This will be used in answering how accurately users can implement a design of a given application using Figma and Framer without having any prior experience and knowledge of the tools **[RQ1]**.

Task 2: Page Interactions

The participants were given a limit of 15 minutes to perform this task. In this task, the participant needs to perform the page interactions on the two iPhone 13 frames which they've created for task 1. This task involves the following steps:

a. Create an interaction using the scan icon to navigate to the second screen.
b. Similarly, create an interaction using the back button to navigate back to the previous screen.

All of the 6 participants were able to complete task 2 successfully without any help but each of them took different time to do this task. We are recording the time taken to complete the task which will be used in answering the efficiency of the learnability of the prototyping and interaction features for new users in Figma and Framer **[RQ2]** . After each task, the participants need to check the completed task on the given handout. Towards the end of the activity, we asked users to fill out a short questionnaire which is used in answering **[RQ3]**. These questions are standard and based on the SUS Questionnaire(System Usability Scale)[4]. Here we have used 5 positive questions and 5 negative questions in order to eliminate the questions to be leading. The questions are ordered from positives and negatives alternatively. The answers to these questions are in a 5 scale format from "Strongly Disagree" to "Strongly Agree".

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

2.3 Data Collection, Method Applicability, and Data Analysis

In this section, a detailed plan is described for collecting and managing data from user studies and relating the data to answer research questions..

The data collection involves collecting screen recordings, voice recordings, task completion time, screenshots of final designs, and questionnaire results. The collected data is stored in shared files in the Google Drive account of OSU where only researchers have access to it. The participant's data are anonymized for IRB purposes. To achieve this, we created a database with all our participant information and encrypted the database using a primary key not related to participant details directly. Furthermore, any tasks or handouts given to the participants are numbered based on these primary keys. Here is how each piece of collected data will be used to answer which parts of the research questions:

RQ1: For task1, our ground truth is the screenshots and how accurately users are able to implement the design in the tool is going to be measured by grading them based on this ground truth. The screenshots of the final output of each task completed by a participant are obtained from screen recordings. These screenshots answer how accurately and effectively users are able to implement the design in the tool which is going to be measured by grading them based on this ground truth. The grading for task 1 is done based on the below Table 1, where we give points to each of the criteria. Then we summed the given grades of the sections for each participant. The maximum score that each participant can earn based on this criteria is 34.

Table 1: Grading rubric for Task 1

Section #	Features Evaluated	Points
Section 1	coloring background, importing icon, coloring the icon	3
Section 2	importing the picture, rounding the picture proportionally, shadow	3
Section 3	font type and size, font alignment, font color	3
Section 4	blank text box insertion, rounding, icon upload	3
Section 5	rounding borders(2), overlaying text box(2), shadows overlay(2)	6
Section 6	color as per reference, layering icon on background, size and position of icon	3
Section 7	inserting the character, blur background, rounding corners	3
Section 8	inserting text, text color, text alignment	3
Section 9	alignment of the stars, equal distance between the stars, coloring the star icons	3
Section 10	rectangle blur effect(2), icon upload, icon blur effect	4

RQ2: The time that each user spends to learn the prototyping features to perform task 2 in these tools is being analyzed to assess the efficiency. The recorded time is used to understand which tool has a faster learnability for prototyping feature.

RQ3: After the tasks are completed, users are asked questions about the ease of use of the tools with a post-study questionnaire (we mentioned these questions in section 2.1). In addition, the observers write down any specific issues related to the participants while performing the tasks.

Hence, the data collected is useful in giving out the desired results by answering our research questions.

We previously discussed how the research questions are linked to our data collection methods. After the studies are completed, we are going to perform an analysis over the data that we have collected as below:

- **RQ1**: We have collected the exported files of the users that shows us their final completed tasks. Then we go over our grading criteria mentioned previously in the study design section and we are going to grade the participants based on that. By comparing the mean of the grades of the two tools, first we make sure that their difference is statistically significant or not. If statistically significant, we conclude that the tool with higher mean of the grades is more efficient than the other tool. If not statistically significant, we conclude that the effectiveness of both of the tools are the same.
- **RQ2**: We have recorded the time taken to complete the tasks for both tools for each participant. Next, we are going to take an average time for both groups Figma and Framer respectively. Finally, we can conclude which tool is more efficient in learning and performing the prototyping features by comparing the results.
- **RQ3**: After the completion of tasks, a post-study questionnaire is given to the users, and they are going to fill out the form. Some of the questions are quantitative, and some are qualitative. Next, we will perform a qualitative analysis over the qualitative questions and perform a statistical analysis like calculating an average from the responses to quantitative questions. Finally, we compare the findings of both the groups and can conclude which has the better ease of use over the other.

2.4 Research Roles

In this section, we'll give a description of roles and assigned team members. In this study, we are having several different roles for researchers that are described as follows:

Recorder (Pavan Sai Nallagoni)
Writes down the observations
Keep track of the time.
Keeps track of important qualitative details about the participants and the study.

Tutorializer (Delyar Tabatabai)
Gives a walkthrough in the tool for the participants
Explains the tasks to the participants.
Answers to the potential questions that the participants may come up with.

Facilitator (Puja Agarwal)
Asks the participants to fill out the pre and post-study questionnaire.
Asks the participants to read through the informed consent form.

2.5 Sandboxing

Before starting the actual user studies, we performed one sandboxing user study in order to make sure that the study design is suitable for participants and find out what problems they will face during the user studies.

First issue that we faced during the sandboxing was the amount of time that the participant spent on figuring what images and icons to use from the given original design. So based on this, we decided to prepare all the images and icons for the participants before the studies begin in order to eliminate the time that participants are going to spend only on finding the icons and images.

Secondly, we figured out that if we divide the whole task into sub-tasks, participants will have a more accurate path towards performing the task and also it would be easier for us to grade the whole design task. So we divided the task1 into 10 sub-sections.

Before starting the sandboxing, we assumed that 30 minutes is enough for a user to be able to design two screens. However after the sandboxing we recognized that our assumption was not correct and users with no prior knowledge needed more time to process what the tool is and what features it has so we increased the first task 's time to 45 minutes.

2.6 Study Procedure

In this section, we'll elaborate on the processes that will help to guarantee non-biased study sessions.

Firstly, there are minimized learning effects since we have a between-subjects study design. Secondly, there are no individual differences here since all the participants have no prior experience with either of the tools. Furthermore, we're going to give a brief tutorial about the tools on which participants perform. This tutorial is made in a way that doesn't involve any subject-expectancy bias i.e., keeping the tutorial short, clear and make the participants familiar with the tool as per the tutorial and not to jeopardize them with the tasks. Also, in the informed consent, we did not ask participants for any prior experience, demographic information and how research goals are going to affect them, which might cause a subject expectancy bias. Participants are told that their answers to the tasks do not put them in any risky positions. We are eliminating the observer expectancy bias by not influencing the participants behavior by looking at them in a power position. Confirmation bias is eliminated in this study as we do not have a prior hypothesis in mind that could affect the way that we analyze the data to confirm our beliefs. The participants recruited were not chosen in any manner that they might be sympathetic to our expectations and research goals. Furthermore, our treatments and group assignment methods were done randomly.

3. ANALYSIS

We collected qualitative and quantitative data from our participants for the lab study between Figma and Framer. We analyzed the quantitative data by using statistical methods. We used the R programming language to perform the statistical analysis. To analyze the qualitative data, we used the observation notes to strengthen our results from quantitative analysis.

RQ1: To answer this question, we graded Task 1 based on the grading scale defined in the methodology part. We can see the detailed grading scale in Table 1 and the outcome of grading of the section tasks in Figure 3. We can also see the design results for each participant in Figure 1. Using these data, we graded the responses using the grading scale. We imported the graded data per section per participant through a .csv file onto the R Studio software for Figma and Framer separately. The data we collected from 15 participants per tool (i.e. Figma and Framer), we tested the data for normality using the Shapiro Wilk Normality Test[6]. Shapiro Wilk Normality Test is powerful for small data. Since we had 15 rows of data for each tool, the Shapiro Wilk Normality Test was the best option for us. After performing the normality test on the data, we found that the data is not normal. We then performed inferential statistics to infer the data. We performed a non-parametric Wilcoxon Test[8] as our data was not normal. We observed that the p value obtained from this test is <0.05, i.e. the differences of the means is statistically significant. We also observed the mean values of Figma is 17 and Framer is 21.93. In order to test the effect size, we are using Cohen's d[5] formula which uses mean and standard deviation of the data.

RQ2: To answer this RQ, we obtained the timestamp from the screen recordings for the time required to perform the interaction task for prototyping, ie. Task 2 in this case. Figure 4 shows the graphical representation of the time required by participants. We saved the data onto a .csv file and imported the file onto R Studio software for statistical analysis. We performed a normality test on the data and chose the Shapiro Wilk Normality Test to test our data because it is powerful for small data. Since we had 15 rows of data for each tool, the Shapiro Wilk Normality Test was the best option for us. On performing the normality test. we observed the p value to be >0.05, which indicated that our data is normal. We went ahead with using parametric t.test[7] to perform the inferential statistics on our normal data. We chose t.test to perform the statistical analysis because firstly, we had numeric data, secondly, t.test is used to identify if the difference in the means of the two groups is statistically meaningful or not and lastly, we had the population size of <50. On execution, we observed that the p value obtained is <0.05. This means that the differences of the means are statistically significant. We also observed the mean of Figma is 212.06342 and mean of Framer is 91.24799. We used Cohen's d[5] function in R to compute the effect size of the data on our sample population.

RQ3: For this research question, we used the validated SUS Questionnaire[4] to test the usability of the two prototyping tools. We divided the questionnaire into 2 sections for the analysis purpose only, one section containing the positives questions and another section containing the negative questions. The questions having a positive lead are termed here as positive questions and the questions having a negative lead are termed as negative questions here. The mean of the responses based on the questionnaire is shown in Figure 5.1 and Figure 5.2 respectively. We quantified the responses of participants on a scale from 0-4 and calculated the sum of all the responses per participant on the excel sheet. We multiplied the sum value by a factor of 2.5 to convert the original scores of 0-40 to 0-100. It was observed that all the values obtained were less than 68, i.e. according to SUS questionnaire research, the value below 68 is considered to be below average. This did not give us a satisfactory analysis. In order to best interpret the results, we recorded this data on a .csv file and uploaded it to the R studio for statistical analysis. We performed the normality test using Shapiro Wilk Test for the same reasons and obtained the data to be normal. We then performed the t-test to find the difference of the mean of the two tools and obtained the p value to be <0.05. We also observed the mean value of Figma to be 42.68057 and Framer to be 51.18386. We used Cohen's d[5] function in R to compute the effect size of the data of our sample.

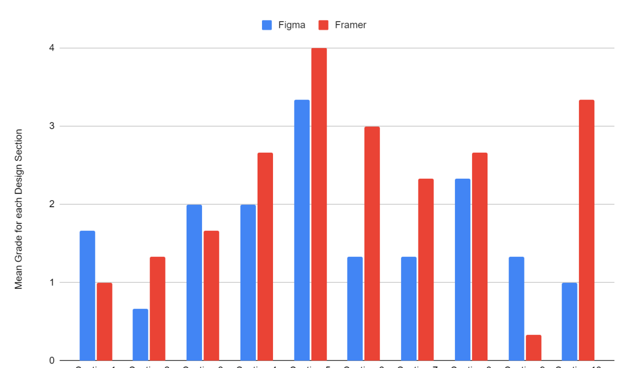

Figure 3: Graphical representation of mean of grading of Task 1 for Figma and Framer

4. RESULTS

RQ1: We performed a number of statistical tests on the graded data of the section tasks to analyze the data collected. We observed that the Wilcoxon Tests gives a statistically significant p value. This implies that the null hypothesis can be rejected for this test, i.e. there is a marked difference between the tools Figma and Framer. We tested the effect size of the data via Cohen's d[5] test and found that the value of d is 2.83. This implies that the difference between Figma and Framer is large on the grades. Hence Framer is more effective than Figma in accurately implementing the given design tasks.

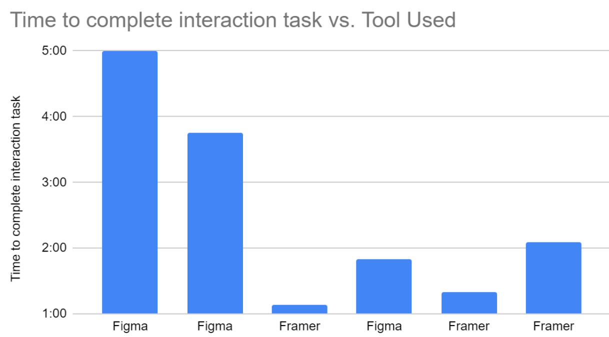

Figure 4: Graphical representation of the interaction time required for Figma and Framer

RQ2: On performing the statistical analysis on the data for interaction task, ie. Task 2, we observed that we have a normal set of data. Also on performing the inferential statistics, we found out that the difference in means of the two tools is statistically significant, ie. the null hypothesis can be rejected. This means that there is a significant difference between the two prototyping tools. Now, in order to test that direction of flow of the means, i.e. effect size, we used Cohen's d[5] test. The value obtained for d in Cohen's test is 149.576. This implies that the difference between Figma and Framer is large in the interaction time. Hence, we could conclude that Framer has a better effect on learnability of prototyping features when compared to Figma.

RQ3: In order to test the usability of the prototyping tool, we used the validated SUS questionnaire. After performing the statistical analysis, we observed that the data we collected from the participants is normally distributed as the p value obtained from the Shapiro Wilk Test is >0.05. From the t-test we concluded that the mean of the difference between the tools is statistically significant. This implies that the null hypothesis can be rejected and there is a marked difference between the effect size of the two tools. In order to calculate the effect size, we used Cohen's d[5] test and found that the value of d is 4.392173. We could conclude that Framer has a better usability compared to Figma for people with no experience.

On a different note, from Figure 1, we can see that the participants of Figma did a fair better job on designing the screens than participants in Framer. But the quality of the designs in Framer was better than that in Figma. This resulted in better grading of the Framer tasks than that of Figma. Hence, although it is visually appealing that Figma was better than Framer, our statistical analysis shows a completely opposite result. The observations done when the participants were performing the task also strengthens our obtained results that Framer is a better tool from an effectiveness, learnability and usability point of view for an audience having no prior experience in designing tools.

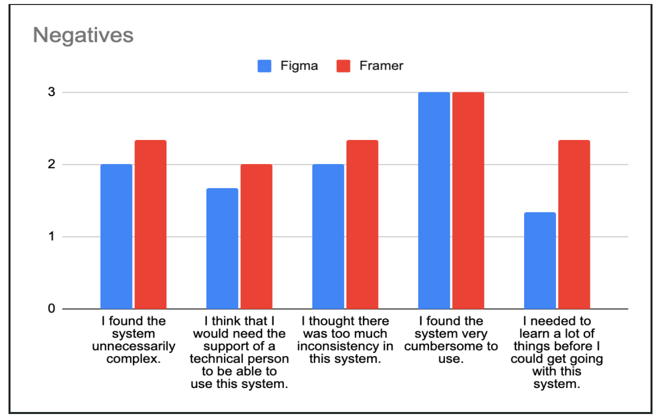

Figure 5.1: Graphical representation of the participants' responses on the negative SUS Questionnaire.

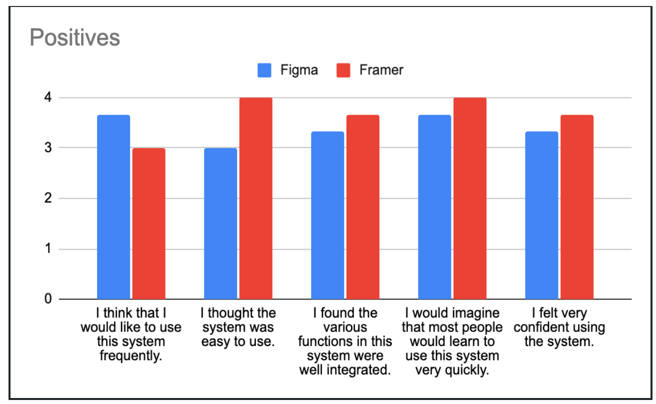

Figure 5.2: Graphical representation of the participants' responses on the positive SUS Questionnaire.

5. THREATS TO VALIDITY

Several possible threats to validity exist in how we conducted our evaluation. First, we want to point out the internal validity threats. In our experiments, we had the limitation of time and participants so it was not possible for us to recruit more than 6 participants for the user studies. So the data gathered from these participants is very limited and causes a threat to the validity of our study. The second internal threat to validity is that we only recruited graduate students of Oregon State University. It is possible that the level of intelligence or understanding of these students is the same with each other so in future research we suggest studying various groups of participants. The next threat to the internal validity is that we could not control variables like the IQ or creativity of the participants so they might be compounding variables to our results as internal validity threats. Another internal threat to the validity is that we did not consider participants' biases. During the studies, many of the students told us about how tired they were; however, we are not analyzing the effect of their fatigue on the studies. Some of the participants were friends of the researchers which makes us think maybe there might be a reason for them to behave differently due to this despite the fact that during the tutorials we mentioned that they need to behave normally and they are not being judged by anyone.

Next, we want to point out the Construct Validity threats. First, we have to say that our data measurements are based on our observations, the questionnaires, and our gradings to the designed prototypes by the participants. These data collections may not be precise because we invented the way of grading designs by only the final screenshot that we get from the participants; however, in fact, the quality of a good design depends on many factors but we are only considering some of them to grade the designs like colors, proportions, location of the elements, sizes, ….

On the other hand, we want to mention what Conclusion Validity Threats we had in our study. So in one part of this study, we had one independent variable which was the tool (Framer vs Figma), and one dependent variable which was the grades that participants got in redesigning a given design. In our statistical analysis, we only compared the mean grade of the two groups of participants with each other. However, a more comprehensive statistical analysis could be performed such as comparing medians. Also since we did not have enough time to gather more data, our collected data and our observations were limited, which leads to having a low statistical power. The limited time to perform the tasks might have caused the participants not to be able to perform their best resulting in false negative results.

Lastly, as the external threat to validity, we have to say that we did our best to make this analysis generalizable but there were some limitations. The low number of participants, the sample population, and many other factors could stop our results from being generalizable. The users who use Figma and Framer in day-to-day life are mostly professional UI/UX designers however we did not have access to them. On the other hand, all of our participants had zero previous experiences with the design tools cause we wanted to be consistent among all of the participants. Some people may not be good at completing tasks in our limited time. So by increasing the variety of participants and the number of them, the study could be done more accurately.

6. DISCUSSION

The main goal of this study was to help new users be able to choose between Framer and Figma when deciding to start design tasks. In this manner we answered three research questions. These research questions build up the conclusion that Framer is easier to use, faster to learn interaction tasks implementation and more desirable to work with among the users who are not familiar with this tool. This result was totally opposite to what we hypothesized because Figma is the tool that is widely used among the UI/UX designers in the industry and in academics as well.. One of our thoughts about this is that since in the industry, the designers are more experienced than our user study participants, they might have different approaches towards preferring a tool in their day to day life. So what we suggest in future studies, is replicating our study design for experienced users to analyze if experienced users and non-experienced users have different preferences on choosing the tool to work with.

In our studies, we did not control variables such as creativity or the intelligence of the participants. In future studies, we suggest analysis with multiple factors. For example factors such as the amount of fatigue can also be measured in future studies.

In the future studies, threats to the validations could be eliminated more. For example this study can be performed on participants who are not only students and come from different backgrounds. Also, study can be expanded to be more generalizable by recruiting more participants in order to gain more accurate results.

The tasks that we have designed in our study were suitable for a two hour study. However it would be interesting to analyze the effect of the task complexity and the length of the tasks on the effectiveness of the participants' results.

We also suggest preparing compensation for the participants to create motivation for them in the studies. In our studies, we were not able to compensate the participants and it might have had an effect on their behaviors.

7. CONCLUSION

In this paper, we study the design and prototyping tools Figma and Framer and conclude the tool which would be more suitable for someone who is new to digital designing and prototyping. We have performed two-hour controlled lab studies among two groups of participants. One group using Figma and the other group using Framer. The participants were all adult students of the Oregon State University and did not have prior experience with any design/prototyping tools. The participants were given two tasks so that we could measure the effectiveness of the tools in designing, the time they spend on learning how to perform an interaction task and their overall satisfaction with the tools. Based on the results we achieved after performing the data collection and statistical analysis, we concluded that Framer is more suitable for new users in all the three aspects that we have analyzed.We observed the mean values of Figma is 17 and Framer is 21.93 in performing the designing tasks. Thus, we can infer that Framer is more effective in finding out how to implement a design for new users than that of Figma. Consequently, the mean time taken to complete the interaction task in Figma and Framer is 212.06342 and 91.24799 respectively. This means that Framer is faster for learning how to make interactions work in Framer than in Figma. Finally, using the System Usability Scale (SUS) [4], the mean scores of post-study questionnaires of Figma is 42.68057 whereas Framer is 51.18386. Hence, our results show that for new users it is more desirable to work with Framer than Figma.

8. REFERENCES

[1] https://www.figma.com/

[2] https://framer.com/

[3] The Comment Picker website https://commentpicker.com/i

[4] System Usability Scale (SUS) https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

[5]Cohen's d https://www.rdocumentation.org/packages/psych/versions/1.7.8/topics/cohen.d

[6] Shapiro test https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/shapiro.test

[7] T test https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/t.test

[8] Wilcox test https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/wilcox.test