THE IMPACT OF ALBUM LAYOUT ON USER EXPERIENCE IN SMARTPHONE MUSIC APPLICATIONS

Molly Scott[1]and Damian Schofield[1]

[1]Department of Computer Science, State University of New York, Oswego, New York, USA

ABSTRACT

The study addresses a significant issue in User Interface (UI) design for smartphone music applications, focusing on the usability impacts of grid-based versus list-based album layouts. Music apps are widely used as smartphones have become central to entertainment consumption. Despite their prevalence, little research has explored how different UI layouts impact user experience in music apps, especially in terms of efficiency, perception, and usability. This gap in knowledge makes the study vital, as optimizing UI layouts can improve the functionality, user satisfaction, and accessibility of music applications, which are integral to millions of users' daily lives.This work addresses this variation by exploring how each layout affects different usability considerations within a music app context, such as the time taken to find a desired album, user error rates, and users' qualitative perceptions of the app. A custom music app interface was created and used for the user testing undertaken in this study. The website guided participants through two usability tests in a simulated music app and automatically collected survey responses and quantitative metrics. This study did not find evidence to support a significant difference in the time it took participants to find a desired album between the two layouts. However, the experimental participants believed that they could find a desired album quickly when using a grid-based layout than when using a list-based layout.

KEYWORDS

Music Applications, information Search, Smartphone Applications,User Testing, User Interface, Usability

1. INTRODUCTION

As smartphone technology has developed, these devices have implemented a wider variety of features and uses that once required additional devices through a process known as convergence. Smartphones are increasingly used to consume different forms of entertainment media across multiple age groups [1], with smartphone-based music listening being especially popular among younger users [2]. Smartphone apps have replaced the need for separate music listening devices such as portable CD players or MP3 players.

The mass adoption of smartphones as music-listening devices has resulted in a proliferation of various music apps and services.Some apps are designed to provide music content as a subscription service, where users can curate a personal library from a music catalog. Other apps are more focused on a traditional e-commerce model where users can purchase music and access their library through an app. A wide range of music app interfaces are available, some are minimalist in design, while others have many more advanced features. Many of these apps have integrated advanced, artificial intelligence based, music recommendation systems to help users discover new music [3].

Despite these many differences, all of the music apps focus on providing an integrated music listening and browsing experience. Hence, the majority of the music apps that are currently available are attempting to solve similar design problems, and must take similar design heuristics and usability considerations into account. In particular, all music apps must consider how to best present lists of music albums, songs, artist names etc to a user, so that a user can quickly find a desired piece of music.

Many music applications have been designed to display information about music albums in either a list format or a grid format. List formats typically feature a smaller album art image next to the album title. In contrast, grid formats typically emphasize the album art and have a smaller album title (Figure 1).

Figure 1. Popular interface design layouts; grid vs. list formats.

These differences in interface format, varying how visual and textual information is displayed, could have potential usability implications. However, little research has been conducted on these differences in a music app context.

1.1. Visual Search

When encountering new content on a web page or within an app, users will automatically scan the user interface to discover important information, landmarks, and affordances. This search process is undertaken in two phases. The first stage consists of scanning the overall interface looking for 'entry points', and the second phase consists of a further inspection of a selected entry point and its immediate surrounding areas [4]. Users will primarily be performing short saccadic eye movements where the eye moves rapidly from fixation to fixation over a period of 20 to 180 milliseconds, and users only fixate on each portion of the interface for 200 to 400 milliseconds [5].

During this visual search, most users will follow a specific scanning pattern with their eye movements. Users who primarily read left-to-right-oriented languages will typically scan content in either an F-shaped pattern or a Z-shaped pattern [6]. On the other hand, users who read in right-to-left-oriented languages will frequently scan content in flipped versions of these patterns. These F-shaped and Z-shaped patterns have been found in users of both desktop and mobile applications [7].

Users habitually conduct visual searches either as part of an 'exploratory' search or a 'goal directed' search. Hence, the visual design of a user interface can affect how users conduct these different types of searches based on the amount of attention different interface elements draw from the user [4]. Some visual designs are created to take advantage the users' natural visual scanning tendencies. Other interfaces influence the way users scan a page through the use of visual hierarchies that are created by applying strong attention-getting characteristics to appropriate elements[8].

1.2. Visual Hierarchy and Grouping

Visual hierarchy and grouping are important design concepts for helping users understand information in any user interface. Visual hierarchies can help give the interface design a focal point and help users to identify the message [6]. Additionally, hierarchy and grouping can help users identify groupings of elements that are similar in their contribution to meaning and direct attention to relationships that are informative, beyond the literal subject matter or visual character of the elements alone. An example of this would be the [8].

Visual hierarchies can be created within a user interface by creating a controlled gradient of difference between items that allows the user to visually rank the importance of each item [8]. These visual hierarchy distinctions can take many forms including :

- These differences can be created through color, positioning, or scaling of individual interface elements [6, 10]
- The perception of hierarchical relationships among elements is also a function of their relative left-right and top-down positions [9].
- Visual groups can be defined by uniform connectedness through a sharedline or common edge or by being enclosed by a shared boundary [6].

1.3. Usability Considerations

Traditionally, user interface designers have relied on techniques like visual hierarchy, grouping, and other gestalt visual processing principles to create usable interfaces [11]. The usability of an interface is heavily dependent on the perception of the user, where users gain an understanding of the interface and its content through gestalt processing. This, in turn, leads users to group and focus on elements appropriately. However, an incorrectly designed interface will cause users to incorrectly group information and other elements [12].

Visual hierarchies provide the simplest method of increasing knowledge about the structure of a system [9]. Additionally, visual hierarchies can be used to properly format information into understandable chunks. This takes advantage of Miller's Law by arranging content into smaller clusters to help users process, understand, and memorize easily [13].

However, an ineffective hierarchy may make it difficult for viewers to know where to look first [6]. If the visual hierarchy fails to create enough contrast between elements, it can result in a design that appears cluttered and flat, which can make it harder for users to find entry points within a user interface [4]. According to Hick's Law, this can place additional cognitive load on users and decrease the usability of an interface [13].

It has been found that the visual scanning pattern users apply to an interface can also affect their experience. Users who follow an F-shaped visual scanning pattern can accidentally miss important information within text if it is not formatted properly, and the information the user

focuses on may vary if the layout of the text changes based on screen size [7]. Users are also less likely to examine content 'below the fold' within a website or interface [4]. This reinforces the idea that user interfaces should place important content within commonly scanned areas so that users do not miss important information.

In addition to visual and cognitive processing considerations, it is also important to consider the physical ergonomics (biomechanic aspects) of the interface as well. According to Fitts' Law, the time required for a user to engage with an object is relative to its size and the distance to it, which implies that usable designs should ensure that targets are appropriately sized and positioned so that users can easily select them [13]. Targets that are small and widely spaced tend to result in longer selection times, and faster movements between targets tend to result in higher error rates [9]. Although this is an important consideration for any interface that requires pointing, Fitts' Law is important to consider when designing a user interface layout on mobile devices where screen real estate is limited[13].

1.4. Impacts of Layout and Content

Previous research has demonstrated that the layout of a web page can influence how users scan text. In experiments undertaken on the differences between one-column and two-column text layouts on a webpage, it was found that although users applied an F-shaped pattern to both layouts, the two-column layout resulted in more content fixations when users were searching the text [14]. Likewise, other studies comparing text-heavy three-column layouts and grid layouts found that users had a greater number of fixations and searched for content faster on the three-column layout than on the grid layout, but that users did not follow a strict F-shaped scan pattern for either layout [15].

It has also been found that users will respond differently to image-based and text-based layouts. In studies comparing text-heavy webpages and a grid layout of image-based content categories, it was found that users responded with the expected F-shaped pattern to the text-heavy content, but that the image-based layout resulted in unique search patterns for each participant [16].

In addition to the differences in how users respond to text and images, it has also been shown that the type of image within an interface can impact user response. Research has shown that users are more likely to focus on images containing faces than on other types of content. This diversion of attention has even been found to affect how long it takes users to find desired information within an interface, with faces above the fold causing a negative impact and faces below the fold improving performance [4].

It has been repeatedly demonstrated that organization and layout can have a significant effect on overall user preference for an interface. An experimental study comparing nine online bookstore websites found that not only did the organization and layout of the webpage have an effect on user preferences for each interface, it had a much greater effect on user preference than aesthetic aspects, such as color and typography [17].

1.5. The Importance of Aesthetics

Although aesthetics may impact the user experience less than layout, it is still an important factorin assessing the perceived usability of an interface. Research found that users tend to correlate the apparent usability of an interface with aesthetic beauty more so than its inherent usability. This indicates that users may be strongly affected by the aesthetic aspect of the interface even when they try to evaluate the interface in its functional aspects [18]. Other research has expanded on these findings, concluding that the degree of system's aesthetics significantly

affected the post-use perceptions of both aesthetics and usability, whereas the degree of actual usability had no such effect [19].

This correlation between the visual beauty of an interface and its perceived usability has been dubbed the aesthetic-usability effect [9]. This effect occurs when the aesthetics of a design create a positive emotional response, which in turn enhances people's cognitive abilities, leading to the perception that the design actually works better [13].

It has been found that the aesthetic-usability effect can be beneficial to aesthetic designs because it can improve users' tolerance to "minor usability issues" in a design [13]. Additionally, aesthetically pleasing designs are more likely to be used than non-aesthetically pleasing designs regardless of their actual usability, and aesthetics can help foster long-term positive attitudes toward a design [9].

Anexperimental study on cell phone interfaces comparing an aesthetically pleasing and a non-aesthetically pleasing design found that not only did the aesthetically pleasing design lead to improvements in perceived usability, but it also led to faster task completion times [20].

However, the aesthetic-usability effect can also have negative impacts on a product or design. More usable, but less-aesthetic, designs may suffer a lack of acceptance that renders issues of usability moot, which in turn leads to future user bias against the product [9]. There can also be negative effects on the overall usability of aesthetically pleasing designs as well. It has been shown that users' improved tolerance for minor usability issues can lead to usability problems potentially being overlooked in crucial usability tests throughout the design process [13].

1.6. Impacts of List and Grid Interface Layouts

List-based and grid-based user interface layouts are commonly found in a variety of user interfaces and are often used to display listings of items with associated information. Research has been previously performed to compareand contrast these user interface layouts.

Experimental studies found that these different layouts affected how efficiently users were able to find items within an e-commerce website, as well as how users perceived the aesthetics of the search results, with the grid layout resulting in faster reaction times and higher values in perceived aesthetics than the list design [21].

Previous research has demonstrated that different layouts help to support different types of browsing behavior. Another experimental study on e-commerce product information layouts found that it benefitted user search efficiency and information recall when the information layout matched with the assigned task. Researchers also discovered that information in a list format was best for browsing products while a grid layout was better suited for searching for specific information [22].

List and grid interface layouts have also been shown to affect how users evaluate each item, as well as how they process the list/grid of items. In an experimental study on how users evaluate the trustworthiness of web search engine results, it was found that users tended to view the items in the list layout more linearly and gave more attention to the first items on the list. The grid layout, however, resulted in users spending a more equal amount of time on each item and resulted in more users correctly selecting the most trustworthy result [23].

1.7. Research on Music Applications

Numerous usability studies have been conducted on music recommendation systems and Music Information Retrieval (MIR) systems, but the majority of these studies have focused on the technical success of these systems, rather than assessing the user experience of an MIR interface [24]. Despite the wide variation in UI designs and the importance of user experience to business success, there has been limited research into mobile music applications[25].

However, the limited existing related research into music app user interfaces has providedsome interesting results. A study on the effects of album art orientation and user swiping directions on app adoption found that atypical designs with album art arranged in a grid fashion led to increased adoption when users were allowed to swipe in a left-to-right direction to interact with the interface [25].

Researchers have also found success in integrating novel features into existing interfaces. An experimental study that implemented a comment and tagging system into an app found that users were able to complete searches more efficiently with the aid of tags [26].

Other research has been conducted using commercially available music services. A heuristic analysis and user interview series of the Spotify mobile and desktop clients and the Pandora desktop client found that users were primarily concerned with issues of design consistency. Different users were likely to follow varied browsing and search patterns. The research also showed that users tended to choose a service that was highly specific to their needs instead of a broadly designed service [24].

Although there is a limited amount of studies in related areas, reviews of existing literature illustrate a lack of user-centered research on mobile music-listening applications. This lack of research combined with the mass adoption of smartphone-based music listening indicates a strong need for further studies of mobile-based usability.

2. METHODOLOGY

This experimental study examined whether grid-based or list-based layouts affect the user experience of a smartphone music application. Experimental participants were tested using a custom-built website that automatically guided them through a series of usability tasks and surveys. Quantitative and qualitative data was collected automatically to address the following research questions:

1. Does the chosen layout influence how quickly users are able to find a target album?
2. Do different layouts result in different numbers of incorrect album selections?
3. Does the layout affect user perceptions of usability factors such as:
    a. Overall ease of use
    b. Speed in finding the target album
    c. Ease of navigation
4. Does the layout affect how users perceive the information in the interface including:
    a. Ease of viewing album art
    b. Ease of viewing album and artist names
    c. The design's overall organization
5. Do different layouts affect user perceptions of:
    a. Overall aesthetic appeal
    b. Desire to use a similar app in everyday life

c. Overall positive impression of the design

2.1. Participants and Recruitment

Participants were recruited through the online social media platforms LinkedIn and Discord. A total of 24 experimental participants were recruited. A pre-screening survey ensured that all participants wereover 18 years of age, fluent in English, and had average or corrected vision. Participants were required to complete the experimental study using a smartphone, this allowed the custom-built testing website to mimic a smartphone application environment.

2.2. Study Procedure

This study was conducted using a custom-built user testing website. The website guided participants through the onboarding steps of the experimental study, the experiment tasks, and the debriefing process was also automated. The website was hosted on a SUNY Oswego server and participant data was collected, encrypted and storedon a secure server.

All album art, album names, and artist names used in this study were AI-generated to avoid any potential existing biases participants may have had due to familiarity with a particularmusic album or artist but also allowed the experiment to maintain a sense of realism. The artwork was created using Copilot Designer (Microsoft), and all artist names and album names were generated using ChatGPT (OpenAI).

Participants accessed the study via a link provided in the online recruitment posts. After consenting to the study, the participants completed a pre-screening survey to confirm that they were 18 or older, fluent in reading English, had average or corrected vision, and were participating in the study using a smartphone. If the participant met these criteria, they were able to move on to the next portion of the study. Potential participants who did not meet these requirements, were automatically redirected to a notice informing them that they have been removed from the study.

The experimental portion of the study followed a within-subjects design, where each participant was tested on their responses to both a list-based and grid-based music application layout. For each layout, the participant completed album search tasks and a post-task survey. The testing order for each layout was randomized to help minimize potential bias.

At the start of each task, the participant was shown instructions to explain that they must memorize an album that will be shown for 15 seconds on the following page and that they will then need to find this album as quickly as possible in the displayed interface. The participants were then shown an album information set (grid or list layout) including the album art, album name, and artist name. The combination of this information was also randomized for each participant and each task to minimize bias.

Each participant had to attempt to find and select the memorized album as quickly as possible. During the task, the website automatically tracked how long it took the participant to find the correct album, as well as how many incorrect album selections were made. If an incorrect album was selected, the participant was notified via a popup screen which prompted them to continue trying until the correct album was found, and the website paused tracking the task completion time until the popup was dismissed.

The other albums displayed in the interface were each created from randomized combinations of album art, album names, and artist names, but all album combinations were unique. A total of 16

albums were displayed to each participant. The targeted album was ordered 10th in the series in each experimental test to minimize potential timing biases (based on Fitts' Law) caused by the participant having to navigate to widely varying positions in the interface.

When the participant selected the correct album in the interface, they were notified that they had successfully completed the task. They then were then presented with a post-task survey consisting of nine Likert scale questions about their perceptions of the usability and organization of the interface, as well as their overall attitudes towards the interface.

The task and post-task survey procedure then were repeated for the other interface layout (grid or list), ensuring all participants complete a task and post-task survey for both layout variations.

2.3. Interface Design

The custom-built experimental website was designed primarily for users accessing the study via a mobile device using touch-based controls. Figure 2 shows an example of the website in use creating both grid and list layouts of AI-generated albums and artists.
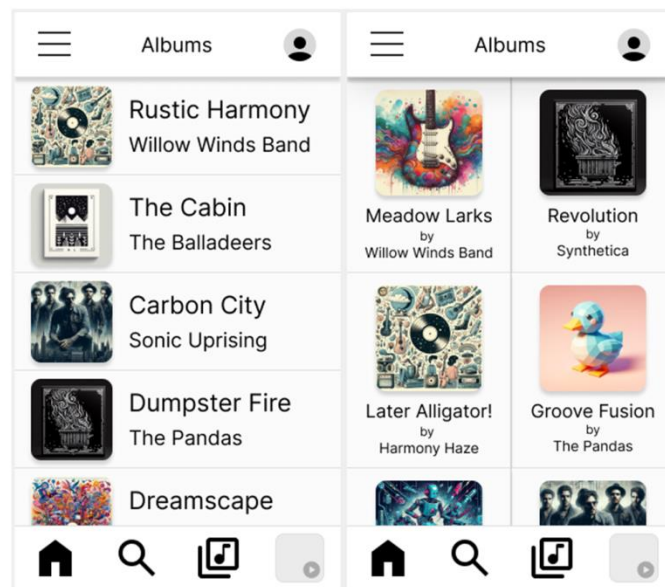

Figure 2. List and grid layouts of AI-generated albums and artists.

Initially a prototype of the music application website was created using Figma (Figma Inc.). The final experimental website was created using custom HTML, CSS, and Javascript, and the Svelte and Typescript web frameworks.

Participant data was stored locally in the participant's browser using JavaScript as they progressed through the experiment, and it was only transmitted to the research team if the participant completed the experiment. Upon submission, the experiment data was encrypted and stored securely using the Google Apps Script platform.

3. RESULTS

A total of 24 valid responses were collected during the experimental study. However, an initial examination of the differences in task completion time between the layout versions found a

couple of potentially extreme outliers that could affect the final analysis. An boxplot of this metric showed two potential extreme outliers, one in the positive, and one in the negative direction (Figure 3). The interquartile method was applied to determine if these outliers could be considered extreme. Thetwo outliers were both determined to fall outside of the interquartile range, and they were removed from the dataset for this study. This left a total of 22 valid responses for further analysis.
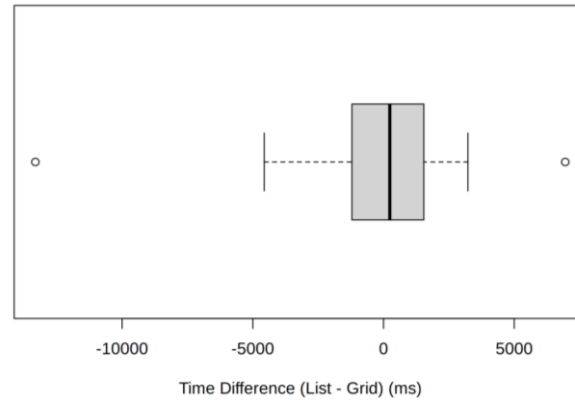

Figure 3. Box plot of task completion time differences before outlier removal

During the study, the custom-built testing website was used to automatically track any album selection errors made by participants during the album search task. The participants in this study did not make any erroneous selections using either layout, so no analysis of errors was undertaken.

3.1. Task Completion Times

Task completion times were tracked automatically by the testing website.

- The completion times for the list layout ranged from 2227 milliseconds to 7888 milliseconds, with a mean of 5019 milliseconds (Table 1, Figure 4).
- For the grid layout, completion times ranged from 2115 milliseconds to 16591 milliseconds, with a mean of 5278 milliseconds (Table 1, Figure 5).

When subtracting the grid layout completion times from the list layout completion times for each participant, the differences ranged from 13318 milliseconds to 6952 milliseconds, with a mean difference of 101.3 milliseconds (Table 1, Figure 6).

Table 1. Summary of task completion times and time differences in milliseconds (ms)

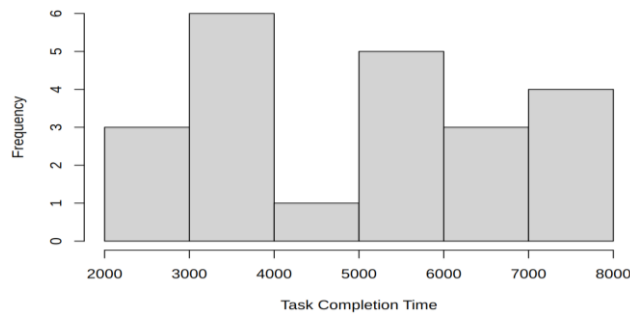| Metric	Min	Q1	Median	Mean	Q3	Max
List Time	2227	3627	5155	5177	6394	10560
Grid Time	2115	3319	4332	5278	5856	16591
Time Difference	13318.0	1187.0	240.0	101.3	1286.5	6952.0

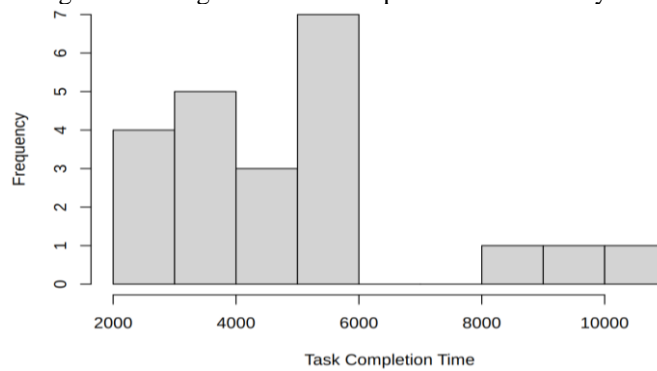Figure 4. Histogram of task completion times: list layout

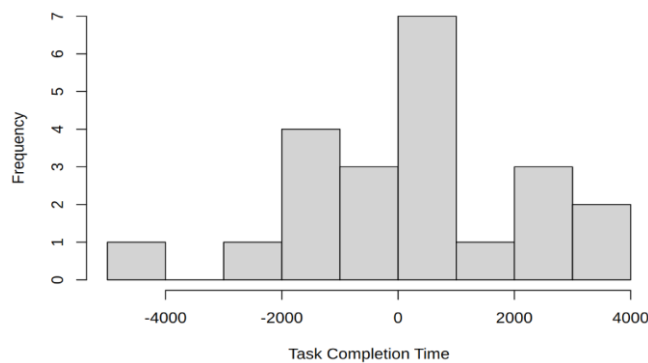Figure 5. Histogram of task completion times: gridlayout

Figure 6. Histogram of task completion time differences (list and grid)

3.2. Qualitative Survey

For each layout, the experimental participants were asked to answer the following nine qualitative survey questions using a seven-point Likert scale ranging from 1 (Strongly Disagree) to 7 (Strongly Agree) :

- Q1: Overall, this task was easy to complete.
- Q2: I was able to find the correct album quickly.
- Q3: The design was easy to navigate.
- Q4: It was easy to see the album art for each album.
- Q5: It was easy to read the album name and artist for each album.
- Q6: Overall, I think the design was well organized.
- Q7: I think the design was pleasing to look at.

- Q8: I would like to use a music app with a similar layout in my everyday life.
- Q9: Overall, I liked using this design.

The 22 experimental participants each completed nine questions per interface layout, which resulted in 198 individual question responses per layout (396 in total). A tabulation of the scores received across all questions per layout shows that the results were mostly positive, with most responses ranking 5 or above on the seven-point scale (Table 2, Figure 7).

The grid interface variant received the highest number of *Strongly Agree* responses (133) compared to the list variant (119). Interestingly, the grid variant also received the highest number of *Strongly Disagree* responses (4), compared to the list variant (0).

Table 2. Total frequency of score responses across all questions by layout

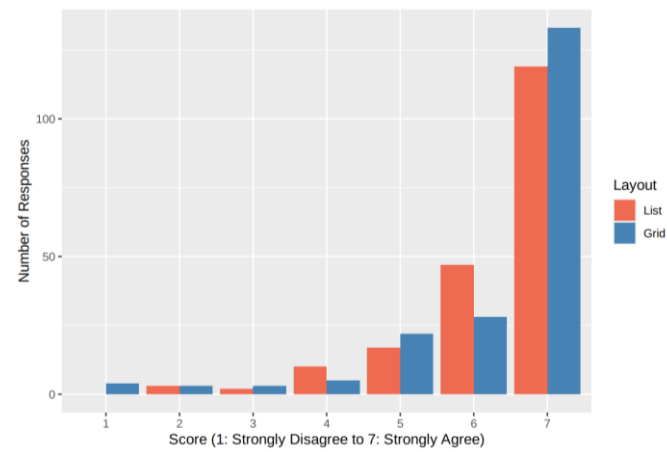| Layout	1	2	3	4	5	6	7
List	0	3	2	10	17	47	119
Grid	4	3	3	5	22	28	133

Figure 7. Score responses across all qualitative survey questions by interface layout

Further examination of the responses for each individual question per layout found that for the list layout, questions five questions received a negative response below a 4 ranking (Table 3). The grid layout also received negative scores for five questions, but one question received two negative scores (Q7), and one receiving three negative scores (Q8) (Table 4).

Table 3. List layout survey responses

Question	1	2	3	4	5	6	7
Q1	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	8 (36%)	14 (64%)
Q2	0 (0%)	0 (0%)	0 (0%)	1 (4.5%)	3 (14%)	6 (27%)	12 (55%)
Q3	0 (0%)	0 (0%)	1 (4.5%)	0 (0%)	0 (0%)	6 (27%)	15 (68%)
Q4	0 (0%)	1 (4.5%)	0 (0%)	1 (4.5%)	1 (4.5%)	3 (14%)	16 (73%)
Q5	0 (0%)	0 (0%)	0 (0%)	0 (0%)	4 (18%)	4 (18%)	14 (64%)
Q6	0 (0%)	0 (0%)	1 (4.5%)	0 (0%)	2 (9.1%)	5 (23%)	14 (64%)
Q7	0 (0%)	0 (0%)	0 (0%)	3 (14%)	2 (9.1%)	3 (14%)	14 (64%)
Q8	0 (0%)	1 (4.5%)	0 (0%)	3 (14%)	3 (14%)	5 (23%)	10 (45%)
Q9	0 (0%)	1 (4.5%)	0 (0%)	2 (9.1%)	2 (9.1%)	7 (32%)	10 (45%)

A plot of the score frequencies for each question shows that the question with the overall highest number of *Strongly Agree* responses was question 2 (I was able to find the correct album quickly) when participants used the grid layout (Figure 8). The question with the most *Strongly Disagree* responses was question 8 (I would like to use a music app with a similar layout in my everyday life).

Table 4. Grid layout survey responses

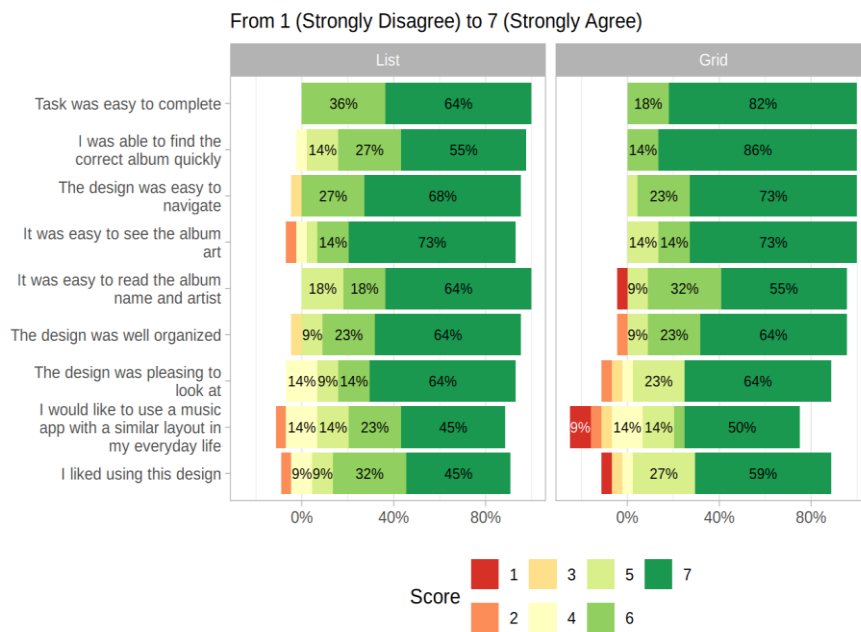| Question	1	2	3	4	5	6	7
Q1	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	4 (18%)	18 (82%)
Q2	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (14%)	19 (86%)
Q3	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (4.5%)	5 (23%)	16 (73%)
Q4	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (14%)	3 (14%)	16 (73%)
Q5	1 (4.5%)	0 (0%)	0 (0%)	0 (0%)	2 (9.1%)	7 (32%)	12 (55%)
Q6	0 (0%)	1 (4.5%)	0 (0%)	0 (0%)	2 (9.1%)	5 (23%)	14 (64%)
Q7	0 (0%)	1 (4.5%)	1 (4.5%)	1 (4.5%)	5 (23%)	0 (0%)	14 (64%)
Q8	2 (9.1%)	1 (4.5%)	1 (4.5%)	3 (14%)	3 (14%)	1 (4.5%)	11 (55%)
Q9	1 (4.5%)	0 (0%)	1 (4.5%)	1 (4.5%)	6 (27%)	0 (0%)	13 (59%)

Figure 8. Survey question responses by layout

In preparation for further analysis using a paired samples methodology, the differences between the question response scores were also analyzed. In a similar manner to the analysis undertaken with the task completion time data, the grid layout scores were subtracted from the list layout scores. The results for each question were plotted using a series of histograms as shown in Figure 9.
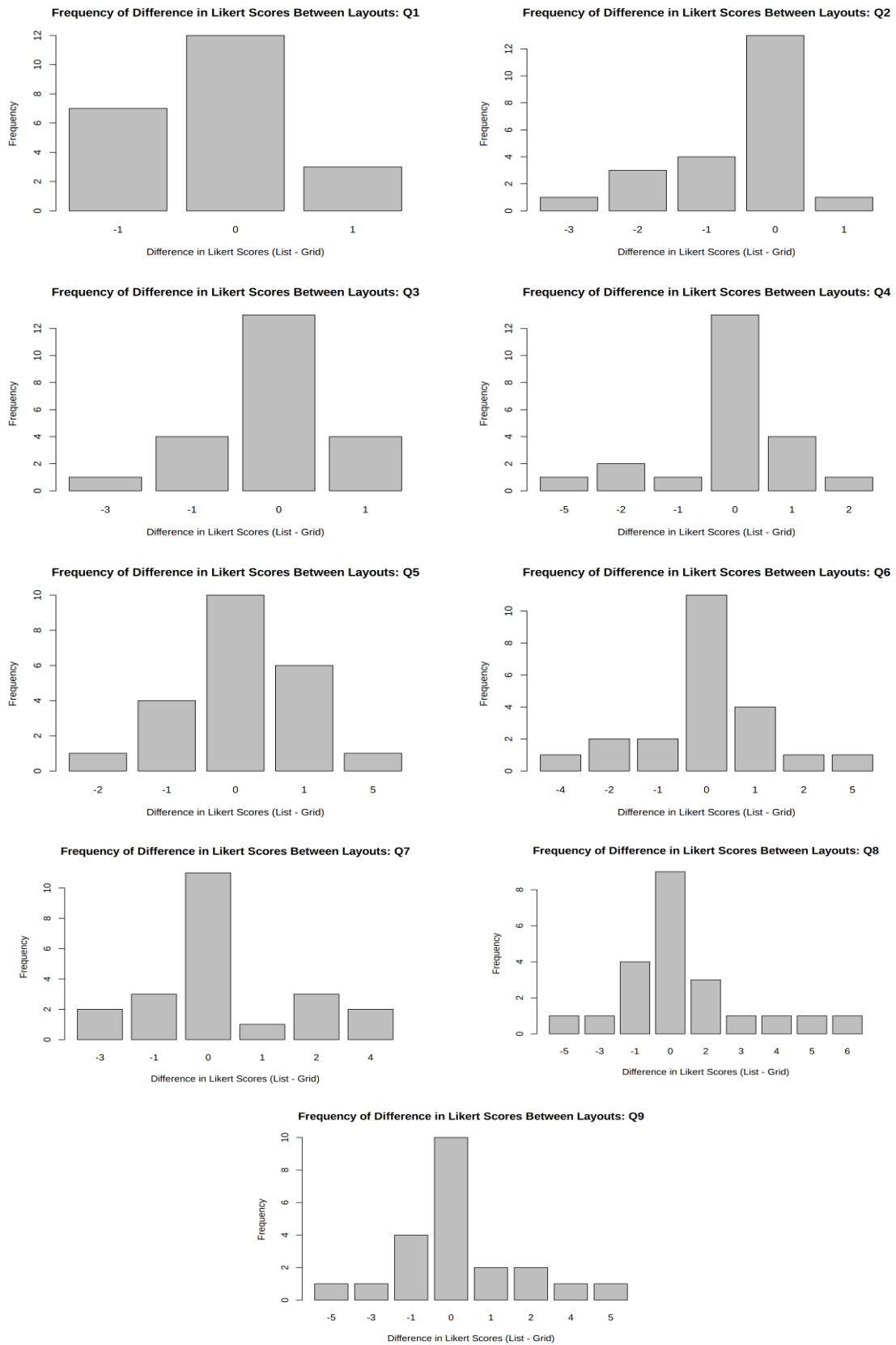
Figure 9. Survey score differences; Questions 1-9

For all questions, the majority of responses yielded a difference of 0, but all questions had differences in both positive and negative directions. For question 1, the differences ranged between positive and negative 1 (Figure 8). In questions 2 and 3, the differences ranged from -3 to +1 (Figure 9). All other questions had much larger ranges, with question 8 having the largest at a range of -5 to +6.

4. ANALYSIS

The analysis of the findings from the experimental study explore the relationship between grid-based and list-based layouts in smartphone music applications. This section examines both quantitative data, such as task completion times and qualitative user perceptions, as captured through survey responses.

Through statistical testing and correlation analyses, the study identifies patterns and anomalies, providing insights into the impact of layout design on user experience. Additionally, the analysis highlights the psychological factors that may influence users' perceptions, offering a deeper understanding of usability considerations in music app interfaces. By integrating these findings, this section aims to bridge the gap between empirical results and broader usability implications.

4.1. Task Completion Time Analysis

Given the relatively small sample size for this study, a Shapiro-Wilk test was chosen to test the normality of the data for the task completion time differences. The data passed the test for normality, as the results failed to reject the null hypothesis that the sample comes from a normally distributed population. The Shapiro-Wilk test statistic that measures how closely the data matches a normal distribution, $W(22) = 0.96$, with a significance level of $p = 0.45$. This indicated that the data was suitable for further analysis.

Following the Shapiro-Wilk test, a two-tailed paired t-test was performed to compare the album search task completion times between the list-based and grid-based layouts. The test results indicated that there was not a significant difference in the task completion times for the list-based layout ($M = 5177$, $SD = 1703.27$) and the grid-based layout ($M = 5278$, $SD = 2216.97$); $t(21) = 0.456$, $p = 0.65$.

Although there was not a significant difference between the task completion times for each layout, further exploration of this relationship was made possible through testing for any existing correlation between these factors. A Pearson correlation test was used to determine if a linear correlation exists between the task completion time for the list layout and for the grid layout. The results of the test indicate that there is a significant positive correlation between these variables $r(20) = 0.59$, $p = 0.004$(Figure 10).

This indicates that both layouts allow users to complete the album search task in approximately the same amount of time, and there is not a significant difference quantitatively in this facet of user experience.

4.2. Qualitative Survey Analysis

Given that the results from the qualitative survey are a form of ordinal data, a two-tailed paired Wilcoxon signed-rank test was performed for each question to analyze the differences in scores between each layout variation (Table 5). For all 9 questions except question 2, the results were

nonsignificant and failed to reject the null hypothesis, indicating that there is likely not a significant difference between the scores for each layout variation (Table 5).
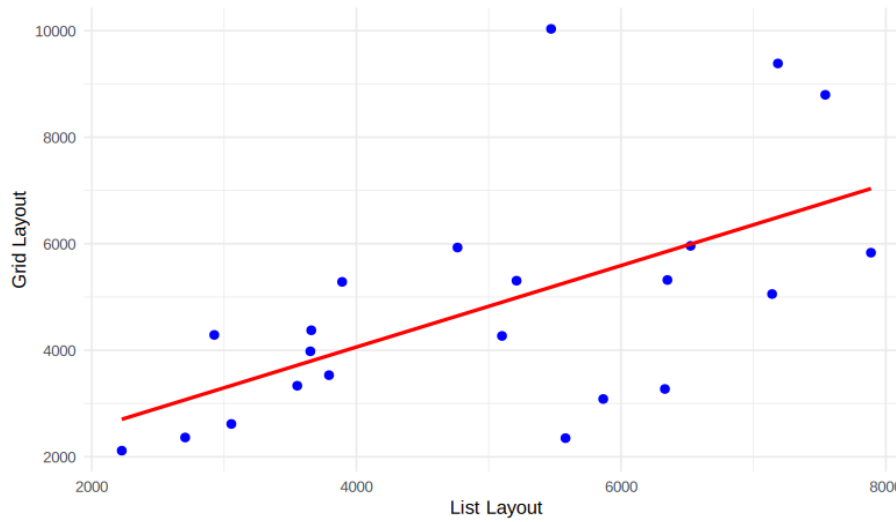

Figure 10. Task completion time - Pearson correlation

However, for question 2 ("I was able to find the correct album quickly."), the results indicated a significant difference between the scores for the list-based layout and the grid-based layout, W = 3.0, p = 0.0215. For this question, both the list and grid layouts received a median score of 7, but the list variant had a mean score of 6.31 and the grid variant had a mean score of 6.86.

The effect size for the Wilcoxon signed-rank test for each of the 9 questions was also calculated (Table 5). Question 2 resulted in a large effect size of r = 0.51. All other questions resulted in small effect sizes.

This indicates that there was a strong significant difference between the layouts in how users perceived the time taken to complete the task, despite there not being an actual difference in the time taken to complete the task.

Table 5. Wilcoxon signed-rank test results by question

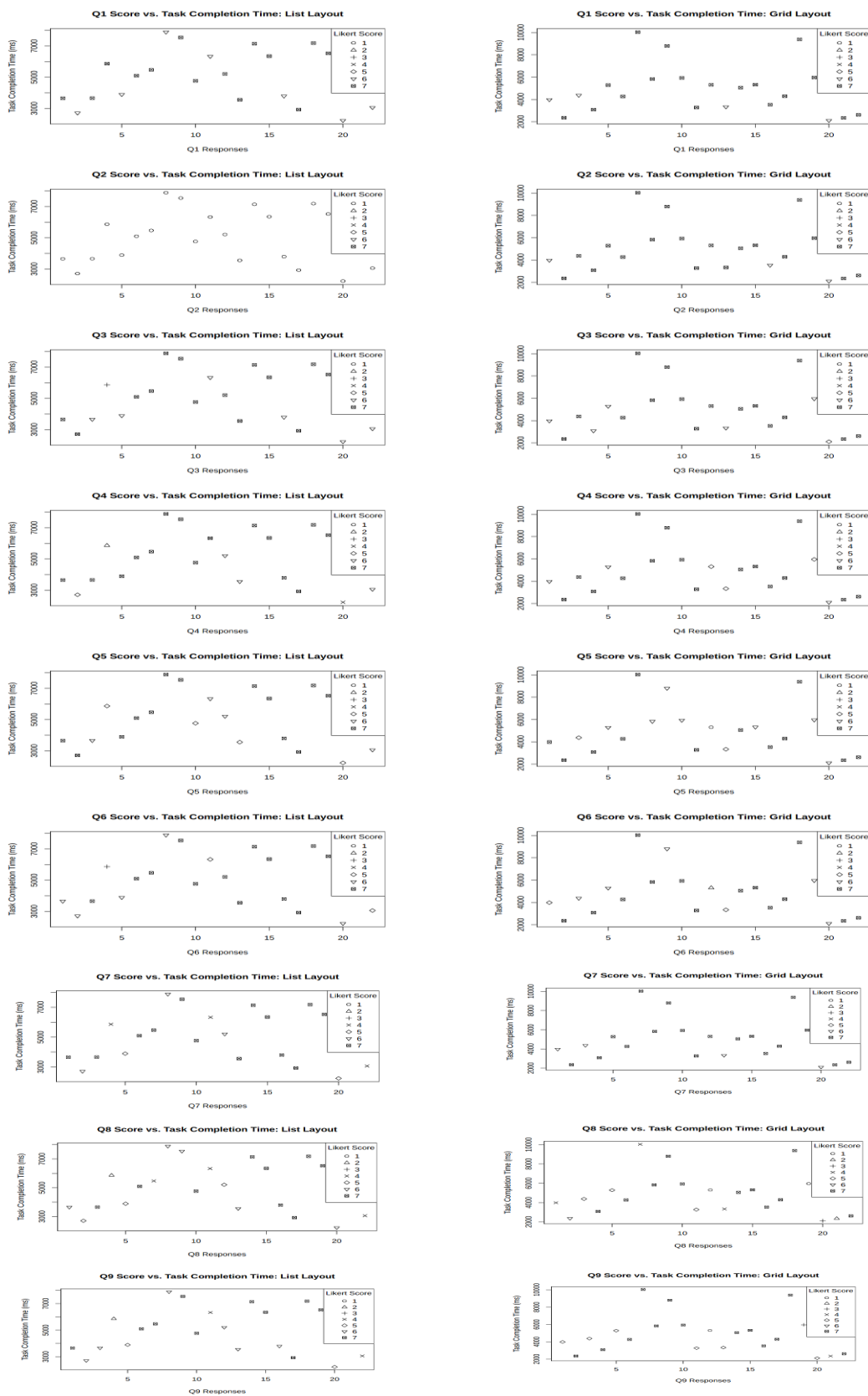| Question	W	p	r
Q1	16.5	.227	.27
Q2	3.0	.0215	.51
Q3	18.0	.608	.09
Q4	19.0	.717	.02
Q5	45.0	.644	.12
Q6	33.0	1.000	.04
Q7	41.5	0.473	.10
Q8	61.0	.292	.12
Q9	43.5	.750	.03

Figure 11. Correlation of question score and task completion time: Q1 -Q9; list and grid layout

4.3. Correlations Between Task Completion Time and Survey Responses

Given that there was not a significant difference in task completion times between the layouts, but there was a significant difference in user perceptions surrounding the time taken to complete the task reported via question 2, it is important to determine whether this difference in perception was influenced by the actual time taken to complete the task. To explore whether task completion time influenced participant responses to question 2, a correlation analysis was conducted to further explore this relationship.

Additionally, correlation testing was conducted between task completion times and the other 8 survey question responses to determine whether a relationship exists despite those survey questions providing nonsignificant results during the Wilcoxon signed-rank test.

A Spearman correlation test was conducted for all 9 survey questions (Figure 11) for both the list layout (Table 6) and the grid layout (Table 7). Question 4 (It was easy to see the album art for each album) had a significantly positive correlation with task completion time in the list layout, $r(20) = 0.46$, $p = 0.030$, but no significant relationship was found for this question for the grid-based layout $r(20) = 0.03$, $p = 0.906$. All other questions were not significantly correlated with task completion time for either layout.

Table 6. Spearman correlations by question: List layout

Question	Spearman Correlation Coefficient	P-Value
Q1	0.24	.286
Q2	0.18	.431
Q3	0.28	.205
Q4	0.46	.030
Q5	0.34	.120
Q6	0.20	.362
Q7	0.16	.470
Q8	0.137	.541
Q9	0.30	0.18

Table 7. Spearman correlations by question: Gridlayout

Question	Spearman Correlation Coefficient	P-Value
Q1	0.32	.15
Q2	0.34	.116
Q3	0.26	.244
Q4	0.03	.906
Q5	-0.29	.19
Q6	-0.05	.813
Q7	0.09	.685
Q8	0.16	.479
Q9	0.20	.31

This indicates that task completion time does not have a significant correlation with how participants perceived the time they took to complete the task. Additionally, it also shows that there does appear to be a relationship between task completion times and participant perceptions of the album art, but only for the list-based layout.

5. DISCUSSION

The results of this study show that within the context of searching for an album, list-based and grid-based do not appear to have a significant effect on how quickly users can find an album. However, it does appear that the layout variant had a significant effect on how users perceived the time it took for them to find an album.

The majority of users felt that both designs allowed them to find the correct album quickly (Question 2), but respondents showed a significantly more positive response to the grid layout in comparison to the list layout.

The responses showed that 82% of respondents strongly felt that they were able to find the album quickly in the grid layout (score of 7), with the remaining participants assigning a score of 6 on the Likert scale. In comparison, the list layout had a much wider range of responses, with only 55% of respondents strongly felt that they were able to find the album quickly in the grid layout (score of 7), and responses ranged down to a neutral 4 (Figure 8).

Interestingly, the Spearman correlation test found that there was no significant correlation between users' responses to Question 2 and the actual time taken to complete the task for either layout (Table 6and Table 7). However, a Pearson correlation test and paired t-test showed that there was not a significant difference between the actual time it took participants to complete the task between the layouts.

This indicates that there exists a discrepancy between how participants perceived the task completion times and the actual task completion times for each layout. This difference was likely due to a range of psychological factors,rather than how long it actually took the participants to complete the task. It appears that the grid layout provided better usability in this instance from a psychological standpoint, however the precise cause of this discrepancy needs to be researched further.

The experimental data showed that there are no other significant differences in the other aspects of user experience between list-based and grid-based layouts. Both layouts resulted in a lack of user errors, and although this was not possible to analyze via statistics, this is positive anecdotal evidence for the overall usability of both designs. Likewise, survey questions 1 and 3 both concerned aspects of ease of use, and both layouts received an approximately equally positive response.

For aspects of organization and legibility, questions 4, 5, and 6 did not see a significant difference between the layout variations. Both layouts again received mostly positive scores for each of these questions. This indicates that both layouts likely provided approximately equal user experiences in terms of browsing and consuming information.

Interestingly, the responses for question 4 for the list layout (It was easy to see the album art for each album) and question 5 for the grid layout (It was easy to read the album name and artist for each album) received a few highly negative responses. Additionally, question 4 had a significant positive correlation with task completion time, but only for the list-based layout.

This may suggestthat list layouts may make it more difficult to view image-based information and grid layouts may make it more difficult to view textual information for some users. However, additional research is needed to make any substantive claims on this.

In terms of aesthetics and user preference, most participants found the designs aesthetically pleasing in question 7 and showed a favorable view of using the design both in their everyday lives and as an overall experience in questions 8 and 9.

Although there was not a statistically significant difference between the scores for these questions between the layout variants, it is interesting to note that 18.2% of respondents indicated that they disagreed with the idea of using the grid-based layout in everyday life, compared to 4.5% of respondents for the list-based layout.

This could potentially indicate that many users have approximately the same opinions about these two layouts, but that some have much stronger preferences. However, there is not enough evidence to draw a definitive conclusion as to whether this is the case.

6. CONCLUSIONS

As mobile devices have become a ubiquitous part of everyday life and have taken on additional functions, music apps have become a major avenue for users to curate their personal music collections. Although research has found evidence for differences between list-based and grid-based layouts in other user interface contexts, there has been a lack of research into how different layouts affect the user experience of music apps.

This study examined whether grid and list layouts affected user experience in music apps by analyzing both quantitative usability metrics and qualitative user perceptions. The results indicate that although users did not take a significantly different amount of time to find a desired album, they felt more strongly that they could complete the task faster when using the grid-based layout. Additionally, it was found that there was no correlation between this reported perception of task completion time and the actual time taken to complete the task for either layout.

The results of the research have practical implications for UI designers and developers of smartphone applications. By identifying user preferences for certain layouts and understanding their perceptions of usability, developers can create interfaces that enhance user satisfaction and engagement. For example, the finding that grid layouts are perceived as faster for task completion (even if the actual time taken isn't significantly different) suggests a psychological advantage that developers can exploit. These insights can be applied to improve not only music apps but also other applications that rely heavily on visual representation and navigation.

Generalizing the methodology used in this study could prove immensely useful for evaluating and designing UI across various app categories. By adopting a similar experimental framework (leveraging controlled tasks, randomized testing orders, and the integration of both qualitative and quantitative measures) developers can assess the usability of alternative layouts in other contexts. The approach could be extended to e-commerce platforms, educational apps, or social media interfaces, where grid and list layouts are also common. This would allow for more user-centered design practices, ensuring interfaces cater to diverse needs while maintaining aesthetic appeal and functional efficiency.

This paper makes a noteworthy contribution to the field of human-computer interaction by shedding light on the usability of list-based versus grid-based album layouts in smartphone music applications. It provides valuable insights into user experience design, emphasizing the interplay between visual perception and interface layout preferences. While the study confirms that task completion times do not vary significantly between layouts, it highlights an intriguing psychological perception where users feel grid layouts are faster, despite the data showing otherwise. By exploring these nuanced user perceptions, the paper extends the understanding of

how layout aesthetics and usability intersect. Moreover, the research methodology offers a replicable framework for evaluating UI design in other application domains, making this study not only impactful for music apps but also broadly relevant across different types of interfaces.

Additional work is needed to determine the precise cause of this difference in actual and perceived completion times. Previous research described a connection between grid layouts and increased perceptions of aesthetics [21] which may indicate that other studies should explore whether the aesthetic-usability effect is a possible cause of this difference.

Incorporating more statistically grounded methods like A/B testing could substantially enhance the robustness of this study. A/B testing would allow researchers to directly compare the performance of two interface layouts (grid and list) across a larger and more diverse pool of participants. By randomly assigning participants to interact with only one layout variant, A/B testing eliminates potential biases that arise in within-subject designs, such as carryover effects or learning from prior exposure. This approach could provide clearer insights into the distinct usability benefits of each layout. Additionally, A/B testing would enable researchers to perform more sophisticated statistical analyses to detect subtle differences in behavior and preferences, such as multivariate regression or causal inference methods.

Another possible avenue of research could include further investigation of how visual search patterns and time perception are affected in this context. Studies have shown that highly visual grid-based designs can result in altered search patterns in comparison to the expected F-shaped scan pattern [16], and that saccadic eye movement can cause illusory distortions to time perceptions via saccadic chronostasis[27]. If grid-based interfaces cause a different saccade pattern than list-based interfaces, perhaps there could be a different cumulative effect on users' overall time perceptions when using a grid-based interface.

REFERENCES

[1] Anderson, M. (2016) "More Americans Using Smartphones for Getting Directions, Streaming TV", Pew Research Center, Streaming TV, https://www.pewresearch.org/short-reads/2016/01/29/us-smartphone-use/

[2] Smith, A. (2015) "U.S. Smartphone Use in 2015", Pew Research Center: Internet, Science & Tech,https://www.pewresearch.org/internet/2015/04/01/us-smartphone-use-in-2015/

[3] Schedl, M., Zamani, H., Chen, C., Deldjoo, Y., & Elahi, M. (2018) "Current Challenges and Visions in Music Recommender Systems Research", International Journal of Multimedia Information Retrieval, 7(2), 95-116.

[4] Romano, B. J., & Schall, A. (Eds.). (2014) "Eye Tracking in User Experience Design", Elsevier Science & Technology.

[5] Ware, C. (2012) "Information Visualization: Perception for Design", Elsevier Science & Tech.

[6] Malamed, C. (2015) "Visual Design Solutions : Principles and Creative Inspiration for Learning Professionals", John Wiley & Sons, Incorporated.

[7] Pernice, K. (2017) "F-Shaped Pattern of Reading on the Web: Misunderstood, but Still Relevant (Even on Mobile)", Nielsen Norman Group, https://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/

[8] Davis, M., & Hunt, J. (2017) "Visual Communication Design : An Introduction to Design Concepts in Everyday Experience", Bloomsbury Publishing USA.

[9] Lidwell, W., Holden, K., & Butler, J. (2010) "Universal Principles of Design, Revised and Updated : 125 Ways to Enhance Usability, Influence Perception, Increase Appeal, Make Better Design Decisions, and Teach Through Design", Quarto Publishing Group USA.

[10] Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012) "A Century of Gestalt Psychology in Visual Perception: I Perceptual Grouping and Figure–Ground Organization", Psychological Bulletin, 138(6), 1172–1217.

[11] Still, J. D. (2018) "Web Page Visual Hierarchy: Examining Faraday's Guidelines for Entry Points", Computers in Human Behavior, 84, 352–359.

[12] Rosenholtz, R., Twarog, N. R., Schinkel-Bielefeld, N., & Wattenberg, M. (2009) "An Intuitive Model of Perceptual Grouping for HCI Design", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1331–1340.

[13] Yablonski, J. (2020),"Laws of UX", O'Reilly Media Inc.

[14] Shrestha, S., & Owens, J. (2008) "Eye Movements on a Single-Column and Double-Column Text Web Page", Usability News, 10(1), 1-8.

[15] Namoun, A. (2018) "Three Column Website Layout vs. Grid Website Layout: An Eye Tracking Study", Design, User Experience, and Usability: Designing Interactions, 271–284.

[16] Schresta, S., & Lenz, K. (2007), "Eye Gaze Patterns while Searching vs. Browsing a Website", Usability News, 9(1), 1-10.

[17] Lee, S., & Koubek, R. J. (2010),"The Effects of Usability and Web Design Attributes on User Preference for E-Commerce Web Sites", Computers in Industry, 61(4), 329–341.

[18] Kurosu, M., &Kashimura, K. (1995), "Apparent Usability vs. Inherent Usability",Proceedings of Conference Companion on Human Factors in Computing Systems - CHI '95, 292–293.

[19] Tractinsky, N., Katz, A. S., & Ikar, D. (2000), "What is Beautiful is Usable", Interacting with Computers, 13(2), 127–145.

[20] Sonderegger, A., & Sauer, J. (2010), "The Influence of Design Aesthetics in Usability Testing: Effects on User Performance and Perceived Usability", Applied Ergonomics, 41(3), 403–410.

[21] Schröder, F., Kleinjan, A., & Brandenburg, S. (2018), "he Effects of Grid- and List design of E-Commerce Result Lists on Search Efficiency and Perceived Aesthetics", Advances in Usability, User Experience and Assistive Technology, 794, 204–211.

[22] Hong, W., Thong, J. Y. L., & Tam, K. Y. (2004), "The Effects of Information Format and Shopping Task on Consumers' Online Shopping Behavior: A Cognitive Fit Perspective", Journal of Management Information Systems, 21(3), 149–184.

[23] Kammerer, Y., & Gerjets, P. (2010), "How the Interface Design Influences Users' Spontaneous Trustworthiness Evaluations of Web Search Results", Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10, 299–306.

[24] Lee, J. H., & Price, R. (2016), "User Experience with Commercial Music Services: An Empirical Exploration", Journal of the Association for Information Science & Technology, 67(4), 800–811.

[25] Rushan, M. R. I. (2021), "The Way You Listen to Music: Effect of Swiping Direction and Album Arts on Adoption of Music Streaming Applications", Behaviour & Information Technology, 40(16), 1806–1827.

[26] Chen, L., Liu, Q., Zhang, C., Huang, Y., Peng, Z., Zeng, H., Sun, Z., Ma, X., & Li, Q. (2024), "Amplifying the Music Listening Experience Through Song Comments on Music Streaming Platforms", Journal of Visualization, 27(3), 401–419.

[27] Yarrow, K., Haggard, P., & Rothwell, J. C. (2010), "Saccadic Chronostasis and the Continuity of Subjective Temporal Experience Across Eye Movements", Space and Time in Perception and Action, 149–163.

AUTHORS

Molly Scott is a recent graduate from the Human-Computer Interaction M.A. program at SUNY Oswego with a background in robotics engineering. Before attending SUNY Oswego, Molly contributed to research at the University of Alabama Huntsville by designing and developing a real-time dashboard website for a TOLNet atmospheric research vehicle. Molly's current research interests include the interactions between user perceptions and usability, and how users process information displayed in various user interfaces.

Damian Schofield is currently the Director of the HCI masters program at the State University of New York (Oswego). Prior to his move to America, Dr. Schofield was the Associate Professor of Computer Games and Digital Media, in the School of Creative Media at RMIT University in Melbourne, Australia. Before his move out to Australia, he was Lecturer in the School of Computer Science at the University of Nottingham, UK. Dr. Schofield is specifically interested in the representation and understanding of visual information in multiple digital environments.