U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

MESTRADO EM CIÊNCIA DA INFORMAÇÃO

Critical Analysis and Evaluation of the User Experience in the ZeroZero's Mobile Version Case Study at ZeroZero

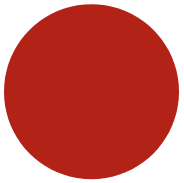
Marta Pinhal

M

2025

UNIDADES ORGÂNICAS ENVOLVIDAS

FACULDADE DE ENGENHARIA
FACULDADE DE LETRAS



FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO
FACULDADE DE LETRAS DA UNIVERSIDADE DO PORTO

Critical Analysis and Evaluation of the User Experience in the ZeroZero's Mobile Version Case Study at ZeroZero

Marta Garranas Do Novo Pinhal

Mestrado em Ciência da Informação

Supervisor: Prof. Alexandre Valle de Carvalho

July 25, 2025

Critical Analysis and Evaluation of the User Experience in the ZeroZero's Mobile Version Case Study at ZeroZero

Marta Garranas Do Novo Pinhal

Mestrado em Ciência da Informação

Approved in oral examination by the committee:

President: Prof. Teresa Silveira
Referee: Prof. Bruno Giesteira
Referee: Prof. Alexandre Valle de Carvalho

July 25, 2025

Resumo

A experiência do utilizador tem emergido como um fator crítico para o sucesso de plataformas digitais, sobretudo na transição de websites para versões otimizadas para dispositivos móveis. Esta dissertação explora os desafios e oportunidades desta transformação na plataforma Zos, a maior base de dados desportiva do mundo, recorrendo a um estudo de caso.

Recorrendo a uma revisão sistemática da literatura, foram identificadas lacunas e tendências nos métodos, métricas e instrumentos utilizados na avaliação da UX, com especial enfoque em plataformas intensivas em dados. A metodologia adotada combina questionários (*System Usability Scale – SUS*), entrevistas estruturadas e análise comportamental com recurso a ferramentas como o Hotjar, permitindo uma avaliação multidimensional da navegação, usabilidade e satisfação dos utilizadores.

O estudo envolveu mais de 140 participantes, abrangendo tanto utilizadores registados como não registados. Os resultados revelam que, apesar de a usabilidade global ter sido positivamente avaliada, persistem problemas recorrentes em áreas como a navegação, desempenho, intrusividade da publicidade e eficácia da funcionalidade de pesquisa. Com base nestes dados, são apresentadas recomendações práticas para a melhoria da experiência móvel, abrangendo desde a simplificação do layout e otimização do desempenho até à personalização da interface e expansão da cobertura estatística.

Este trabalho contribui para o conhecimento académico em torno da avaliação da UX em contextos móveis e oferece orientações práticas de design para plataformas digitais complexas e centradas em dados.

Palavras-Chave: Experiência do Utilizador; Usabilidade; Questionários SUS; Entrevistas Estruturadas; Hotjar; Otimização móvel; Métodos de avaliação da experiência do utilizador

Abstract

UX has emerged as a critical success factor for digital platforms, particularly in the transition of websites to mobile-optimized versions. This dissertation investigates the challenges and opportunities of this transformation in Zos, the world's largest sports database, through a case study approach.

Using a systematic literature review, gaps and trends in UX evaluation methods, metrics, and tools were identified, with a focus on data-intensive platforms. The methodology combines questionnaires (System Usability Scale – SUS), structured interviews, and behavioral analysis using tools like Hotjar, providing a multidimensional assessment of navigation, usability, and user satisfaction.

The study engaged over 140 participants across both registered and non-registered user groups. Findings reveal that while overall usability was positively rated, recurring issues were identified in areas such as navigation, performance, advertising intrusiveness, and search functionality. Based on these insights, the dissertation presents a set of actionable recommendations to improve the mobile user experience, ranging from layout simplification and performance optimization to enhanced personalization and expanded statistical coverage.

This work contributes to the academic discourse on UX evaluation in mobile contexts and offers practical design guidance for complex, data-driven digital platforms.

Keywords: User Experience; Usability; SUS questionnaires; Structured Interviews; Hotjar; Mobile Optimization; UX Evaluation Methods

United Nations Sustainable Development Goals

This dissertation, focused on the evaluation and improvement of the user experience (UX) of ZeroZero's mobile platform, contributes to several of the United Nations Sustainable Development Goals (SDGs). These goals aim to address global challenges such as access to information, digital inclusion, innovation, and equality.

Identified SDGs

The following SDGs are directly or indirectly aligned with the scope and objectives of this research:

- **SDG 4 – Quality Education**: By improving the accessibility and usability of a major digital platform, this project enhances users' access to reliable sports information, fostering digital literacy and informal learning.
- **SDG 9 – Industry, Innovation, and Infrastructure**: The methodological and technical improvements proposed contribute to innovation in mobile UX and to the development of more resilient digital infrastructure.
- **SDG 10 – Reduced Inequalities**: Enhancing accessibility and usability in mobile platforms can reduce digital inequality, particularly among users with different devices and technical skills.
- **SDG 17 – Partnerships for the Goals**: This research was conducted in collaboration with a private sector organization (ZeroZero), reflecting the importance of partnerships between academia and industry in promoting innovation and sustainable development.

Specific Targets and Contribution

Within the identified **SDGs**, specific targets have been considered. These include Target 4.4 (increase relevant skills for employment), Target 9.5 (enhance scientific research and upgrade the technological capabilities of industrial sectors), Target 10.2 (empower and promote the social, economic, and political inclusion of all), and Target 17.17 (encourage and promote effective public, public-private and civil society partnerships).

This dissertation contributes by proposing improvements that support inclusive and user-centered design, reinforcing the value of data-driven decision-making and collaborative innovation.

Performance Indicators and Metrics

The table below summarizes the SDGs addressed, the specific contribution of the project, and potential performance indicators:

SDG	**Goal**	**Contribution**	**Performance Indicators and Metrics**
SDG 4	Quality Education	Enhancing access to structured sports information and promoting digital literacy through improved UX design.	Average time spent on platform; user satisfaction scores (SUS); increased use among students or younger users.
SDG 9	Industry, Innovation and Infrastructure	Development and application of innovative UX evaluation methods for mobile contexts.	SUS score improvements; number of implemented recommendations; reduction of usability issues identified by Hotjar.
SDG 10	Reduced Inequalities	Improving platform accessibility across diverse user profiles and device types.	Device diversity metrics; performance differences by device; Hotjar metrics (e.g., reduced U-turns, rage clicks).
SDG 17	Partnerships for the Goals	Collaboration between academia and the private sector to address real-world challenges in UX.	Documented project outcomes; frequency of industry-academic feedback loops; practical adoption of proposed changes.

Table 1: Alignment of this project with the United Nations Sustainable Development Goals

Acknowledgements

As I reach the end of this challenging and rewarding journey, I want to express my heartfelt gratitude to the people who supported me and made this thesis possible.

First and foremost, I extend my deepest thanks to my supervisor, **Professor Alexandre Valle de Carvalho**, for his unwavering support, guidance, and availability throughout this entire process. Your dedication and encouragement were instrumental in helping me navigate this project, and for that, I am forever grateful.

To **Professors Teresa Silveira** and **Carla Teixeira Lopes**, thank you for your invaluable guidance in research methodologies. You always knew how to calm my moments of despair and helped me make sense of the concepts that were so crucial for this thesis. Your patience and encouragement made all the difference.

To **ZeroZero**, thank you for the opportunity to carry out this project in collaboration with your team. Having access to a real-world context was an important part of this journey and provided valuable insights that helped shape this work.

To my father, **Nuno Pinhal**, I owe everything. Thank you for always believing in me, standing by my side, and offering your unconditional support when I needed it most. Your confidence in me kept me going, even during the hardest moments.

To my cat, **Spooky**, for being my constant companion during long nights of writing, for reminding me to take breaks (mostly to feed her and rest), and for curling up on my lap, offering comfort and calm with her quiet presence and soft purrs.

To my boyfriend, **Fábio Nunes**, thank you for being my rock and reminding me to take breaks when I needed them most. Your care and insistence on helping me find balance saved me from countless breakdowns and allowed me to finish this journey with a clear mind. Thank you for listening to my endless rants about UX, for your hugs when things felt overwhelming, and for celebrating every tiny victory with me. Your love and support made everything more bearable and even joyful.

To my boss, **Denise Marques**, thank you for being such an incredible support throughout this journey. Your understanding, encouragement, and flexibility (especially in allowing me time and space to work on this thesis) meant more to me than words can express. Thank you also for the wonderful opportunity of working with you. You are truly one of the people who made this possible.

To my beloved friends **Rodrigo Couto** and **Sofia Soares**, thank you for always being there to make me laugh, distract me when I needed it, and cheer me on in every step of this journey. Your presence in my life is a constant reminder that I am never alone.

To my dear friend and companion from the Information Science degree, **Ana Pontes**, thank you for being the best partner to share moments of stress, laughter, and mutual support. I couldn't have asked for a better person to navigate this academic rollercoaster with.

To my group of friends from the **Information Science undergraduate program**, thank you for being there during those foundational three years and for staying in my life even though some

of you did not pursue the master's program. You are one of the most precious gifts that university has given me, and your support, then and now, has been incredibly meaningful. I am deeply grateful for the moments we shared and the continued presence you have in my life.

I would like to thank **FEUP** for providing me with an unforgettable master's experience. To all the professors who accompanied me from my undergraduate degree to the master's program, I extend my heartfelt gratitude. Each of you played a vital role in my academic journey, and your trust in me has been an invaluable source of motivation.

Lastly, I would like to thank myself, for the resilience, dedication, and for not giving up during the most challenging moments. This journey demanded both academic and personal growth, and reaching this milestone fills me with pride and gratitude.

To all of you, I dedicate my sincere thanks. Without your support, belief, and encouragement, this achievement would not have been possible.

Marta Pinhal

"Designers are not users."

Jakob Nielsen, 1993

Contents

List of Figures

List of Tables

List of Acronyms

ACM	Association for Computing Machinery
ECG	Electrocardiogram
FEUP	Faculty of Engineering, University of Porto
GDPR	General Data Protection Regulation
HCI	Human-Computer Interaction
HE	Heuristic Evaluation
SPSS	IBM SPSS Statistics (Statistical software for data analysis)
IEEE Xplore	Institute of Electrical and Electronics Engineers
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
QA	Quality Assessment
SDGs	United Nations Sustainable Development Goals
SLR	Systematic Literature Review
SOTA	State of the Art
SUS	System Usability Scale
TAM	Technology Acceptance Model
UC-CW	User-Centered Cognitive Walk-through
UEQ	User Experience Questionnaire
UX	User Experience
WCAG	Web Content Accessibility Guidelines
WoS	Web of Science
xG	Expected Goals
Zos	ZeroZero

Chapter 1

Introduction

This dissertation investigates the user experience of the ZeroZero mobile website through mixed-methods evaluation and derives actionable recommendations grounded in direct user feedback.

1.1 Contextualization

In recent years, mobile devices have become the primary gateway for accessing the Internet, leading to an exponential increase in the demand for mobile-friendly platforms. Studies have highlighted that mobile users expect seamless navigation and functionality tailored to smaller screens, making mobile optimization a critical focus for digital platforms (Nielsen; Budiu, 2012). Companies worldwide are compelled not only to adapt their websites to fit smaller screens but also to redefine their UX strategies to align with new interaction patterns. This adaptation extends beyond a simple visual redesign; it involves a comprehensive overhaul of the interface, navigation logic, and functionality to ensure optimized, efficient, and enjoyable usage (ISO, 2018).

For platforms like Zos, recognized as the world's largest sports database, this transition poses unique challenges. The platform's extensive repository of sports data necessitates meticulous organization to maintain usability and accessibility in its mobile version. The adaptation process involves addressing technical and design hurdles, such as interface simplification, response time optimization, and ensuring an intuitive user journey across devices with varying screen sizes and processing capabilities. Research emphasizes that user satisfaction plays a critical role in the success and continued relevance of online platforms, underscoring the importance of these transformations (Jain et al., 2023 and Rohrer, 2022).

Despite the critical nature of these changes, there is limited research examining the specific impacts of website-to-mobile transformations on UX, particularly for data-intensive platforms like Zos. This study aims to fill this gap by focusing on Zos's mobile adaptation as a case study, exploring user challenges, perceptions, and opportunities for improvement.

1.2 Problem statement

The transition from a desktop-centric website to a mobile-friendly platform can introduce significant usability challenges. Research highlights common issues such as navigation difficulties, interface clutter, and diminished user satisfaction due to mismatched user expectations (Nielsen; Budiu, 2012 and Rohrer, 2022). While the mobile-first approach is widely advocated, the specific impacts of these transitions on user behavior and satisfaction remain insufficiently explored, particularly for platforms with extensive and complex datasets.

The case of Zos exemplifies these challenges. As the world's largest sports database, it requires a careful balance between maintaining functional completeness and delivering an intuitive, efficient user experience. Ensuring that users can seamlessly navigate and access relevant information, even within the constraints of smaller screens and varying device capabilities, is a critical objective. Addressing these usability issues is essential to guarantee the platform's ongoing relevance and success, as user satisfaction has been consistently linked to the long-term viability of digital platforms (ISO, 2018 and Jain et al., 2023).

1.3 Motivation

This research is driven by the increasing importance of UX and usability within digital environments, particularly in the context of mobile platforms. HCI emphasizes the design of systems that are both intuitive and efficient, aligning with established goals to enhance user satisfaction and technology adoption, as highlighted by Jain et al., 2023 and ISO, 2018.

The adaptation of websites into mobile-friendly formats is a critical topic, considering the ubiquity of mobile devices and the demand for seamless navigation experiences. Jakob Nielsen's studies illustrate the enhanced performance and satisfaction rates achieved through mobile-optimized designs (Nielsen; Budiu, 2012). These transformations require comprehensive adjustments, not only to interface aesthetics but also to the overall user journey, ensuring accessibility and satisfaction across diverse devices.

By addressing a real-world case, this dissertation seeks to contribute to UX research by identifying usability gaps and proposing improvements informed by established methodologies and metrics (Rohrer, 2022). This approach aligns with the broader objectives of advancing user-centric design practices, particularly for complex platforms such as Zos, which requires tailored strategies to accommodate its extensive data and dynamic user needs.

1.4 Aim and Goals

The primary aim of this research is to evaluate the impact of Zos transformation into a mobile-friendly platform on the user experience, with the ultimate goal of proposing actionable improvements. This research will contribute to the academic discussion on UX evaluation and provide practical insight to optimize mobile platforms.

The specific objectives of this research are as follows:

1. Conduct a comprehensive literature review and SOTA to identify best practices in UX evaluation for mobile platforms.
2. Develop and substantiate methodological options for assessing the user experience of mobile platforms.
3. Identify usability gaps and propose targeted improvements based on user feedback and empirical findings.

1.5 Research Questions

This research seeks to answer the following question:

RQ: *In the transformation of a website into a mobile version, what feature-related design decisions mostly impact user experience—specifically learnability, efficiency, memorability, errors and satisfaction?*

These five elements (learnability, efficiency, memorability, error rate, and satisfaction) are considered key usability attributes, as defined by Nielsen, 1993. While they are often mentioned alongside Nielsen's usability heuristics, they serve a distinct role: rather than acting as prescriptive design guidelines, they offer measurable dimensions for evaluating the quality of user interaction. These particular attributes were selected due to their strong relevance in evaluating user experience during the transition from desktop to mobile environments, where intuitive design, task performance, and user satisfaction are critical success factors.

To address this research question, a review of the relevant literature and state-of-the-art approaches will be carried out in order to identify usability evaluation methods appropriate for mobile interfaces. Building upon this foundation, a methodological framework will be developed to assess the usability of the mobile version of ZeroZero's website, with a focus on feature-related design decisions and their impact on user experience—specifically in terms of learnability, efficiency, memorability, and satisfaction.

The study will involve selecting appropriate techniques and metrics, followed by the design of an empirical evaluation within a specific usage context. This is intended to support the identification of usability issues and potential areas for improvement in the mobile experience.

1.6 Dissertation structure

In addition to the introduction, this dissertation is structured into five more chapters, each contributing to the overall narrative of the study. Chapter 2 presents the literature review, providing an overview of the research conducted, including the databases selected, the criteria used for inclusion, and the rationale behind these choices. Chapter 3 describes the state of the art in the field,

showcasing related work and highlighting key advancements relevant to the study. Chapter 4 outlines the methodology, detailing the research paradigm, data collection techniques, and analytical approaches employed in the study. Chapter 5 focuses on the design and implementation of the data collection methods, including the UX evaluation questionnaire, the use of Hotjar for behavioral analytics, and the planning of focus groups as a qualitative complement. Chapter 6 presents the results of the questionnaires, hotjar and stuctured interviews. Chapter 7, focused on elaborating on the findings and their implications for the field of UX and the Zos mobile version. Chapter 8 delivers actionable recommendations grounded in user feedback and interviews.Finally, Chapter 9 concludes the dissertation by summarizing the research outcomes, emphasizing the study's contributions, and suggesting avenues for future research.

Chapter 2

Literature review

2.1 Introduction

This chapter reviews the literature on methods and metrics for evaluating the UX on websites that have gone mobile. The main objective is to identify and analyze UX evaluation methodologies, focusing on the difficulties and gaps associated with this transition.

The research was conducted using a systematic review process, structured to identify and analyze relevant studies. The methodology followed the PRISMA protocol [1] to ensure transparency, reproducibility, and rigor in the identification, screening, and inclusion of studies.

Although this work centers on a systematic literature review (SLR), it was preceded by a brief narrative review to provide context and introduce the subject matter. The references Nielsen; Budiu, 2012, Garrett, 2011, Holzinger, 2005, Jain et al., 2023, Kohavi et al., 2009, Moran et al., 0000, Rohrer, 2022, ISO, 2018, Nielsen, 1993, Brooke, 1996 and Sauro et al., 2012, Norman, 2005, Norman, 2013, originate from the narrative review, as they were not identified in the systematic process due to falling outside the specified year range or database scope. However, they proved vital to the research due to reasons such as understanding the evolution of core concepts in UX, HCI, and usability, particularly in the context of mobile interfaces, refining the comprehension of UX evaluation methods, metrics, and tools, ensuring a robust methodological foundation, and highlighting the importance of mobile versions in bridging usability challenges and aligning design strategies with user needs.

Although this chapter focuses on the systematic literature review, a small set of foundational works fell outside the SLR scope yet proved essential for other phases of the research. Yin, 2008 and Burrell et al., 2019 were drawn upon to underpin the case (study design and to frame the sociological dimensions of user behaviour in the mobile) web context. Likewise, Field, 2018 and Sauro et al., 2012 informed the interpretation of empirical findings, providing statistical techniques and quantitative benchmarks for analyzing questionnaire and interview data in the results chapter.

[1] https://www.prisma-statement.org/

2.2 Theoretical Framework

The theoretical foundation of this review focuses on HCI and UX evaluation methodologies. Core concepts such as usability evaluation methods, testing frameworks, and evaluation metrics are explored to contextualize the research within existing knowledge.

2.3 Databases and Queries Used

The review utilized major academic databases, including WoS Scopus, ACM Digital Library, IEEE Xplore, and Science Direct, due to their extensive collections of usability and user experience studies. The queries strings used are presented in table 2.1, they are designed to capture relevant literature.

Table 2.1: Search strategy for databases.

Database	**Query String**	**Period**	**Document**	**Language**	**Looking to find**
Wos, Science Direct, Scopus, IEEE Xplorer	("HCI" OR "Human-Computer Interaction") AND ("ux" OR "user experience") AND ("usability evaluation methods" OR "evaluation methods" OR "testing methods")	2019-2024	Article, Conference Paper, Proceding Paper, or Book	English	• About HCI and UX area • Various methodological options for evaluating UX
ACM	[All: [all: usability evaluation methods]] AND [All: [all: user experience web application]] AND [All: [ccs 2012: hci design and evaluation methods]] AND [E-Publication Date: Past 5 years]	2019-2024	Article, Conference Paper, Proceding Paper, or Book	English	• About HCI and UX area • Various methodological options for evaluating UX

In databases such as WoS, Science Direct, Scopus, and IEEE Xplorer, using an internal classification system similar to ACM's CCS is uncommon or unnecessary. Hence, the query relies solely on the main keywords ("HCI", "Human-Computer Interaction", "ux", "user experience", "usability evaluation methods", etc.) without referring to internal classifications, capturing a broader range of results. The difference between the query used for the ACM database and the queries applied to other databases (WoS, Science Direct, Scopus, and IEEE Xplorer) lies in the specific classification and organization system in ACM, requiring a tailored approach to capture relevant documents accurately. In ACM, using classification codes from the 2012 "Computing Classification System" (CCS) is essential. This system categorizes documents into granular knowledge areas, facilitating the identification of articles specifically in Human-Computer Interaction (HCI) and UX evaluation methods. Including the code "CCS 2012: HCI design and evaluation methods" in the query directly helps narrow the search effectively.

2.3.1 Query Refinement Process

The iterative refinement of the queries, as summarized in Table 2.2, was essential for improving the precision and relevance of the retrieved articles. Each query was adjusted to address specific challenges identified during the systematic review process.

Initially, the queries (QS1 and QS2) employed broad terms such as "HCI" and "User Experience". While these queries captured a wide range of articles, many were unrelated to the objectives of this study. For instance, QS1 retrieved articles with general discussions on HCI and UX but lacked focus on mobile contexts or usability evaluation methods. Similarly, QS2 added terms like "Satisfaction" and "Usability Improvement", narrowing the scope slightly toward user satisfaction and experience, but still failing to target mobile usability.

Subsequent refinements introduced more specific terms. QS3 included "mobile website" and "mobile" to shift the focus toward mobile platforms. However, this query still lacked precision in identifying studies related to usability evaluation methods. QS4 addressed this by incorporating terms like "usability evaluation methods" and "testing methods", which directly aligned with the objectives of identifying concrete evaluation approaches in HCI and mobile contexts.

For the ACM Digital Library, the refinement process followed a similar logic but required adjustments specific to its structure. The initial query in ACM used broad terms such as "evaluation user experience", which retrieved many irrelevant articles. To improve relevance, a second query incorporated the ACM classification code [CCS 2012: HCI design and evaluation methods], allowing for more targeted retrieval of studies. The final query maintained this classification and introduced additional keywords, such as "usability evaluation methods", to focus on identifying concrete methods and tools.

For the selected queries, precision was calculated to evaluate their effectiveness in retrieving relevant articles. Precision was determined using the formula:

$$\text{Precision} = \frac{\text{Number of Relevant Results}}{\text{Total Number of Results}}$$

This calculation was performed only for the final queries (QS5 across all databases and QS3 in ACM) due to the feasibility of analyzing their results comprehensively. QS5 achieved a precision of 11%, while QS3 in ACM reached 5%. These values reflect the refinement's success in aligning the retrieved articles with the study's objectives.

This iterative refinement process, illustrated in Table 2.2, highlights the importance of tailoring search queries to the unique structures of each database, ensuring the retrieval of articles that directly address the research focus.

Table 2.2: Query Refinements

Wos, Science Direct, Scopus, IEEE Xplorer	
Query ID	**Query String**
QS1	(HCI OR "Human-Computer Interaction") AND ("User Experience" OR UX)
QS2	("User Experience" OR UX) AND (Satisfaction OR "User Satisfaction") AND (Usability OR "Usability Improvement")
QS3	(HCI OR "Human-Computer Interaction") AND ("User Experience" OR UX) AND "Usability" AND ("mobile website" OR "mobile")
QS4	("HCI" OR "Human-Computer Interaction") AND ("ux" OR "user experience") AND ("usability evaluation methods" OR "evaluation methods" OR "testing methods")
QS5	("HCI" OR "Human-Computer Interaction") AND ("ux" OR "user experience") AND ("usability evaluation methods" OR "user experience evaluation metrics" OR "user experience evaluation instruments")
ACM	
QS1	[All: [all: evaluation user experience]] AND [All: [all: user experience web application]] AND [All: [hci design and evaluation methods]]
QS2	[All: evaluation user experience] AND [All: user experience web application] AND [CCS 2012: HCI design and evaluation methods]
QS3	[All: usability evaluation methods] AND [All: user experience web application] AND [CCS 2012: HCI design and evaluation methods]

2.3.2 Search Process

As mentioned in subsection 2.3.1, the search was structured in distinct phases for each database group (WoS, Science Direct, Scopus, IEEE Xplorer) and ACM, as seen in table 2.2. This separation was necessary due to each database's specific requirements for query format and topic categorization.

2.3.3 Inclusion and Exclusion Criteria

The inclusion and exclusion criteria were defined to ensure that only studies of high relevance and quality were considered. These criteria are listed in subsection 2.3.3.1 and 2.3.3.2.

2.3.3.1 Inclusion Criteria

The inclusion criteria were established to ensure the selection of studies that are directly relevant, recent, and of high quality, meeting the specific needs of this research. Each criterion plays a critical role in filtering the most suitable literature:

- **Direct relevance to the study's objective**: Ensuring studies align with the research objectives is essential to avoid unrelated or tangential material, focusing only on literature that contributes directly to the understanding and evaluation of UX methodologies.
- **Publications from 2019 to 2024**: Restricting the publication period to recent years guarantees that the research incorporates the latest advancements, trends, and tools in UX evaluation, which are particularly relevant in a rapidly evolving field like HCI.
- **Full-text access**: Full access to the documents is vital to gain a comprehensive understanding of each study's methodology, results, and implications. Abstracts alone often lack the depth required for critical evaluation.
- **Written in English**: Selecting studies in English broadens the scope to include a wider range of global research, as English is commonly used in scientific literature, thus enhancing the diversity and applicability of insights.
- **Use of specific keywords**: Keywords like "UX evaluation methods" "evaluation methods," "assessment methods," "evaluation frameworks," "evaluation techniques," and "testing methods" target studies focused specifically on methodologies for UX evaluation. This focus helps in identifying studies that discuss relevant evaluation approaches, frameworks, or techniques for UX assessment.
- **Inclusion of various document types (Articles, Conference Papers, or Books)**: Including different types of documents enriches the review by incorporating perspectives from foundational research (books), the latest findings (conference papers), and peer-reviewed studies (journal articles), offering a well-rounded understanding of the field.

2.3.3.2 Exclusion Criteria

The exclusion criteria were carefully designed to remove studies that do not meet the quality or relevance standards required for this research. Each criterion ensures that only meaningful and valuable studies are retained:

- **Duplicates**: Removing duplicate entries prevents redundancy in the literature review, ensuring that each study is reviewed only once and avoiding inflated data due to repeated analyses.
- **Lack of abstract or keywords**: Studies without abstracts or keywords are excluded, as these elements provide essential information for quickly assessing relevance. Without them,

determining the focus and scope of a paper becomes difficult, especially in systematic reviews.

- **Irrelevant titles**: Titles serve as the first indicator of a study's focus. Excluding studies with irrelevant titles helps streamline the selection process, allowing a more efficient focus on directly relevant literature.
- **Irrelevant abstracts**: Even if titles seem relevant, abstracts provide a clearer insight into a study's content. Excluding studies with irrelevant abstracts prevents the inclusion of materials that don't align with the specific UX evaluation focus of this research.
- **Lack of full-text access**: Studies without full-text availability are excluded, as partial information hinders an in-depth analysis of methodologies, findings, and limitations, which are essential for a comprehensive literature review.
- **Uninteresting conclusions**: Papers with conclusions that do not offer insightful findings or relevant contributions to UX evaluation methods are excluded to maintain a focus on studies that add substantial value to the research.
- **Not summaries of workshops, seminars, lectures, or tutorials**: These types of documents are excluded because they often lack rigorous research methodologies and empirical evidence. Instead, they tend to summarize knowledge without the detailed analysis or validation needed for a thorough academic review.

2.3.4 Identification

As shown below in the table 2.3, a total of 429 initial results were retrieved from these databases. After applying the inclusion and exclusion criteria, 39 articles were identified as relevant for further analysis, representing 9.1% of the total results. The remaining 390 results were excluded, accounting for 90.9% of the total results, as represented in the image 2.1 below.

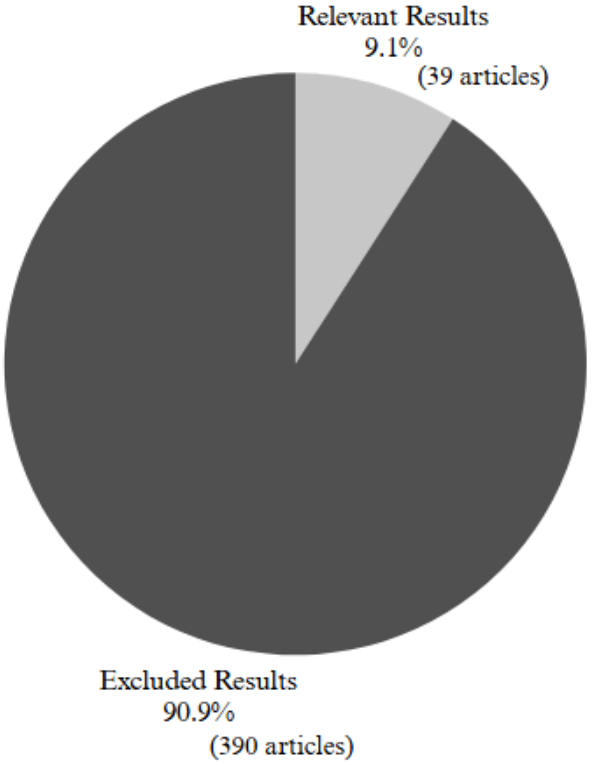


Figure 2.1: Final Results after Inclusion and Exclusion Criteria Application

Among the databases, Science Direct yielded the highest number of initial results (291), followed by ACM Digital Library (117). Despite this, Science Direct also had the highest exclusion rate, with only 26 articles identified as relevant. The ACM Digital Library contributed 6 relevant articles, highlighting its importance as a source of high-quality studies on usability and UX evaluation.

Scopus and Web of Science produced more modest results, with 14 and 3 initial results, respectively. From these, 5 articles were deemed relevant from Scopus, and 2 articles from Web of Science. IEEE Xplore, although a recognized database, did not contribute any relevant articles to the review.

Table 2.3: Results of the Systematic Review by Database

Database	**Initial results**	**Relevant Results**	**Excluded Results**
Web of Science	3	2	1
Scopus	14	5	9
Science Direct	291	26	265
ACM	117	6	111
IEEE Xplore	4	0	4
No. Total Results	**429**	**39**	**390**
%	**100%**	**9%**	**91%**

As seen in the table 2.3, the inclusion and exclusion process was critical in ensuring that only the most methodologically sound and contextually relevant studies were retained for the review. This highlights the importance of systematic filtering when conducting a robust literature review.

The percentages shown in Table 2.3 were calculated using the following formula:

$$\text{Percentage} = \frac{\text{Number of Results in Category}}{\text{Total Number of Initial Results}} \times 100$$

For instance, the percentage of relevant results (9%) was calculated as:

$$\text{Percentage of Relevant Results} = \frac{39}{429} \times 100 = 9\%$$

Similarly, the percentage of excluded results (91%) was obtained as:

$$\text{Percentage of Excluded Results} = \frac{390}{429} \times 100 = 91\%$$

To provide a clear and structured overview of the systematic review process, the PRISMA flow diagram (Figure 2.2) was utilized. This diagram visually represents the identification, screening, and inclusion phases of the study selection, highlighting the number of records processed at each stage and the reasons for exclusion.

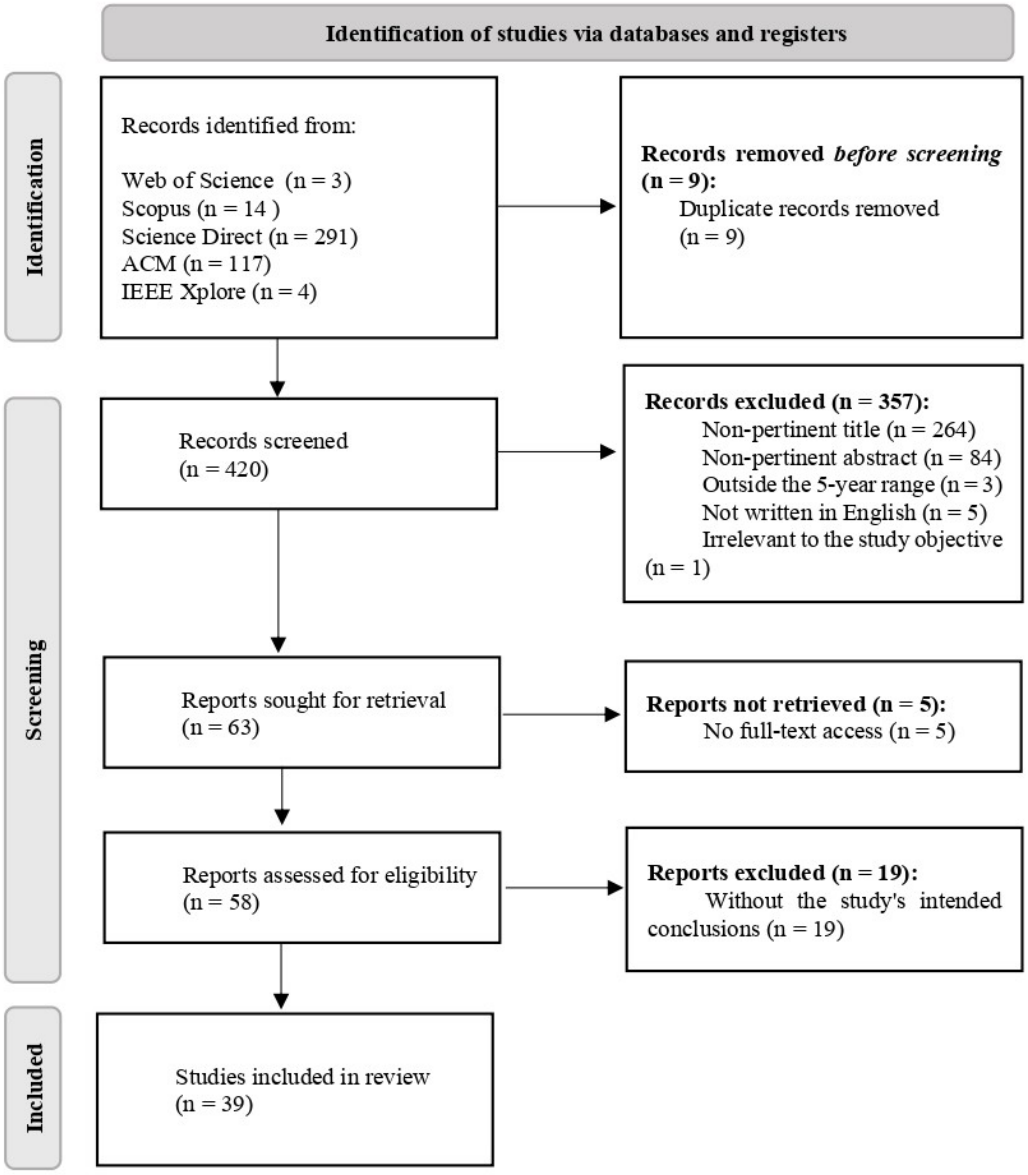


Figure 2.2: PRISMA Flow Diagram

The exclusion criteria applied during the systematic review process are detailed in Table 2.4. The table summarizes the total number of studies excluded based on specific criteria, highlighting the steps taken to ensure that only the most relevant and high-quality articles were included in the review.

During the initial search, a range of years and specific types of documents of interest were predefined within the selected databases. This filtering process was intended to reduce the inclusion of irrelevant studies, such as workshop summaries, surveys, seminar papers, or tutorials. Consequently, no documents of these types were identified in the results, as indicated by the exclusion count of zero for this category. However, despite setting a range of years, a small number of studies (3 articles) outside the desired range were still retrieved and subsequently excluded.

The most frequent reason for exclusion was titles not being pertinent, accounting for 264 articles. This reflects the importance of the title as a first indicator of relevance. Similarly, 84 articles were excluded due to summaries not aligning with the objectives of the study, 19 articles were excluded due to uninteresting or irrelevant conclusions, and 1 article was excluded due to be irrelevant to the study objective.

Additionally, 9 duplicate articles were removed to avoid redundancy, while 5 articles were excluded due to a lack of full-text access. Lastly, 5 studies were excluded as they were not written in English, which was a necessary inclusion criterion to ensure consistency in analysis.

This rigorous exclusion process ensured that only highly relevant studies remained for detailed review, allowing the focus to remain on methods, tools, and practices for evaluating usability and user experience in mobile platforms.

Table 2.4: Results Excluded by Criteria

Exclusion Criteria	**Total Count**
Duplicate	9
Irrelevant to the study objective	1
No abstract/keywords	0
Non-pertinent title	264
Non-pertinent abstract	84
Uninteresting conclusions	19
No full-text access	5
Other document types outside the inclusion criteria (e.g., workshop summaries, surveys, seminars, tutorials)	0
Outside the 5-year range	3
Not written in English	5

To enhance the understanding of the exclusion criteria applied during the systematic review, a graph in figure 2.3 is included to visually represent the distribution of excluded results across the defined categories. This graphical representation complements Table 2.4 by providing a clearer overview of the most common reasons for exclusion, such as non-pertinent titles and abstracts, which accounted for the majority of the excluded articles. The graph facilitates the comparison between categories, highlighting the relative frequency of each criterion and emphasizing the rigor of the filtering process.

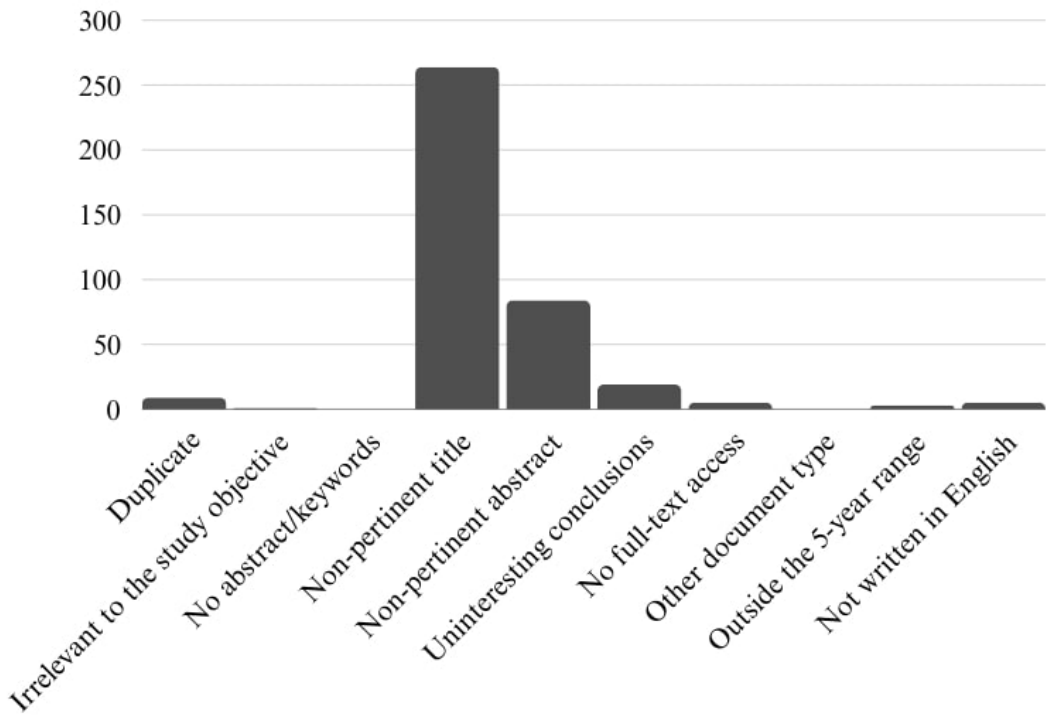


Figure 2.3: Excluded results by criterion

The chart below (figure 2.4) illustrates the distribution of document types among the studies included in this review. Articles constitute the majority, accounting for 74.4% of the total. This is followed by conference papers at 12.8%, and review articles, which represent 10.3%. Book chapters form the smallest category, comprising only 2.6%. This distribution highlights a strong reliance on articles as primary sources, with less emphasis on book chapters and review articles. The predominance of journal articles indicates the central role of scholarly journals in disseminating research findings within the field.

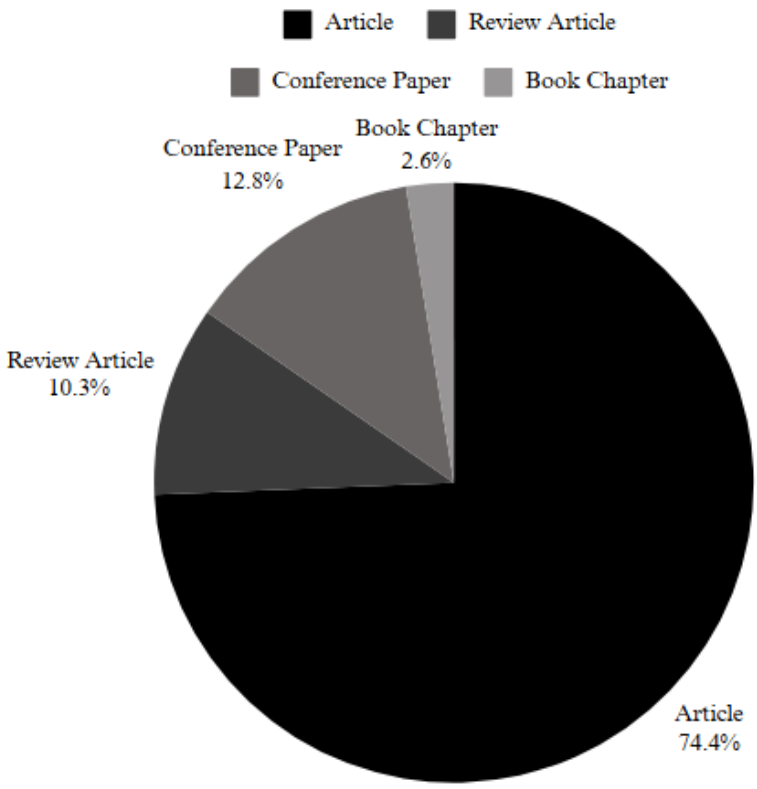


Figure 2.4: Results Included by Document Type

The distribution highlights the importance of peer-reviewed publications in ensuring the credibility and relevance of the included studies, although the relatively smaller share of conference papers and book chapters may point to limitations in covering more niche or emerging topics discussed in such formats.

To further enrich the process of identifying and contextualizing the 39 included articles, a citation and temporal analysis was conducted using LitMaps [2]. This analysis explores the relationships among the articles based on their publication dates and citation counts, providing insights into their influence and interconnections.

The metadata of the selected articles was imported into LitMaps to generate a chronological and citation-based map. This map highlights:

- **Chronology:** The X-axis positions articles by their publication dates, showing how research has evolved over time.
- **Citation Impact:** The Y-axis reflects the number of citations, identifying influential articles that have significantly shaped the field.

[2] LitMap link for this analysis: https://app.litmaps.com/shared/84170347-932d-4c90-9e3b-be17b70f811c

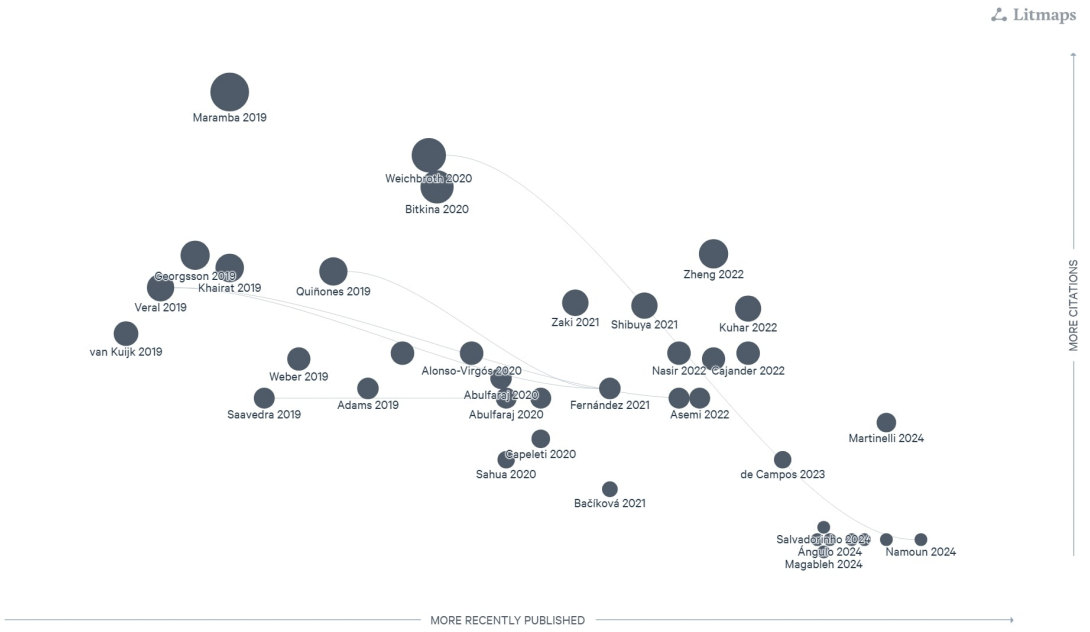


Figure 2.5: Authors relations on Lit Maps

As visualized in Figure 2.5 and Tables 2.5 and 2.6, the analysis revealed several notable patterns:

- **Influential Works:**
  - Articles such as Weichbroth, 2020 and Bitkina et al., 2020 are prominent due to their high citation counts, marking them as core contributions to usability and UX evaluation research.
- **Emerging Research:**
  - Recent articles, including Martinelli et al., 2024, Campos et al., 2024, and Namoun et al., 2024, represent emerging research directions, particularly in AI-driven usability and advanced methodologies. Their lower citation counts may be reflective of their recency rather than lack of relevance.
- **Evolution of Themes:**
  - Early works, such as Georgsson et al., 2019 and Khairat et al., 2019, focus heavily on healthcare applications, while later works show a shift towards diverse areas like mobile applications from Zheng et al., 2022 study and psychophysiological measures from Zaki et al., 2021 study.
- **Interconnections:**
  - Citation links between articles illustrate intellectual relationships. For instance, Saavedra et al., 2019 and Quiñones et al., 2019 are interconnected through their focus on heuristic evaluation methods, forming a thematic cluster within the literature.

Table 2.5: Bibliometric Overview of Selected Articles *(part 1)*

Author(s)	**Title**	**References**	**Citations**
Alvarez et al., 2024	Development of a Community-Based Disaster Awareness Web-based Application for HCI: Pulse PH	0	0
Baciková et al., 2021	Domain Usability Evaluation	72	1
Kuijk et al., 2019	Drivers of usability in product design practice	108	16
Cheng et al., 2022	EasyGaze: Hybrid eye tracking approach for handheld mobile devices	48	11
Georgsson et al., 2019	Employing a user-centered cognitive walkthrough to evaluate a mHealth diabetes application	72	49
Nakai et al., 2024	Evaluation of Tourism Website User Experience Based on Psychophysiological Measures	15	0
Kuhar et al., 2022	Exploring user experience in digital libraries through questionnaire and eye-tracking data	49	23
Fernández et al., 2021	Heuristic-Based Usability Evaluation Support: A Systematic Literature Review	80	7
Capeleti et al., 2020	Heuristic evaluation and user tests of wearable mobile health monitoring applications	25	3
Martins et al., 2024	Heuristic Evaluation of a Web-Based Application for Older Adults	15	0
Shibuya et al., 2022	Mapping HCI research methods for studying social media interaction	158	24
Maramba et al., 2019	Methods of usability testing in the development of eHealth applications	42	471
Zaki et al., 2021	Neurological and physiological measures to evaluate usability and UX of IS	84	25
Namoun et al., 2024	Predicting the usability of mobile applications using AI tools	23	0
Magableh et al., 2024	Psychometric evaluation of the Arabic SUPR-Q	43	0
Veral et al., 2019	Supporting user-perceived usability benchmarking through a developed quantitative metric	38	31
Alonso-Virgós et al., 2020	Test usability guidelines and follow conventions	59	12
Campos et al., 2024	Usability and User Experience Evaluation of Touchable Holographic Solutions	80	2
Bitkina et al., 2020	Usability and user experience of medical devices	71	127
Bokingkito et al., 2022	Usability evaluation of a real-time water quality monitoring mobile application	16	6
Nasir et al., 2022	Usability inspection: Novice crowd inspectors versus expert	59	12

Table 2.6: Bibliometric Overview of Selected Articles *(part 2)*

Author(s)	Title	References	Citations
Weber et al., 2019	Usability of development tools: a case-study	21	11
Weichbroth, 2020	Usability of Mobile Applications: A Systematic Literature Study	192	168
Yuhui et al., 2024	Using Animation to Develop a Usability Questionnaire	137	0
Cajander et al., 2022	UX professionals' learning and usage of UX methods in agile	65	12
Zheng et al., 2022	UX Research on Conversational Human-AI Interaction	230	50
Martinelli et al., 2024	UX Research practices related to Long-Term UX	50	4
Angulo et al., 2024	Workspace metrics and evaluation	46	0
Adams et al., 2019	Co-created evaluation: Identifying how games support police learning	25	10
Abulfaraj et al., 2020a	Coherent heuristic evaluation (CoHE): Toward increasing the effectiveness of heuristic eval	40	6
Saavedra et al., 2019	A Set of Usability and User Experience Heuristics for Social Networks	42	12
Quiñones et al., 2019	Applying a methodology to develop user eXperience heuristics	43	39
Sahua et al., 2020	A Systematic Review of Usability Evaluation Methods and Tools for ATM Interfaces	22	2
Nowakowski, 2020	Analysis and Evaluation of Information Usefulness and User Experience	16	12
Abulfaraj et al., 2020b	Detailed Usability Heuristics: A Breakdown of Usability Heuristics	37	7
Salvadorinho et al., 2024	A talent management platform for HR 4.0 challenges: usability evaluation	15	0
Khairat et al., 2019	A mixed-methods evaluation framework for electronic health records usability studies	29	41
Asemi et al., 2022	A judgment-based model for usability evaluating of interactive systems using fuzzy Multi Factors Evaluation (MFE)	52	6
Krawiec et al., 2020	A comparison of heuristics applied for studying the usability of websites	20	6

2.3.5 Results

The selected studies encompass a wide range of usability evaluation methods, including heuristic evaluations, user testing, and mixed-method approaches. They cover applications in diverse fields such as health, social media, disaster management, and conversational AI. These studies provide knowledge into effectiveness, efficiency, and user satisfaction metrics as part of the evaluation process.

Evaluating UX is essential to ensure that systems and applications are effective, efficient, and satisfying to users. The following tables provide a structured synthesis of the Methods, Metrics,

and Instruments used for UX evaluation, based on 39 reviewed articles. Each table details characteristics, applications, limitations, and supporting citations for each item presented.

2.3.6 Quality Assessment

The Quality Assessment (QA) process is a fundamental step in ensuring the reliability and robustness of a SLR. In the context of this dissertation, the QA was designed to evaluate the relevance, rigor, and comprehensiveness of the selected studies that focus on UX evaluation methods for mobile websites. By identifying strengths and weaknesses within the body of literature, the QA process provides critical insights into the state of the art and guides the subsequent analysis and discussions.

2.3.6.1 QA Methodology

The QA was conducted using a structured questionnaire consisting of six key questions (QA1 to QA6), each addressing a specific dimension of quality in the evaluated studies. The questions and their respective purposes are outlined as follows:

- **QA1 (Purpose):** Does the article explicitly address the purpose and application of UX evaluation methods?
- **QA2 (Extend of State of the Art):** Does the study expand on the state of the art in UX evaluation or introduce new methods/metrics?
- **QA3 (Details on Methodology):** Does the study clearly describe the methodology used for UX evaluation, including tools, metrics, and specific processes?
- **QA4 (Use of Proven Tools):** Does the study employ well-established tools and frameworks for UX evaluation?
- **QA5 (Application):** Does the study present practical applications or real-world examples of UX evaluation methodologies?
- **QA6 (Limitations and Future Work):** Does the study discuss its limitations and provide suggestions for future research?

The evaluation criteria for each question were categorized as *Yes*, *Partially*, or *No*, indicating the extent to which the study met the respective criterion. A total of 39 studies were assessed based on this framework.

2.3.6.2 QA Results

The results of the QA are summarized in Table 2.7 and illustrated in Figure 2.6, which provides a graphical representation of the distribution of responses across the six questions.

Question	Response	Count
QA1	Yes	27
	Partially	4
	No	2
QA2	Yes	20
	Partially	10
	No	5
QA3	Yes	21
	Partially	3
	No	15
QA4	Yes	24
	Partially	7
	No	3
QA5	Yes	30
	Partially	2
	No	11
QA6	Yes	24
	Partially	11
	No	1

Table 2.7: QA Results Summary

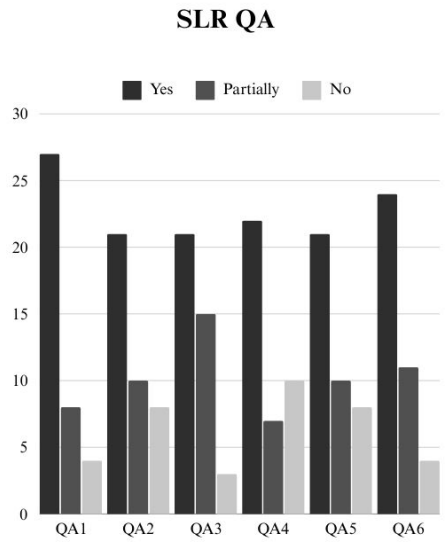


Figure 2.6: Distribution of QA Results

As for the results mentioned in Table 2.7 and Figure 2.6, the following observations were made:

- **QA1 (Purpose):** The majority of studies (69%) explicitly address the purpose and application of UX evaluation methods, demonstrating a strong focus on UX in mobile contexts.
- **QA2 (State of the Art):** Only 54% of studies contribute new advancements, with many revisiting or adapting existing approaches.
- **QA3 (Methodology):** While most studies (54%) provide detailed methodologies, 38% only partially describe their processes, and 8% lack clarity.
- **QA4 (Proven Tools):** A significant portion of studies (26%) fail to utilize established tools, potentially affecting the credibility of their findings.
- **QA5 (Application):** Practical applications were demonstrated in 54% of studies, though a notable 20% provided limited or no real-world examples.
- **QA6 (Limitations and Future Work):** Most studies (62%) discuss limitations and propose future directions, highlighting an awareness of the field's evolving nature.

2.3.6.3 Key Insights from the QA

The QA process has highlighted notable strengths and identified critical gaps within the evaluated studies, identified below.

The identified **strengths** were categorized into three distinct groups:

- **Clear Purpose:** Most studies (69%) explicitly address the purpose and application of UX evaluation methods, indicating strong alignment with research objectives.
- **Acknowledgment of Limitations:** A majority of studies (62%) openly discuss their limitations and propose future directions, showcasing a critical understanding of challenges in UX evaluation.
- **Adoption of Proven Tools:** Over half of the studies (56%) utilize well-established tools and frameworks, contributing to the reliability and comparability of their findings.

The identified **gaps** were as follows:

- **Lack of Innovation:** Only 54% of studies present advancements in the state of the art, indicating limited progress in introducing novel methodologies or metrics.
- **Methodological Inconsistencies:** Nearly 38% of studies partially describe their methodologies, and 8% fail to provide sufficient detail, leading to questions about reproducibility.
- **Underrepresentation of Practical Applications:** A significant number of studies (46%) either lack real-world examples or provide limited practical insights, reducing the applicability of their findings.

The QA results, detailed in the accompanying Excel file in appendix A, include specific classifications for each article based on the QA1 – QA6 criteria. These findings provide a comprehensive understanding of how each study contributes to the research on usability and user experience.

2.3.6.4 Discussion and Implications

The QA results reveal both strengths and weaknesses within the reviewed literature. While the majority of studies effectively address the purpose and application of UX evaluation, there is room for improvement in the innovation of methods (QA2) and the use of proven tools (QA4). Furthermore, the relatively high proportion of partially addressed methodologies (QA3) suggests that methodological rigor remains inconsistent.

These findings have implications for the reliability of the SLR and underline the need for clear, comprehensive methodologies in future UX research. Additionally, the limited adoption of proven tools in some studies underscores the importance of standardization within the field.

2.3.6.5 Limitations

The QA process itself is subject to limitations. For instance, the categorization of *Partially* responses involves a degree of subjectivity that may affect the consistency of results. Additionally, the framework used does not account for the relative importance of each question, potentially overlooking nuances in study quality.

2.4 Multidimensional analisys of the selected articles

To ensure a comprehensive and structured analysis of the literature, this review adopts a thematic approach, with elements of a methodological perspective where applicable. The studies are organized around key themes, reflecting the primary focus areas in the literature. This approach allows for a clear understanding of how UX evaluation methodologies have been applied and developed across different contexts, while also highlighting significant methods, tools, and trends.

2.4.1 Thematic Approach

To gain a deeper understanding of the key themes and interconnections within the selected 39 articles, a term co-occurrence analysis was conducted using VOSviewer [3]. This tool allows the visualization and clustering of frequently used terms, helping to identify dominant themes and relationships within the corpus.

The metadata of the 39 selected articles, including titles, abstracts, and keywords, was exported into RIS format and analyzed using VOSviewer. The analysis focused on identifying terms with a minimum frequency of occurrence to ensure relevance while reducing noise. The terms were then clustered into thematic groups based on their co-occurrence and link strengths.

[3] https://www.vosviewer.com/

The term analysis provided valuable insights into the dominant topics and recurring themes across the literature, facilitating a better understanding of the research landscape. Additionally, it helped to visually identify gaps and overlaps, guiding the structuring of the thematic approach adopted in this review.

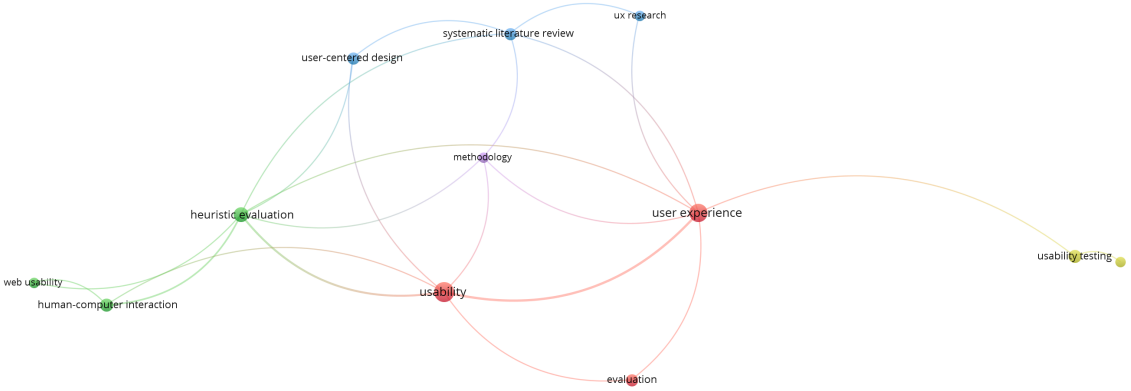


Figure 2.7: Term Co-occurrence Network of Selected Articles (VOSviewer)

The results of the analysis are visualized in Figure 2.7, which highlights the primary clusters and their relationships:

- **Green cluster:** Related to heuristic evaluation and web usability, highlighting the importance of evaluation techniques specific to web interfaces.
- **Blue cluster:** Focused on systematic literature reviews and user-centered design, emphasizing methodological rigor.
- **Red cluster:** Terms such as usability, user experience, and evaluation dominate this cluster, representing the core focus of the articles.
- **Yellow cluster:** Isolated topics like usability testing, reflecting practical application methods.
- **Purple cluster:** Highlighted methodology, showcasing standalone discussions about frameworks and approaches.

By visualizing the relationships between key terms, the analysis provided a comprehensive overview of how different research areas, methodologies, and applications intersect across the corpus. This visualization allowed for the identification of clusters, which directly informed the categorization of articles in the accompanying table.

Table 2.8 serves as a detailed classification of the articles based on their thematic characteristics, namely: Heuristic Evaluation, Web Usability, User-Centered Design, Usability Testing, and Methodology. Each article was systematically analyzed and categorized into one or more of these themes, reflecting its primary focus and contributions.

For example, articles categorized under Heuristic Evaluation include foundational works discussing frameworks and tools for assessing usability based on predefined heuristics. These studies align closely with terms clustered in the "Heuristic Evaluation" group in the VOSviewer analysis. Similarly, the Web Usability category encompasses studies focusing on usability challenges specific to web interfaces, echoing the connections identified in the visualization.

The categorization also underscores the interdisciplinary nature of some studies. For instance, articles that fall under both User-Centered Design and Usability Testing illustrate a methodological focus on participatory approaches while simultaneously emphasizing practical application through testing. Likewise, the inclusion of articles in the Methodology category reflects contributions that propose novel frameworks or metrics for evaluating UX.

Table 2.8: Classification of Articles by Thematic Characteristics

Authors	**Heuristic Evaluation**	**Web Usability**	**User-Centered Design**	**Usability Testing**	**Methodology**
Baciková et al., 2021			X		X
Weichbroth, 2020			X	X	X
Abulfaraj et al., 2020b	X		X		
Abulfaraj et al., 2020a	X		X		X
Sahua et al., 2020	X			X	X
Weber et al., 2019	X		X		X
Saavedra et al., 2019	X				
Zaki et al., 2021			X	X	X
Magableh et al., 2024			X		X
Cajander et al., 2022			X	X	X
Quiñones et al., 2019	X		X		X
Martinelli et al., 2024	X		X		X
Kuhar et al., 2022			X	X	X
Shibuya et al., 2022			X	X	X
Bitkina et al., 2020	X		X	X	X
Nowakowski, 2020	X		X	X	
Martins et al., 2024	X	X	X		
Namoun et al., 2024				X	X
Cheng et al., 2022			X		X
Krawiec et al., 2020	X	X			X
Veral et al., 2019			X	X	X
Adams et al., 2019			X	X	X
Alonso-Virgós et al., 2020	X	X	X		X
Asemi et al., 2022	X		X		X
Nasir et al., 2022	X		X	X	X
Yuhui et al., 2024			X	X	X
Maramba et al., 2019	X		X	X	X
Salvadorinho et al., 2024			X	X	X
Georgsson et al., 2019			X	X	X
Angulo et al., 2024	X		X	X	X
Kuijk et al., 2019			X		X
Bokingkito et al., 2022				X	X
Khairat et al., 2019				X	X
Zheng et al., 2022			X		X
Nakai et al., 2024			X		X
Capeleti et al., 2020	X			X	
Campos et al., 2024				X	X
Alvarez et al., 2024		X	X		
Fernández et al., 2021	X				X

2.4.2 Chronological Elements

While the primary structure is thematic, there is also a recognition of chronological progress in UX evaluation methodologies. Studies prior to 2019 generally focused on traditional methods, such as heuristic evaluations and user testing. These methods were foundational but lacked adaptability to emerging technologies and platforms. From 2019 onwards, there has been a noticeable shift toward integrating advanced tools, such as eye-tracking and psychophysiological measures, reflecting a growing emphasis on real-time and data-driven UX evaluation. Recent studies (2024) demonstrate an increasing reliance on standardized metrics and instruments to address the lack of universal benchmarking tools. This chronological perspective highlights the evolution of UX evaluation methods to address emerging challenges and technologies, particularly the transition to mobile platforms and the incorporation of advanced evaluation tools.

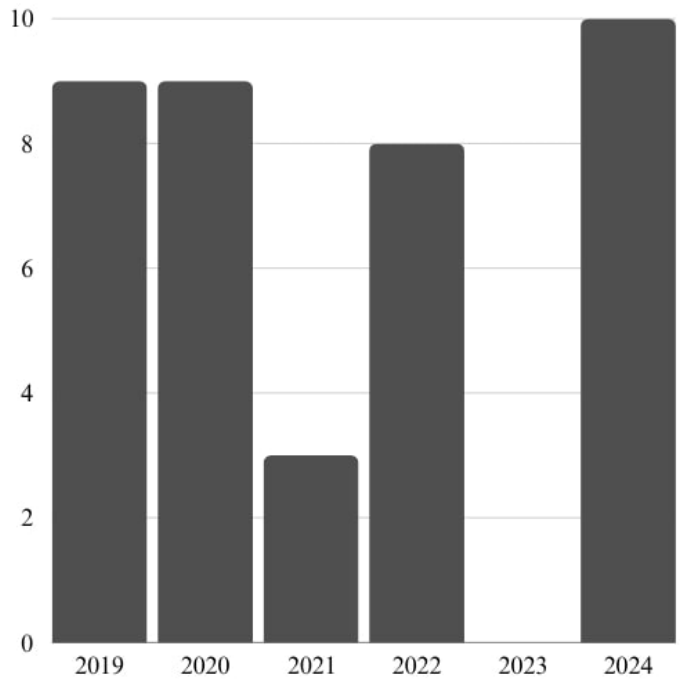


Figure 2.8: No. Results Included by publication Year

2.4.3 Methodological Perspective

In addition to the thematic organization, a methodological lens is applied to analyze how the studies conducted their evaluations. The methods are examined for their rigor, practical application, and contributions to the state of the art. Key methodologies identified include:

1. **Systematic Literature Reviews**

- Several studies employed systematic reviews to synthesize UX evaluation methods, providing comprehensive overviews of trends, gaps, and future directions like in the studies of Weichbroth, 2020 and Martinelli et al., 2024.

2. **Empirical Studies**

   - Empirical research focused on real-world applications of UX evaluation methods, testing tools and frameworks across specific domains such as healthcare, social media, and mobile platforms, as evidenced in the studies from Bitkina et al., 2020 and Alvarez et al., 2024.

3. **Experimental Methods**

   - Experimental approaches involved user testing, eye-tracking, and psychophysiological measures to evaluate UX under controlled conditions, offering objective insights into user behavior, as highlighted in the work of Nakai et al., 2024 and Capeleti et al., 2020.

2.4.4 Summary of Approach

By combining thematic organization with a methodological analysis and integrating chronological progress, this section provided a structured and holistic review of the literature. This multidimensional approach ensured that the key contributions, trends, and gaps in UX evaluation methods were identified, forming a strong foundation for the subsequent analysis.

2.5 Critical Analysis

2.5.1 Trends and Patterns

The literature reveals a strong reliance on established methods such as heuristic evaluations and user testing like mention in Cajander et al., 2022, Krawiec et al., 2020, Alonso-Virgós et al., 2020, Capeleti et al., 2020. These methods are often used to evaluate usability and user experience in various contexts, including websites, mobile applications and medical devices.

However, recent trends point to the integration of psychophysiological measures and mixed-method approaches to gain deeper insights into UX (Nakai et al., 2024, Zaki et al., 2021). This includes the use of tools such as eye-tracking (Alonso-Virgós et al., 2020), electroencephalography (EEG), electrocardiography (ECG) and other physiological measures to assess users' emotional and cognitive responses during interaction with systems as mentioned by Nakai et al., 2024: *"Traditional usability evaluation methods, such as questionnaires and surveys, require users to temporarily interrupt their website use or rely on recall to assess the usability of each content piece. This may cause a deviation between the real-time experiences and those expressed in the evaluation. By incorporating physiological indicators into the evaluation process, it is feasible to continuously assess the usability associated with individual content on the web-site. The usability of the respective website content is accumulated as UX. Therefore, the integration of physiological*

indicators allows for the ongoing assessment of UX for each piece of content, thereby enhancing evaluation accuracy.".

There is an increase in the use of questionnaires and interviews, often in combination with other techniques, to gain a more comprehensive understanding of the user experience. Questionnaires are often used to measure user satisfaction, while interviews provide qualitative information about their experiences. Methods such as Think-Aloud and focus groups are also used, although less frequently as mentioned by Maramba et al., 2019: *"The methods used for usability testing, in decreasing order of frequency were: questionnaires (n = 105), task completion (n = 57), 'Think-Aloud' (n = 45), interviews (n = 37), heuristic testing (n = 18) and focus groups (n = 13).".*

2.5.2 Gaps Identified

Several gaps emerged during the review:

1. **Lack of heuristic structures specifically adapted to mobile web platforms:** Although there are many usability and UX heuristics, there is a need for specific heuristics for mobile platforms and other particular contexts. Many sets of heuristics refer to Nielsen's heuristics, but there are original proposals for different interfaces (Krawiec et al., 2020, Abulfaraj et al., 2020b).

2. **Limited comparative evidence between different UX evaluation methods:** Although there are many UX evaluation methods, more studies are needed to compare the effectiveness and results of different approaches. The results of different methods can vary and the evaluation context is essential when comparing different methods. Some studies suggest that combining different methods can provide more valuable information as mentioned by Shibuya et al., 2022: *"Regarding data-collection methods, we find reviews on user experience evaluation methods by HCI scholars (Pettersson et al., 2018; Saket et al., 2016). Such works are more concentrated on the type of methods used and the reasons behind these choices. For instance, Pettersson et al. (2018) finds a research trend to conduct mixed methods in the field of HCI in recent years. In particular, the combination of questionnaire and interview data has been the most common way to mix methods. Simi-larly, Saket et al. (2016) reviewed HCI evaluation metrics specifically in visualization. The authors argued that evaluations of visualization need to utilize not only traditional usability-focused evaluation methods but the user-experience-focused evaluation methods, such as visualizations' memorability, engagement, and the sense of enjoyment and fun.".* Heuristic evaluation can also be combined with user testing as mentioned by Abulfaraj et al., 2020a *"Another technique is the hybrid usability methodology (HUM) proposed in [35], which suggests integrating HE with usability testing (UT). Combining both HE and UT increases the number of usability issues found.".*

3. **Insufficient integration of advanced psychophysiological tools:** Despite the growing interest in psychophysiological measures, their application in UX evaluation is still limited

and not widespread. The use of multi-modal approaches, combining psychophysiological measures with traditional methods, is still under development as mention by Martinelli et al., 2024: *"Some of the systematic reviews and mappings are often quoted in the literature [1,24,45]. Bargas-Avila and Hornbæk [1] was one of the first studies that investigated UX Research and included the moments when the research was carried out with the user (i.e., before, during, and after the usage). Papers between 2005 and 2009 were analyzed, and 92% of the publications referred to experiments only after the use of the product (i.e., Episodic UX) [1]. There are similar studies [24,45] in so far as they both investigate the integration of UCD activities during agile development."*. More research is needed to explore the effectiveness of these tools, as well as to better understand how the different measures can complement each other and provide a richer and more complete view of the user experience (Nakai et al., 2024).

While traditional methods such as heuristic evaluations and user testing are widely used, there is growing recognition of the importance of integrating psychophysiological measures and mixed-method approaches for a more complete and accurate assessment of the user experience. There is a clear need for more research to address gaps such as the lack of specific heuristic frameworks, the lack of comparative evidence between methods and the need for greater integration of advanced psychophysiological tools.

2.6 Summary

This section consolidates the main takeaways from the preceding subsections of the literature review, emphasizing critical findings, identified gaps, and trends:

- **Theoretical Framework (2.2):** Highlighted the core principles of HCI, Usability, and UX, emphasizing their interrelations and importance for designing user-centered systems. The foundations set here support the methodological and practical aspects of UX evaluation.
- **Databases and Queries Used (2.3):** The systematic search strategy across databases (WoS, ACM, IEEE, etc.) was refined iteratively to ensure precision. The inclusion and exclusion criteria narrowed down 429 initial results to 39 relevant articles, emphasizing the importance of relevance and methodological rigor in study selection.
- **Multidimensional Analysis (2.4):**
  - Thematic Approach: Key clusters identified include heuristic evaluation, web usability, and user-centered design, reflecting dominant themes in UX literature.
  - Chronological Elements: A shift from traditional heuristic methods to advanced tools (e.g., eye-tracking) was observed, aligning with technological advancements post-2019.

  - Methodological Perspective: Empirical studies and experimental methods gained prominence, showcasing an increased reliance on data-driven evaluation techniques.

- **Critical Analysis (2.5):**

  - Trends and Patterns: UX evaluation increasingly integrates psychophysiological measures and mixed-method approaches for richer insights.
  - Identified Gaps: Lack of tailored frameworks for mobile platforms, limited comparative evidence, and insufficient integration of advanced tools were notable.

This summary encapsulates the progress and limitations within UX evaluation, forming a basis for the subsequent analysis. It underscores the necessity of advancing standardized tools for mobile contexts while addressing usability challenges.

Chapter 3

State of the Art

3.1 Theoretical Framework

To understand the context of this work, it is essential to define and clarify the fundamental concepts that underpin the research. This section introduces the key concepts of HCI, UX, and Usability. These concepts form the backbone of the methodologies, theories, and practices employed to evaluate and improve user interactions with systems.

3.1.1 Human-Computer Interaction

HCI refers to the study of interactions between people (users) and computational systems. The focus of HCI is on designing user interfaces that are efficient, usable, and capable of providing a good user experience.

As Jain et al., 2023 refers that *"HCI is the study of how people interact with technology."* A primary goal of HCI is to develop systems that are user-friendly, improving technology adoption and satisfaction by creating intuitive, effective, and efficient designs, as Jain et al., 2023 mentioned: *"One of the key goals of HCI research is to develop user-friendly systems. This involves creating technology that is easy to use, efficient, and effective. By designing systems that are intuitive and user-friendly, researchers can improve user satisfaction and ensure that technology is more widely adopted."*.

3.1.2 User Experience

UX encompasses all aspects of a user's interaction with a product, service, or system, including their perceptions, emotions, and satisfaction. UX goes beyond functional usability to include the user's broader response to the system.

As defined by ISO, 2018 [1], *"User's perceptions and responses that result from the use and/or anticipated use of a system, product or service."* .

[1] ISO 9241-11:2018

3.1.3 Usability

Usability refers to the effectiveness, efficiency, and satisfaction with which users can achieve specific goals within a system in a defined context of use. It is a core aspect of UX evaluation and directly impacts how user-friendly a system is.

According to ISO, 2018, *"Extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."*.

Nielsen Norman Group simplifies the concept of usability as: *"Definition of Usability = how easy & pleasant these features are to use."*. [2]

Additionally, usability plays a critical role in designing educational technologies, as highlighted by Jain et al., 2023, *"Educational technology must be designed to be user-friendly and easy to use. HCI can ensure that educational technology is designed with the needs and preferences of learners in mind."*.

3.1.4 Evaluating User Experience

Evaluating user experience involves analyzing and measuring how users interact with a system, product, or service to enhance key aspects of usability, such as learnability, efficiency, memorability, error reduction, and satisfaction. This evaluation emphasizes not only functional and aesthetic aspects but also emotional perceptions, user expectations, and their needs during the interaction process.

The key attributes of UX evaluation, as defined by Nielsen, 1993, are summarized in Table 3.1:

[2] Retrived from: https://www.nngroup.com/articles/usability-101-introduction-to-usability/

Table 3.1: 5 Key attributes to evaluate usability

Attribute	Characteristics
Learnability	*"The system should be easy to learn so that the user can rapidly start getting some work done with the system."* by Nielsen, 1993
Efficiency	*"The system should be efficient to use, so that once the user has learned the system, a high level of productivity is possible."* by Nielsen, 1993
Memorability	*"The system should be easy to remember, so that the casual user is able to return to the system after some period of not having used it, without having to learn everything all over again."* by Nielsen, 1993
Errors	*"The system should have a low error rate, so that users make few errors during the use of the system, and so that if they do make errors they can easily recover from them. Further, catastrophic errors must not occur."* by Nielsen, 1993
Satisfaction	*"The system should be pleasant to use, so that users are subjectively satisfied when using it; they like it."* by Nielsen, 1993

3.2 Mobile Optimization and Usability: Practices and Theories

In the context of transitioning full websites to mobile versions, Jakob Nielsen advocates for creating mobile-optimized sites as a key practice to enhance user experience and increase user satisfaction. In studies conducted by the Nielsen Norman Group in 2009 and 2011, it was observed that user success rates were 64% when using mobile-optimized sites, compared to only 53% and later 60% for full sites, demonstrating that mobile adaptations lead to better performance and greater user satisfaction: *"When users are successful and satisfied, they're likely to come back"* Nielsen; Budiu, 2012. This improvement in usability reinforces the importance of an optimized mobile site.

Nielsen presents several key guidelines for creating an effective mobile experience, including building a separate mobile-optimized site, redirecting users to the mobile version instead of the full site whenever possible, and ensuring clear and simple navigation between versions. He recommends *"build a separate mobile-optimized site (or mobile site) if you can afford it"* and *"cut features to eliminate functionalities that are not core to the mobile use case" and "cut content to reduce word count and defer secondary information to secondary pages."*. However, simplification must be done carefully to avoid limiting the selection of products or services, ensuring that users can still find what they are looking for. As Nielsen states, *"A mobile site should have less information about each product and support fewer tasks that users can do with the products, but the*

range of items should remain the same as on the full site" Nielsen; Budiu, 2012. These practices help prevent frustration and promote a more enjoyable and effective browsing experience.

Nielsen's findings underscore the importance of optimizing user experiences specifically for mobile devices, as evidenced by the clear performance advantages observed with . His research illustrates that users not only succeed more often but also report greater satisfaction when interacting with sites designed to accommodate the unique constraints and affordances of mobile environments. Building upon this foundation, it is crucial to contextualize these insights within the broader theoretical frameworks of HCI, UX, and Usability. These fields provide the concepts and methodologies necessary to systematically evaluate and improve user interactions with mobile websites, ultimately informing best practices for designing user-centered mobile experiences.

3.3 UX evaluation

UX evaluation is a critical process to understand and improve how users interact with a system, ensuring it meets their needs and expectations. It provides valuable insights into usability, satisfaction, and overall user engagement, which are essential for designing and refining digital products.

For this evaluation, three distinct components were considered: UX evaluation methods (3.3.1), metrics (3.3.2), and instruments (3.3.3). Each of these components plays a unique role in assessing the UX from different perspectives:

- **Methods:** These represent strategies or approaches to conduct the evaluation. The primary focus here is on *how* the UX is assessed. This aligns with the *"Why & How to fix"* quadrant in the research methods landscape (Figure 3.2).
- **Metrics:** These are quantitative or qualitative data points collected during the evaluation. They focus on *what* aspects of UX are measured, such as task completion time or user satisfaction. This corresponds to the *"What people do"* and *"What people say"* quadrants, depending on whether the metrics capture behavioral or attitudinal data (Figure 3.2).
- **Instruments:** These are tools or frameworks employed during the evaluation process. Their focus lies on *what* is used to conduct the assessment, such as heuristic checklists or eye-tracking devices. These instruments are part of the *"What people do"* and *"What people say"* quadrants, as they help capture specific data for analysis (Figure 3.2).

Each of these components provides a complementary perspective, together forming a comprehensive framework to evaluate UX effectively. The combination of methods, metrics, and instruments ensures that both qualitative and quantitative aspects of UX are thoroughly examined, allowing for informed decisions to enhance user experience.

3.3.1 Evaluation Methods

UX evaluation methods serve as strategies to guide the assessment process and are fundamental to gaining insights about user interactions with a system. This study builds on the framework

proposed by Rohrer, 2022 [3], which organizes methods along a Three-Dimensional Framework, as seen in figure 3.1. This framework categorizes methods based on three key axes:

- **Attitudinal vs. Behavioral:** This axis contrasts *"what people say"* (attitudinal) with *"what people do"* (behavioral). Attitudinal methods explore user beliefs, preferences, and opinions, while behavioral methods focus on actual user actions and interactions.
- **Qualitative vs. Quantitative:** This dimension differentiates methods that generate open-ended, descriptive data (qualitative) from those that produce numerical, measurable data (quantitative). Qualitative methods are ideal for exploring the *"why"* behind user actions, whereas quantitative methods excel at answering *"how many"* or *"how much"* questions.
- **Context of Product Use:** This axis considers the context in which the product is used during evaluation. It ranges from natural use of the product in real-world settings to tightly scripted, controlled usage, or even scenarios where the product is not used at all.

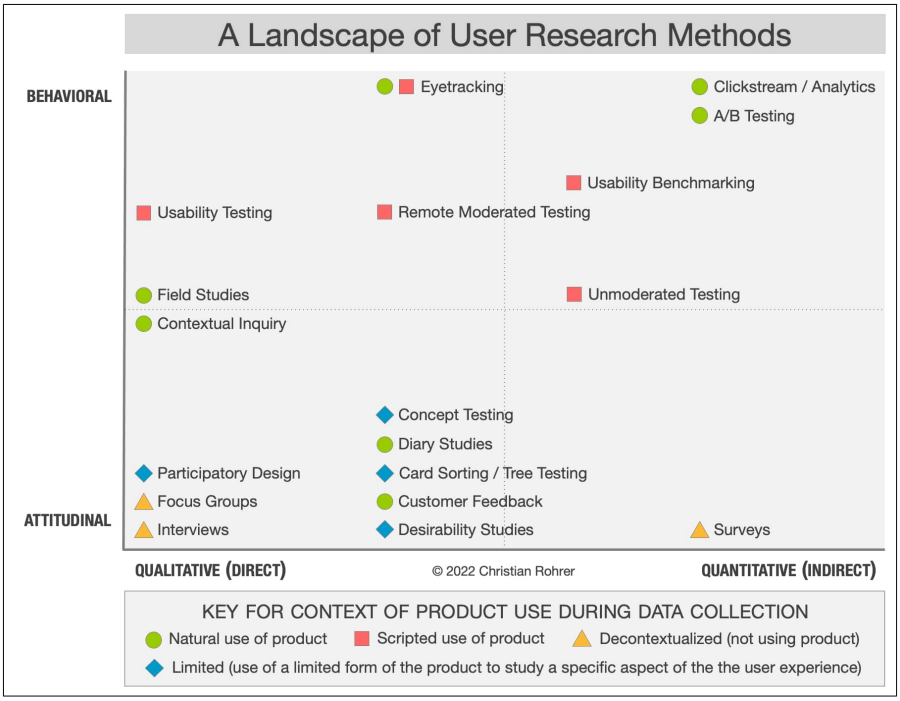


Figure 3.1: A Landscape of User Research Methods. *Retrieved from Rohrer, 2022.*

Rohrer, 2022 further emphasizes the importance of combining methods to achieve a comprehensive understanding of user experience, stating that: *"While it's not realistic to use the full set of methods on a given project, nearly all projects would benefit from multiple research methods and from combining insights. Unfortunately, many design teams only use one or two methods that they are most familiar with. The key question is what to use when."*.

[3] Retrived from: https://www.nngroup.com/articles/which-ux-research-methods/

The application of multiple methods ensures that both attitudinal and behavioral insights are captured, as well as qualitative and quantitative data, resulting in a richer understanding of the user experience, as seen in figure 3.2.

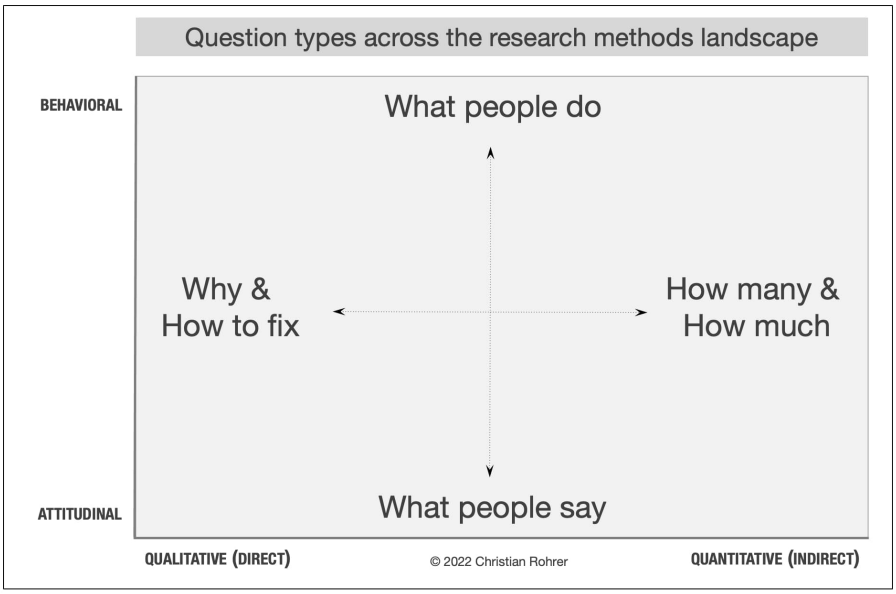


Figure 3.2: Question types across the research methods landscape. *Retrieved from Rohrer, 2022.*

Moreover, these methods can be mapped to different phases of the product development cycle:

- **Strategize (Generative Methods):** Early-stage methods used to identify new opportunities and define the direction of the product.
- **Design (Formative Methods):** Methods applied during the design phase to improve usability and refine features.
- **Launch and Assess (Summative Methods):** Methods used post-launch to measure performance and compare the product to competitors or earlier versions.

[3]Figure 3.1 and 3.2 retrieved from https://www.nngroup.com/articles/which-ux-research-methods/

Table 3.2: Phases of Product Development and Example Methods. *Retrieved from Rohrer, 2022.*

Product-Development Stage	**Strategize**	**Design**	**Launch & Assess**
Research goal	Find new directions and opportunities	Improve usability of design	Measure product performance against itself or its competition
Research methods	Generative research methods	Formative research methods	Summative research methods
Example methods	Field studies, diary studies, interviews, surveys, participatory design, concept testing	Card sorting, tree testing, usability testing, remote testing (moderated and unmoderated)	Usability benchmarking, unmoderated UX testing, A/B testing, clickstream/analytics, surveys

The following table presents 20 user experience evaluation methods, as categorized and described by Rohrer, 2022, offering a comprehensive view of the diverse strategies used in UX research.

[3]Table 3.2 Retrieved from https://www.nngroup.com/articles/which-ux-research-methods/

Nº	Methods	Features	Application (Geral)	Application (Zos)	References
1	Usability testing (aka usability-lab studies)	Participants are brought into a lab, one-on-one with a researcher, and given tasks of specific interest within a product or service	Identifying navigation bottlenecks in user interfaces and validating design solutions.	Identifying navigation bottlenecks in Zos mobile; evaluating task completion rates.	Rohrer, 2022; Weber et al., 2019; Cajander et al., 2022; Nowakowski, 2020; Nasir et al., 2022; Yuhui et al., 2024
2	Field studies	Researchers study participants in their natural environments, capturing real-world interactions.	Studying user behavior in real-world contexts, such as workplace or home.	Understanding how sports enthusiasts use Zos mobile in day-to-day scenarios.	Rohrer, 2022; Weber et al., 2019; Cajander et al., 2022; Martinelli et al., 2024; Nasir et al., 2022
3	Contextual inquiry	Combines observations and interviews to deeply analyze user workflows.	Examining workflows and challenges in task-oriented systems.	Investigating user workflows in accessing real-time sports data on Zos mobile.	Rohrer, 2022; Nakai et al., 2024; Cajander et al., 2022
4	Participatory design	Users collaborate in the design process using creative materials to express preferences.	Co-creating solutions with end-users for enhanced user engagement.	Designing sports data dashboards on Zos mobile tailored to user preferences.	Rohrer, 2022; Nakai et al., 2024
5	Focus groups	Groups of 3–12 participants discuss topics with a moderator, providing verbal and written feedback.	Gathering insights into user preferences for design and feature decisions.	Gathering insights on user preferences for sports statistics presentation on Zos mobile.	Rohrer, 2022; Weber et al., 2019; Martinelli et al., 2024; Shibuya et al., 2022; Bitkina et al., 2020; Adams et al., 2019; Maramba et al., 2019
6	Interviews	One-on-one discussions that explore user needs, preferences, and pain points in detail.	Exploring deep insights about user satisfaction, pain points, and needs.	Investigating navigation satisfaction and pain points with Zos mobile users.	Rohrer, 2022; Abulfaraj et al., 2020b; Bitkina et al., 2020; Maramba et al., 2019; Zheng et al., 2022
7	Eyetracking	Tracks where participants look while interacting with digital interfaces.	Identifying areas of attention or confusion in visual-heavy interfaces.	Evaluating focus areas on dynamic sports content, such as live scores, on Zos mobile.	Rohrer, 2022; Weichbroth, 2020; Weber et al., 2019; Cheng et al., 2022; Nakai et al., 2024; Kuhar et al., 2022; Shibuya et al., 2022
8	Usability benchmarking	Quantitative metrics track usability improvements over time or across versions.	Measuring improvements across iterative design cycles or comparing products.	Benchmarking usability improvements for Zos mobile interface updates over time.	Rohrer, 2022; Kuijk et al., 2019; Weber et al., 2019; Bitkina et al., 2020
9	Remote moderated testing	Usability studies are conducted remotely using video conferencing, screen-sharing, and other digital tools.	Conducting usability testing with geographically distributed participants.	Testing geographically distributed user groups for Zos mobile usability improvements.	Rohrer, 2022; Weber et al., 2019
10	Unmoderated testing	Automated methods record participant behaviors while they perform predefined tasks or goals.	Scaling usability studies for large samples without direct moderation.	Collecting usability metrics, such as task success rate and time-on-task, for Zos mobile platform.	Rohrer, 2022

Table 3.3: Methods for user experience evaluation *(part 1)*

Nº	Methods	Features	Application (Geral)	Application (Zos)	References
11	Concept testing	A researcher shares an approximation of a product or service that captures the key essence (value proposition).	Testing initial ideas for product features or visual layouts.	Evaluating new layout ideas for sports dashboards or navigation in Zos mobile.	Rohrer, 2022; Sahua et al., 2020; Cajander et al., 2022
12	Diary studies	Participants use a mechanism to record and describe aspects of their experiences over time.	Tracking user behavior and challenges over extended periods.	Understanding how users engage with Zos mobile in various contexts over time.	Rohrer, 2022; Weber et al., 2019; Shibuya et al., 2022
13	Customer feedback	Open-ended and/or close-ended information is collected from users through feedback forms or buttons.	Gathering feedback to identify recurring user complaints or feature requests.	Collecting ongoing feedback on navigation, statistics, and usability of Zos mobile.	Rohrer, 2022
14	Desirability studies	Participants evaluate visual-design alternatives by associating them with predefined attributes or adjectives.	Testing design aesthetics and emotional appeal of interfaces.	Comparing alternative designs for sports data presentation on Zos mobile.	Rohrer, 2022
15	Card sorting	Participants organize items into groups and assign categories, helping reveal their mental models.	Refining content organization and improving navigation in digital products.	Optimizing the categorization of sports-related content on Zos mobile.	Rohrer, 2022; Martinelli et al., 2024
16	Tree testing	Quantitative testing of an information architecture to determine how easy it is to find items in the hierarchy.	Validating and improving the hierarchy of website menus and structures.	Testing navigation hierarchies in sports statistics and team data on Zos mobile.	Rohrer, 2022; Moran et al., 0000
17	Analytics	Analyzes user behavior such as clicks, form completion, and time spent on pages.	Monitoring user engagement, traffic patterns, and interaction trends.	Understanding user interaction patterns and engagement metrics for Zos mobile.	Rohrer, 2022; Weber et al., 2019; Martinelli et al., 2024
18	Clickstream analytics	A specific type of analytics analyzing the sequence of pages users visit.	Mapping user flows to identify frequent paths and points of drop-off.	Mapping how users navigate between team pages, statistics, and live scores on Zos mobile.	Rohrer, 2022; Martinelli et al., 2024
19	A/B testing (multivariate testing)	A method of testing different designs by randomly assigning groups of users.	Optimizing layouts, calls-to-action, or feature placements in digital interfaces.	Testing layout and content variations to improve user retention and engagement on Zos mobile.	Rohrer, 2022; Martinelli et al., 2024; Moran et al., 0000
20	Surveys	Quantitative measures of user attitudes collected via surveys.	Collecting user perceptions, satisfaction, and suggestions at scale.	Gathering feedback from sports fans on usability and content relevance in Zos mobile.	Rohrer, 2022; Weber et al., 2019; Kuijk et al., 2019; Nakai et al., 2024; Baciková et al., 2021; Sahua et al., 2020; Martinelli et al., 2024; Shibuya et al., 2022; Bitkina et al., 2020; Yuhui et al., 2024; Maramba et al., 2019; Zheng et al., 2022; Campos et al., 2024

Table 3.4: Methods for user experience evaluation *(part 2)*

3.3.2 UX Evaluation Metrics

The following tables 3.5 and 3.6 focuses on the metrics used to assess UX. Metrics are essential for quantifying usability, user satisfaction, and system effectiveness. The table includes standardized tools like the System Usability Scale (SUS) and performance-related metrics like task completion time.

Metrics	Features	Relevance	Application (Geral)	Application (Zos)	Refs
Click-through rate	Measures the number of clicks required to complete a task.	Indicates navigation efficiency and task complexity.	Measuring user navigation flow and identifying interaction bottlenecks in digital systems.	Understanding user navigation patterns through team pages, live scores, and statistics on Zos mobile.	Khairat et al., 2019
Dwell time / Task Completion Time	Measures the time spent on a specific page or task.	Key indicator of task efficiency and usability.	Evaluating efficiency in completing critical tasks or accessing information.	Analyzing time spent by users retrieving sports data or navigating between pages on Zos mobile.	Khairat et al., 2019; Zheng et al., 2022; Salvadorinho et al., 2024
Success rate	Percentage of users who successfully complete a task.	Indicates overall effectiveness of the system.	Assessing system effectiveness in enabling users to achieve their goals.	Evaluating how often users can successfully find sports information (e.g., match details) on Zos mobile.	Khairat et al., 2019
Ease-of-use rating (post-task question)	After each task, the participants choose an option from a scale to indicate how easy or difficult the task was, and then those responses are averaged.	Helps assess the perceived difficulty of a task and provides an overall indication of system usability.	Used to measure ease of use after task completion in usability studies and interface testing.	Evaluates how users perceive ease of use when performing specific tasks in the Zos mobile plataform, such as navigation or searching for results.	Moran et al., 0000
NASA TLX	Measures cognitive load across six dimensions: mental, physical, time, frustration, performance, and effort.	Highlights user workload and usability challenges.	Evaluating the cognitive demands of completing complex or multi-step tasks.	Measuring cognitive load for tasks such as finding team comparisons or historical sports statistics on Zos.	Angulo et al., 2024; Khairat et al., 2019
NASA RTLX	Simplified version of NASA TLX that eliminates the weighting of dimensions.	Enables faster evaluation of cognitive load with reduced complexity.	Quickly assessing cognitive demands of tasks in usability studies.	Measuring the cognitive load users experience when navigating new mobile layouts, such as multilayer menus or filtering team data.	Georgsson et al., 2019
Error rate	Measures the frequency of user mistakes during tasks.	Identifies usability issues that lead to task failures and frustration.	Identifying specific interaction points where users commonly face errors.	Identifying errors when users interact with new features, such as the updated search system or real-time filtering on the mobile platform.	Zheng et al., 2022
Heart Rate Variability (HRV)	Measures stress or comfort during interactions.	Provides physiological feedback on user experience.	Measuring emotional response during interaction with high-stress or time-sensitive features.	Monitoring stress levels during live match updates or data filtering on Zos mobile.	Nakai et al., 2024; Zaki et al., 2021
Retention rate	Tracks how often users return to the platform over time.	Indicates long-term user engagement and satisfaction.	Measuring user loyalty and identifying what keeps users coming back.	Evaluating user retention rates based on recurring visits to explore sports news or team updates on Zos.	Moran et al., 0000

Table 3.5: Metrics for user experience evaluation *(part 1)*

Metrics	Features	Relevance	Application (Geral)	Application (Zos)	Refs
Gaze Density (Eye Tracking)	Measures user focus and attention areas on interfaces.	Identifies areas of visual interest or confusion.	Assessing visual engagement and optimizing content placement on screens.	Identifying areas where users focus most when exploring statistics and dynamic content on Zos mobile.	Nakai et al., 2024; Zaki et al., 2021; Khairat et al., 2019
Task Load Index	Measures task-related effort and performance metrics.	Highlights task difficulty and usability improvements.	Evaluating user effort in completing tasks and identifying potential inefficiencies.	Measuring effort required to interact with filters and data comparisons on Zos mobile.	Salvadorinho et al., 2024
Bounce Rate	Measures the percentage of users leaving after viewing a single page.	Reflects engagement and content relevance issues.	Monitoring user disengagement and identifying problematic pages.	Monitoring pages where users disengage, such as after viewing a live score page on Zos mobile.	Moran et al., 0000
GQM (Goal Question Metric)	A structured framework to define and assess metrics that align with goals and key questions.	Helps link metrics directly to project objectives, ensuring focused evaluation.	Developing custom metrics for projects, aligning them with organizational goals.	Defining goals for the new mobile platform and evaluating metrics for content accessibility and user satisfaction.	Bokingkito et al., 2022; Weber et al., 2019
Performance Metrics	Includes "Wait Time" and "Cycle Time" to assess system performance.	Provides critical insights into system efficiency and user satisfaction by measuring delays or bottlenecks that can affect the overall experience.	Monitoring system speed and responsiveness under varying loads.	Evaluating response times for live match updates and real-time data comparisons on the mobile platform.	Angulo et al., 2024
ECG	Measures electrical activity of the heart to assess stress or comfort during interactions.	Provides insights into stress responses during critical user interactions.	Tracking user stress levels while navigating complex live scores or rapidly changing data dashboards.	Monitoring physiological stress during high-pressure scenarios, such as navigating rapidly changing live score updates or managing complex filtering in real-time.	Zaki et al., 2021
Facial expression/emotion analysis	Tracks facial expressions to assess emotional responses during interactions.	Identifies user satisfaction, frustration, or delight with interface elements.	Evaluating user satisfaction, frustration, or emotional responses to interactive elements or workflows during usability testing.	Identifying emotional reactions to specific features, like interactive filters or visually dynamic data presentations.	Zaki et al., 2021
Electrodermal activity (EDA)	Measures skin conductance to evaluate physiological arousal and emotional intensity.	Captures stress or engagement during critical interactions.	Monitoring user stress and engagement levels during high-pressure tasks or complex interactions in usability studies.	Monitoring physiological arousal during critical interactions, such as team comparisons or data filtering.	Zaki et al., 2021
Perceived Usability	Captures user perceptions of ease of use and overall system intuitiveness.	Provides subjective feedback on the usability of the new interface.	Capturing subjective feedback on ease of use and system intuitiveness.	Gathering user insights on the intuitiveness of the redesigned navigation system and filters on Zos mobile.	Salvadorinho et al., 2024

Table 3.6: Metrics for user experience evaluation *(part 2)*

3.3.3 UX Evaluation Instruments

The following tables 3.7 and 3.8 highlights the tools and instruments employed in UX evaluation. These include both traditional methods like think-aloud protocols and modern approaches such as biometric sensors and eye tracking.

Instruments	Features	Relevance	Application (Geral)	Application (Zos)	Refs
Interaction recording tools	Hotjar, Crazy Egg	Records user paths, clicks, scrolls, and time spent on sections.	Observing user navigation and identifying usability bottlenecks.	Tracking how users navigate team pages, live scores, and sports statistics on Zos mobile.	Namoun et al., 2024
Think-Aloud Protocol	UC-CW; Cognitive Walkthrough	Encourages users to verbalize thought processes while performing tasks.	Identifying usability challenges and gathering qualitative insights on user behavior.	Evaluating how users articulate difficulties when searching sports data or navigating pages on Zos mobile.	Baciková et al., 2021; Weichbroth, 2020; Martinelli et al., 2024; Bitkina et al., 2020; Alonso-Virgós et al., 2020; Maramba et al., 2019; Georgsson et al., 2019
User Questionnaires	SUS, UEQ (User Experience Questionnaire), UEQ-S (User Experience Questionnaire - Short), SUPR-Q, AttrakDiff, Mojoleaf, Lemtool, Reaction Cards (Benedek e Miner, Barnum e Palmer, Merčun, etc.), Questionnaire for User Interaction Satisfaction (QUIS / SUS survey variable); TAM	Collects subjective user feedback through structured surveys.	Measuring user satisfaction, usability perceptions, and emotional responses.	Capturing feedback on ease of use and satisfaction with navigation and content in Zos mobile.	Magableh et al., 2024; Martinelli et al., 2024; Kuhar et al., 2022; Veral et al., 2019; Adams et al., 2019; Asemi et al., 2022; Yuhui et al., 2024; Maramba et al., 2019; Khairat et al., 2019; Alvarez et al., 2024
Likert Scale	Questionnaires, mentioned in User Questionnaires	Collects subjective feedback using a structured scale (e.g., 5 or 7 points). Measures attitudes or satisfaction levels.	Used in surveys or questionnaires to evaluate user satisfaction, usability, or perception.	Can be influenced by response bias; interpretation may vary among participants.	Nakai et al., 2024
Emotion Word Prompt List (EWPL)	Encourages users to associate interfaces or design elements with emotional adjectives from a predefined list.	Capturing emotional responses to assess the perceived desirability or emotional appeal of interfaces.	Measuring user engagement, navigation efficiency, and content effectiveness across digital platforms.	Understanding user emotions and preferences related to visual layouts or navigation systems in Zos mobile.	Kuhar et al., 2022

Table 3.7: Instruments for user experience evaluation *(part 1)*

Instruments	Features	Relevance	Application (Geral)	Application (Zos)	Refs
Heuristic Checklists for HE	E.g.: Nielsen (10 heuristics); HESC - Heuristic Evaluation System Checklist; HRI Heuristic Validation; etc	Guides expert evaluations of usability issues.	Identifying usability problems through structured evaluations.	Evaluating Zos mobile design consistency and alignment with usability principles.	Abulfaraj et al., 2020b; Abulfaraj et al., 2020a; Weber et al., 2019; Saavedra et al., 2019; Quiñones et al., 2019; Bitkina et al., 2020; Nowakowski, 2020; Martins et al., 2024; Krawiec et al., 2020; Alonso-Virgós et al., 2020; Angulo et al., 2024; Capeleti et al., 2020; Baciková et al., 2021; Weichbroth, 2020; Sahua et al., 2020; Martinelli et al., 2024; Asemi et al., 2022; Nasir et al., 2022; Maramba et al., 2019; Fernández et al., 2021
Biometric Sensors	Heart rate sensors ECG (Electrocardiograms. Electrocardiography), Eye trackers; BITalino System; ProComp Infiniti System; NeuroSky MindWave device; Emotiv EPOC	Monitors physiological responses such as stress, focus, and engagement.	Providing real-time feedback on emotional and cognitive load during interactions.	Analyzing stress or focus levels when interacting with live matches or data filters on Zos mobile.	Nakai et al., 2024; Zaki et al., 2021; Alonso-Virgós et al., 2020
Video Recording	Lookback, Dovetail	Captures user interactions with the system, allowing for replay and analysis.	Reviewing task flows and user behavior for usability problem detection.	Recording user behavior during sports data retrieval or interaction with key features on Zos mobile.	Angulo et al., 2024
Analytics Platforms	Google Analytics, Piwik	Tracks user behavior metrics such as clicks, page views, and session duration.	Analyzing user behavior trends and identifying patterns of engagement or disengagement.	Measuring engagement on key pages, such as team profiles, live scores, and statistics pages on Zos mobile.	Moran et al., 0000
Card Sorting	OptimalSort, XSort	Organizes information based on user mental models. Types: Open: Participants create their own categories. Closed: Participants sort items into predefined categories. Hybrid: Combines open and closed methods.	Improves Information Architecture (IA). Identifies naming/labeling issues for items or categories. Enhances navigation and user experience.	Helps refine navigation and terminology in the Zos mobile version by allowing users to organize content like team stats, players, and live scores into intuitive categories, either freely (open sorting) or within predefined options (closed sorting).	Moran et al., 0000

Table 3.8: Instruments for user experience evaluation *(part 2)*

Instruments	Features	Relevance	Application (Geral)	Application (Zos)	Refs
Eye Tracking Devices	Tobii TX300, EyeLink	Tracks gaze points and heatmaps to evaluate visual engagement.	Identifying visual attention hotspots and evaluating interface design.	Analyzing focus areas for live score updates, team data, or statistics displays on Zos mobile.	Nakai et al., 2024; Zaki et al., 2021; Martinelli et al., 2024
Remote Testing Tools	UserTesting, UserZoom, Userlytics, Lookback	Allows usability testing with remote participants through screen sharing and recordings.	Expanding usability testing to geographically dispersed user groups.	Conducting remote usability tests to evaluate mobile experience with sports enthusiasts across regions.	Namoun et al., 2024
Workshops	—	Collaborative sessions where stakeholders and users work together to brainstorm and solve design challenges.	Generating ideas for user-centric solutions and validating interface concepts.	Brainstorming user needs for new mobile features and co-creating solutions for better usability on Zos mobile.	Cajander et al., 2022
Prototyping	Wireframes (Figma); Uizard, v0, Visily	Tools for creating interactive or static representations of an interface for early testing and feedback.	Validating design concepts and gathering feedback before full development.	Testing and iterating on new navigation flows or data presentation designs in Zos mobile version.	Martinelli et al., 2024; Namoun et al., 2024
UX Scorecards	Google HEART Framework	Framework for systematically evaluating the UX performance of a platform across defined dimensions.	Monitoring progress in UX improvements over time and benchmarking against goals.	Tracking improvements in navigation efficiency, task completion rates, and satisfaction levels after mobile updates.	Martinelli et al., 2024; Moran et al., 0000
Data Analysis Tools	Mamdani Fuzzy Logic Model, Google Analytics, Adobe Analytics, or Pendo	Capture behavioral data (clicks, time on page, navigation flows).	Used to understand user interaction with digital products, monitor traffic, and engagement.	Used to capture how users interact with sports statistics and improve the experience with the content provided in the mobile version.	Asemi et al., 2022; Moran et al., 0000

Table 3.9: Instruments for user experience evaluation *(part 3)*

3.3.4 Application in this research

Taking into account the tables presented in Sections 3.3.1, 3.3.2 and 3.3.3, which outline the evaluation methods, metrics, and instruments, the following methods were applied in this research:

- Questionnaires (SUS)
- Focus Groups
- Hotjar

These methods are particularly aligned with the *Strategize* phase, as demonstrated in Table 3.2, discussed in subsection 3.3.1. Their suitability stems from their ability to gather qualitative

and attitudinal data, which is essential to identify new directions and opportunities in the context of the Zos mobile version.

3.3.4.1 Questionnaires (SUS)

Questionnaires are attitudinal and quantitative methods, positioned in the bottom-right quadrant of figure 3.2 (*"What people say" quantitatively*). Among these, the System Usability Scale (SUS) is a standardized usability evaluation method that derives a usability score from user responses. This method provides a foundational understanding of the user experience, which is then enriched and complemented by the qualitative insights gathered through focus groups. The combination of these approaches ensures that user feedback is captured both quantitatively and qualitatively, offering a holistic perspective on the platform's usability.

3.3.4.2 Focus Groups

Focus groups are attitudinal and qualitative methods, positioned in the bottom-left quadrant of figure 3.2 (*"What people say"*). They are ideal for exploring user opinions, preferences, and expectations. Focus groups allow for rich discussions among participants, capturing diverse perspectives on the usability and features of the mobile platform. This method aligns with the *Strategize* phase outlined in the SOTA, which involves generative methods to identify new opportunities and directions for improvement.

3.3.4.3 Hotjar

Hotjar is a behavioral and quantitative tool, positioned in the top-right quadrant of figure 3.2 (*"What people do" quantitatively*). It tracks user interactions such as clicks, scrolls, and navigation patterns, providing real-time data on how users engage with the mobile version of ZeroZero. By identifying usability bottlenecks and interaction trends, Hotjar offers valuable insights into user behavior. This method complements the qualitative findings from focus groups and the usability assessment provided by SUS, contributing to a comprehensive understanding of the user experience.

3.3.4.4 Methods combination

The integration of focus groups, SUS evaluation, and Hotjar ensures a comprehensive assessment of the user experience by addressing different dimensions:

- **Attitudinal Insights:** Focus groups provide in-depth exploration of user opinions, emotions, and expectations.
- **Usability Assessment:** The SUS serves as a standardized method for measuring perceived usability and user satisfaction.

- **Behavioral Data:** Hotjar captures real-time user interactions, revealing navigation patterns and potential usability challenges.

A more detailed explanation of how these methods will be implemented, including precise and practical applications, is provided in section 4.

3.4 Summary

This chapter provided an in-depth exploration of the theoretical foundations and current practices in UX evaluation, with a focus on mobile platforms. The main takeaways are as follows:

- **Theoretical Framework** *(Section 3.1)*

  UX evaluation is rooted in principles of Human-Computer Interaction (HCI), emphasizing usability dimensions such as effectiveness, efficiency, and satisfaction.

  Established standards, such as ISO 9241-11 and Nielsen's heuristics, provide a robust foundation for usability evaluation.

- **Mobile Optimization and Usability** *(Section 3.2)*

  Mobile optimization significantly enhances user satisfaction by addressing the constraints and opportunities of mobile interfaces.

  Jakob Nielsen's guidelines highlight the importance of mobile-specific design practices, including simplified navigation and reduced content complexity.

- **UX Evaluation: Methods, Metrics, and Tools** *(Section 3.3)*

  Diverse methods, including surveys, usability testing, and behavioral analysis, enable a comprehensive understanding of user interactions (Rohrer, 2022; Nielsen; Budiu, 2012).

  Metrics like task completion time, error rate, and user satisfaction are widely adopted to quantify UX performance (Nielsen, 1993; Khairat et al., 2019).

  Tools such as SUS and UEQ provide standardized approaches for collecting user feedback, while advanced methods like eye tracking and psychophysiological measures offer emerging possibilities for deeper insights (Weichbroth, 2020; Zaki et al., 2021).

The insights from this chapter serve as a foundational reference for evaluating the Zos mobile platform. They guide the research design and the choice of evaluation methods and tools, supporting a structured and user-centered analysis.

Chapter 4

Methodology

4.1 Introduction

This chapter provides a detailed explanation of the methodological approach used in this study to explore the perceptions and experiences of users regarding the new mobile version. It begins by discussing the research paradigm that underpins the study, followed by an explanation of the mixed methods approach adopted to ensure a comprehensive analysis. The chapter then outlines the specific data collection techniques, including surveys and focus groups, and discusses their relevance and justification. Additionally, the chosen research methodology (case study) is elaborated upon, with a focus on its suitability for this research context. Finally, the research population, sampling strategy.

4.2 Research Paradigm

The research paradigm adopted for this study is 'Interpretivism'. This paradigm is particularly suited to the objectives of the study as it takes into account the subjective nature of human experiences and the social context in which they occur. Since the aim is to understand the perceptions and experiences of users regarding the new mobile version, the interpretivist paradigm provides the appropriate lens to explore these aspects in depth. By focusing on the meanings and insights derived from participants, this paradigm facilitates a nuanced understanding of user interactions and feedback.

As noted by Burrell et al., 2019, the interpretivist paradigm seeks to understand the world from the point of view of the individuals directly involved in the social process, emphasizing the subjective experiences and meanings that individuals assign to their actions and interactions.

> *"The interpretive paradigm is informed by a concern to understand the world as it is, to understand the fundamental nature of the social world at the level of subjective experience. It seeks explanation within the realm of individual consciousness and*

> *subjectivity, within the frame of reference of the participant as opposed to the observer of action."* (Burrell et al., 2019 [1])

4.3 Research Methodology

The chosen research methodology for this study is the case study. This approach is particularly effective for gaining an in-depth understanding of a specific phenomenon within its real-life context. In this instance, the case study focuses on the adoption and usage of the new mobile version, allowing for detailed insights into user interactions and behaviors. Using this methodology, the study aims to identify actionable recommendations that address user needs and improve the usability of the platform. As observed by Yin, 2008, the case study methodology excels in investigating modern phenomena, particularly when the context and the phenomenon itself are intricately linked and not easily separable.

> *"The case study is preferred in examining contemporary events, but when the relevant behaviors cannot be manipulated. (...) the case study's unique strength is its ability to deal with a full variety of evidence—documents, artifacts, interviews, and observations—beyond what might be available in historical research."* (Yin, 2008 [2]).

4.4 Plan Experiment

4.4.1 Approach

The proposed approach integrates a mixed-method strategy to evaluate the user experience of the new mobile version of the Zos platform. Combining quantitative surveys, qualitative focus groups and behavioral analysis using tools such as Hotjar allows for a more complete and nuanced understanding of user perceptions and interactions.

The originality of this approach lies in:

- Addressing usability gaps identified in the literature.
- Proposing improvements over existing solutions based on real user feedback.
- Balancing breadth (through questionnaires), depth (through focus groups), and behavioral insights (through Hotjar) in the analysis.

This approach aims to support the research objectives of enhancing user satisfaction and improving navigation efficiency.

[1] *page 28*
[2] *page 39*

4.5 Methods

4.5.1 Purpose

The purpose of this study is to evaluate user perceptions and experiences regarding the new mobile version of the Zos platform. By identifying usability challenges and gathering user feedback, the research aims to propose actionable improvements that enhance the overall user experience, navigation efficiency, and satisfaction. The methodology aligns with the research question and addresses the gaps identified in the literature.

4.5.2 Research Population

The target research population of this study consists of Portuguese users of the ZeroZero.pt platform who access the website through mobile devices. The study deliberately focuses on this segment to ensure the relevance of feedback to the specific context of the mobile version's usability. Desktop users and users accessing from other countries were excluded from the study's scope, as the primary goal was to evaluate the mobile experience of local users.

4.5.3 Sample

A non-probabilistic, convenience sampling strategy was adopted. Participants were not randomly selected but instead voluntarily responded to the questionnaires made available directly on the mobile version of the ZeroZero.pt website (figure 4.1).

Figure 4.1: Mobile interface banner inviting users to participate in the study

The questionnaires were presented in two distinct ways, depending on the user's login status:

- **Questionnaire 1 – Registered Users:** Displayed only to logged-in users. It was made available from April 29 to May 12, 2025, resulting in a total of 85 valid responses.
- **Questionnaire 2 – Non-Registered Users:** Displayed to anonymous users (not logged in) on the mobile homepage. It was launched on May 27, 2025 and remained open until June 9, 2025. The collected responses are included in the overall analysis presented in this dissertation.

While this sampling approach does not provide statistical generalizability, it enables the inclusion of a wide range of mobile users and supports exploratory and comparative analysis. However,

potential limitations related to sampling bias, voluntary participation, and unequal group sizes are acknowledged and further discussed in subsection 4.5.5.

4.5.3.1 Limitations of the Sample

Despite the contextual richness of the sample, the study faces certain limitations:

- **Voluntary Participation:** The open distribution of the questionnaire means participants are self-selected, which may lead to overrepresentation of more motivated users.
- **Limited Access to Final Version:** Depending on the deployment stage of the mobile platform, users may still be interacting with an intermediate version.

4.5.4 Research Approach and Data Collection Techniques

To ensure a comprehensive and robust analysis, this study adopts a mixed-methods approach, integrating both qualitative and quantitative methodologies. In light of this paradigm, the data collection techniques employed include questionnaires (Appendix B) , focus groups (Appendix C), and the Hotjar UX evaluation tool.

4.5.4.1 Questionnaires

The quantitative component of the study relies on surveys to gather a broad overview of user opinions. Questionnaires allow for the collection of structured and measurable data, making it possible to identify trends and patterns across a large user base. This technique is particularly valuable for quantifying user satisfaction and pinpointing common challenges with the new mobile version.

To assess usability and user experience, the SUS will be employed (Yuhui et al., 2024 and Maramba et al., 2019). Alternative instruments such as the User Experience Questionnaire (UEQ) (Alvarez et al., 2024) and AttrakDiff[3] were considered during the research planning phase, but ultimately not included in the final questionnaire in order to maintain a concise and focused survey experience for participants. The decision to use only the SUS was based on its suitability for the study's objectives and the need to balance depth of insight with participant effort.

To ensure participants can respond comfortably and accurately, the questionnaires will be administered in Portuguese. This decision aligns with the considerations outlined in subsection 4.5.5, where potential challenges related to language are addressed.

4.5.4.2 Focus Groups

The qualitative aspect of the study involves conducting focus groups [4] to delve deeper into users' experiences. This technique is ideal for exploring nuanced opinions, uncovering challenges, and

[3] AttrakDiff: Website
[4] Focus Groups: NNG Focus Groups Video

gathering detailed suggestions. This will provide a collaborative environment that fosters open discussion, enabling participants to share insights that might not surface in a questionnaire setting. This qualitative data complements the quantitative findings, offering a richer and more detailed understanding of user experiences.

Deviations from the original plan

Originally, a traditional focus group was planned, with a discussion guide designed to encourage participant interaction and co-construction of insights. However, on the scheduled date for the first group, only two of the four confirmed participants attended. For the second session, two participants expressed interest, but only one ultimately participated. Consequently, both sessions were conducted as structured interviews, following the original focus group script. Although this format change reduced the interactive dynamics characteristic of focus groups, it still yielded rich, in-depth qualitative insights. This deviation is acknowledged in the limitations section (Subsection 4.5.5) and has been taken into account in the interpretation of the results.

4.5.4.3 Hotjar Analysis

As part of the data collection process for the first questionnaire (targeted at registered users), the Hotjar tool[5] was used to observe real user interactions with the mobile version of the platform. Hotjar offers features such as heatmaps, session recordings, and user feedback forms, which allow researchers to visually track navigation behaviors, identify points of confusion, and detect usability bottlenecks.

In this study, the tool was employed during a limited period under a free trial plan, which restricted the quantity and duration of recorded sessions. Due to these limitations, Hotjar was not available during the second questionnaire aimed at non-registered users. Nevertheless, the data gathered during the first phase offered valuable behavioral context, complementing the structured results from the questionnaire and the qualitative insights from focus groups.

Although not comprehensive, this behavioral layer helped illustrate real-time interaction patterns and supported the triangulation of findings across methods.

4.5.5 Limitations

The study acknowledges several limitations:

- **Sample Size:** Although purposive sampling ensures relevance, the number of focus group participants might not fully represent all user perspectives.
- **Response Bias:** Questionnaire responses might be influenced by personal opinions or social desirability.
- **Scope:** The research focuses on the Zos platform, limiting generalization to other contexts.

[5]Hotjar: https://www.hotjar.com/

- **Limited Focus Group Execution:** Despite careful planning, both of the scheduled focus group sessions experienced low attendance (only two participants attended the first session and just one the second). As a result, these sessions were conducted as structured interviews using the original discussion guide. While this adaptation allowed for the collection of relevant qualitative data, the absence of group dynamics may have limited the depth of interaction, spontaneous idea exchange, and the emergence of contrasting perspectives typically expected in a focus group setting.

These limitations will be mitigated through careful sampling, transparent reporting, and triangulation of data sources (combining quantitative and qualitative findings).

4.5.6 Ethical Considerations

This study adheres to rigorous ethical research standards to ensure the rights, dignity, and well-being of all participants are safeguarded. The following principles guide the research process:

- **Informed Consent:** All participants will receive clear and detailed information about the purpose, methods, expected outcomes, and use of the collected data. The consent process will ensure participants fully understand their rights, including the ability to withdraw at any time without justification. Written consent forms will explicitly outline these conditions and include a statement about data confidentiality and anonymity.
- **Anonymity and Confidentiality:** Participant identities will be anonymized through data processing techniques to prevent any direct or indirect identification. In accordance with the GDPR [6], personal data will be collected, processed and stored responsibly. All participant data will be anonymized to ensure that no individual can be directly or indirectly identified. This approach safeguards privacy while maintaining compliance with data protection regulations.
- **Transparency and Access to Results:** Participants will be informed of how the results will be shared and have the opportunity to access the findings upon study completion. This includes an option to receive an email with the study's conclusions and a link to the repository where the results are published. Participants will also be assured that their responses will be used solely for academic purposes. This commitment enhances transparency and ensures participants are kept informed of the study's outcomes.
- **Reversibility of Consent:** Participants retain the right to withdraw their consent at any time, even after their participation, unless their data has already been anonymized and integrated into the analysis. This ensures respect for their autonomy throughout the research process.

In all aspects of communication and data collection, the principles of clarity, simplicity, and objectivity will be prioritized to ensure that participants are fully informed and their participation is both voluntary and conscious.

[6] General Data Protection Regulation

4.5.7 Mitigation Plan

To address potential challenges and limitations, the following mitigation strategies will be implemented:

- **Linguistic Barriers:** The research involves participants whose primary language is Portuguese, while the thesis is written in English. To mitigate potential misinterpretations and ensure participants can express themselves freely, all questionnaires and focus groups will be conducted in Portuguese. The collected data will later be translated into English, which may introduce subtle nuances or variations in meaning despite efforts to preserve the original intent. [7]
- **Data Reliability:** Ensuring the clarity and consistency of data may be challenging. To address this, questionnaires will be pre-tested, and qualitative insights from focus groups will be triangulated with quantitative survey data to enhance the reliability and robustness of findings.
- **Participant Fatigue:** Extended focus group discussions and lengthy questionnaires may lead to participant fatigue, potentially affecting the quality of responses. To minimize this, focus groups will be limited to 60–90 minutes, and questionnaires will be designed to be concise and focused.

These measures will help maintain the rigor, transparency, and validity of the research findings.

4.5.8 Expected Results

The expected results include:

- Identification of key usability challenges that users face with the mobile version.
- Quantitative understanding into satisfaction levels, navigation clarity, and areas needing improvement.
- Qualitative findings that provide actionable recommendations to improve the user experience.

The study aims to suggest potential design improvements that address the identified issues and enhance the usability of the platform.

[7] To ensure transparency and maintain the integrity of the data, all original questionnaires and focus group transcriptions in Portuguese, as well as their translated versions in English, are included in the appendices. This approach aims to address potential concerns regarding the nuances of translation and to facilitate access for readers fluent in both languages.

4.6 Project Plan

To ensure an organized execution of the study, a Gantt chart was developed to illustrate the project timeline, task dependencies, and milestones. This chart provides a high-level overview of the key phases of the project, as depicted in Figure 4.2.

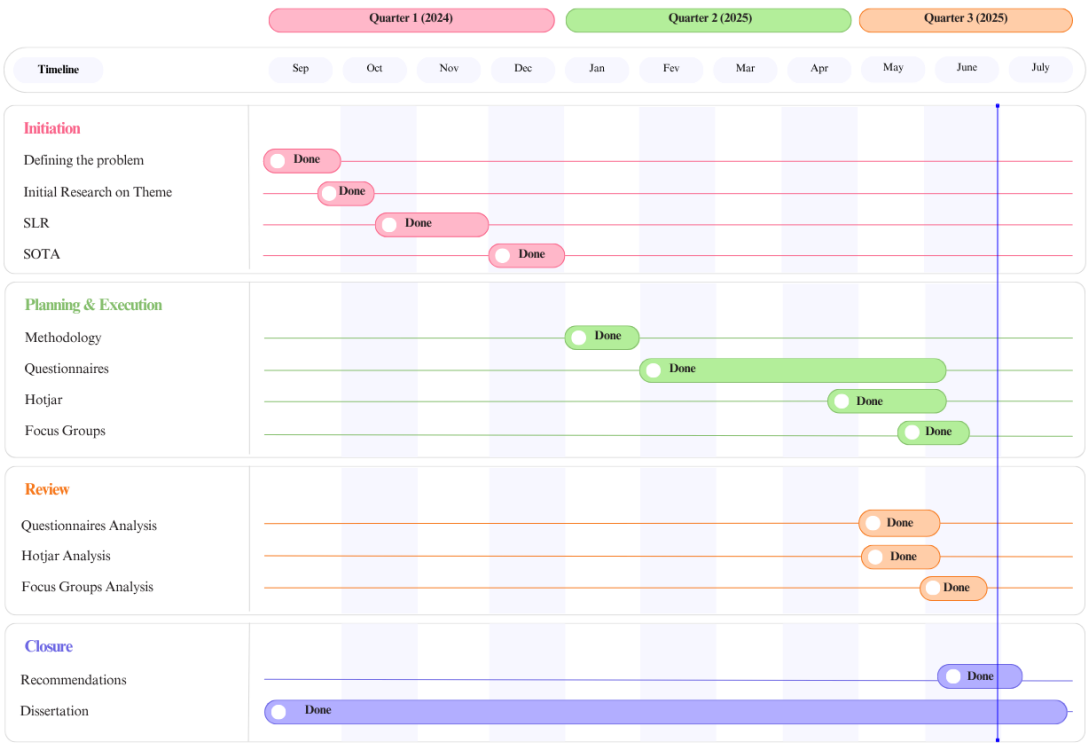


Figure 4.2: Gantt Chart for the project plan's activities

The project is structured into four main phases, each with specific milestones and deliverables:

1. **Initiation:** This phase laid the foundation for the study, focusing on problem definition and early exploratory research. The key milestones and deliverables included:

   - Problem definition and scope clarification. *(Milestone)*
   - Initial research on the theme, including the Systematic Literature Review (SLR) and State of the Art (SOTA). *(Milestone)*
   - Completion of the Introduction (Chapter 1), Literature Review (Chapter 2), and SOTA analysis (Chapter 3). *(Deliverable)*

   Status: Completed in Q1 (2024).

2. **Planning & Execution:** This phase involved the preparation of research instruments and subsequent data collection. It was carried out across Q2 (2025) and Q3 (2025). The key milestones and deliverables included:

- Development of the research methodology (Chapter 4). *(Deliverable)*
- Design and implementation of questionnaires. *(Milestone)*
- Request and coordination of the Hotjar setup with the ZeroZero development team, followed by usage of the collected data during the first questionnaire phase. *(Milestone)*
- Preparation and facilitation of focus groups. *(Milestone)*

Status: Completed.

3. **Review:** This phase focuses on the analysis of collected data. It includes:

   - Analysis of questionnaire responses. *(Milestone)*
   - Analysis of behavioral data from Hotjar. *(Milestone)*
   - Thematic analysis of focus group discussions. *(Milestone)*

   Status: Completed (Q3 2025).

4. **Closure:** The final phase includes the synthesis of findings and the consolidation of final outputs. While the main analyses and recommendations were concentrated in Q3 (2025), the writing of the dissertation was carried out continuously throughout the project. Key deliverables include:

   - Formulation of UX recommendations based on findings. *(Deliverable)*
   - Writing and submission of the dissertation. *(Deliverable, developed progressively across all phases).*

   Status: Completed.

The primary goals and deliverables of the study are as follows:

- Validated data collection instruments (e.g., questionnaires, focus group guides).
- Analysis of behavioral data collected via Hotjar (previously set up by the ZeroZero team).
- Complete analysis of quantitative, qualitative, and behavioral data.
- A set of actionable UX recommendations to enhance the ZeroZero mobile experience.
- Final written dissertation.

4.7 Summary

This chapter outlined the methodological approach adopted for evaluating the user experience of the Zos mobile platform. The main takeaways from this chapter, which directly inform the next phase of the research, are as follows:

- **Research Paradigm and Methodology** *(Subsection 4.2 and 4.3)*

  The interpretivist paradigm and the case study methodology were chosen to explore user perceptions and behaviors in depth. These foundations ensure that the collected data reflects the real-world context of ZeroZero's mobile platform and provides actionable insights.

- **Mixed-Methods Approach** *(Subsection 4.4.1)*

  By combining quantitative (e.g., SUS questionnaires) and qualitative (e.g., focus groups) methods with behavioral analysis (e.g., Hotjar), the study ensures a comprehensive evaluation of usability and user satisfaction. These methods complement each other and provide a multidimensional understanding of the platform's strengths and areas for improvement.

- **Data Collection Tools** *(Subsection 4.5.4)*

  The tools and techniques (questionnaires, focus groups, and behavioral analysis) were selected to capture diverse aspects of user experience, from subjective satisfaction to objective navigation patterns. This diverse dataset will be crucial in identifying usability gaps and proposing targeted recommendations.

- **Ethical Considerations and Limitations** *(Subsection 4.5.6)*

  The ethical framework ensures participant privacy and informed consent, enhancing the reliability and credibility of the findings. Acknowledging the limitations, such as sample size constraints and potential biases, strengthens the transparency and validity of the study.

This summary provide the groundwork for the analysis and discussion presented in the next chapter. The data collected through these methodologies will be analyzed to identify key usability challenges, user behavior patterns, and actionable improvements for the Zos mobile platform. This analysis will form the core of Chapter 6, where the results are interpreted and contextualized to address the research objectives.

Chapter 5

Design and Operationalization of Data Collection Methods

The development of the UX evaluation form for the mobile version of ZeroZero involved multiple stages of discussion, refinement, and methodological considerations. This section outlines the key decisions, dilemmas, and rationale behind the implementation of the chosen methodology (Chapter 4).

5.1 Designing the UX Evaluation Questionnaire

5.1.1 Conceptualization and Initial Considerations

Before constructing the form, it was essential to define the core objectives of the evaluation. The primary focus was to assess usability, accessibility, and overall satisfaction with the mobile platform. Several initial considerations were discussed, including:

- Question Format: Whether to use open-ended, multiple-choice, or Likert scale questions.
- User Demographics: Collecting basic demographic data to characterize the participant profile.
- Evaluation Criteria: Establishing relevant aspects of usability to assess, including navigation ease, content clarity, and overall satisfaction.

5.1.2 Challenges and Dilemmas in Form Construction

Throughout the development process, several dilemmas emerged, necessitating iterative refinement:

- Balancing Depth and Brevity: A more comprehensive form provides richer insights but risks user fatigue and reduced completion rates.

- Subjectivity vs. Objectivity: Some qualitative aspects of UX require subjective user input, whereas quantitative measures enhance comparability.
- Context-Specific Questions: Ensuring questions were tailored to the ZOS mobile platform while maintaining generalizability to UX principles.

5.1.3 Methodological Implementation

The final structure of the form was shaped through iterative testing and expert feedback.

The methodology included:

- Pilot Testing: Conducting a preliminary round of data collection to refine question clarity and relevance.
- User-Centered Design Principles: Incorporating usability principles to enhance user engagement with the form.
- Mixed-Methods Approach: Combining qualitative and quantitative questions to provide a holistic evaluation.

5.1.4 Questionnaire Overview and Refinement Process

The final version of the questionnaire was designed to evaluate multiple dimensions of user experience with the ZeroZero.pt mobile website. The form included structured tasks, demographic profiling, behavioral assessments, and standardized usability metrics. It was developed with a mixed-methods perspective to collect both quantitative and qualitative data.

5.1.4.1 Structure and Content

The questionnaire comprised the following key sections:

- **Informed Consent and Study Context:** Participants were introduced to the study's objectives, confidentiality guarantees, and GDPR-compliant data handling.
- **User Profile and Usage Context:** Demographic and behavioral data, including age, field of interest/study, frequency of use, topics accessed, and device type.
- **Task-Based Usability Evaluation:** Participants completed five real-world tasks (e.g., finding match results, player statistics, and accessing standings or news). For each, users reported success, time spent, and perceived difficulty.
- **Open-Ended Questions:** Included prompts for frustrations, suggestions, and general feedback.
- **System Usability Scale (SUS):** A full SUS questionnaire with 10 items, rated on a 5-point Likert scale.

- **Structured Interview Opt-In:** Participants could indicate willingness to participate in a follow-up qualitative session.

5.1.4.2 Pilot Test and Iterative Refinement

Before being distributed publicly, the questionnaire was tested with 21 internal collaborators from ZOS. Based on their feedback, several modifications were made:

- Clarified task instructions, especially for navigation and search actions.
- Differentiated between academic degree and field of study.
- Adjusted the SUS scale to ensure single-answer selection per item.
- Included missing options and revised terminology for clarity (e.g., replacing *None* with *I was not aware of it*).
- Ensured better consistency between questions and answer choices.

These refinements improved both the reliability and user-friendliness of the instrument, contributing to the overall methodological quality of the study.

5.2 Hotjar as a Behavioral Validation Tool

Hotjar is a behavior analytics tool that enables the collection of real-time user interaction data through heatmaps, session recordings, and click tracking. In this study, Hotjar will be implemented as a complementary method to validate findings from the questionnaire.

5.2.1 Objective of Implementation

The primary goal of using Hotjar is to compare users' self-reported experiences with their actual behavior on the ZeroZero mobile platform. Specifically, it aims to:

- Validate whether the patterns identified in the questionnaire align with real user interactions.
- Reduce potential bias, given that the questionnaire is answered by employees who may have a different perception compared to regular users.
- Provide an objective data source based on real user behavior rather than subjective assessments.

5.2.2 Methodology of Implementation

Hotjar will be applied to analyze user behavior across selected pages of the mobile platform. The key metrics to be observed include:

- Heatmaps: To identify the most and least engaged areas of the interface.
- Session Recordings: To track navigation patterns and points of friction.
- Click and Scroll Data: To assess interaction trends and detect usability issues.

Hotjar was activated in May and ran concurrently with the questionnaire distribution period. This simultaneous data collection allowed for a more integrated comparison between users' real-time behavior on the ZeroZero mobile platform and their self-reported experiences captured in the questionnaire responses.

5.2.3 Potential Limitations

While Hotjar is a valuable tool for behavioral analysis, it has some limitations:

- It does not capture subjective feedback, meaning qualitative insights from users are still necessary.
- It provides limited scope by focusing only on specific monitored pages.
- The sample of recorded users may not be entirely representative of the broader ZOS user base.

5.2.4 Conclusion

Hotjar serves as a crucial validation tool in this research, helping bridge the gap between user perceptions (via the questionnaire) and real-world interactions. By leveraging behavioral analytics, the study aims to refine UX recommendations and gain a more accurate understanding of user engagement on the platform.

5.3 Structured Interviews

Structured interviews are a qualitative research method used to collect detailed insights into users' perceptions, experiences, and expectations. In this study, they were employed to deepen the understanding of how users interact with the mobile version of ZeroZero, complementing the quantitative and behavioral data obtained through the questionnaire and Hotjar.

Unlike focus groups, structured interviews emphasize individual responses to a predefined set of questions, ensuring consistency across participants while allowing space for elaboration. This method is particularly useful when participant availability is limited or when group dynamics might introduce bias.

5.3.1 Objective of Implementation

The structured interviews aimed to:

- Explore users' experiences with the mobile version of ZeroZero, including navigation, usability, and satisfaction with key features.
- Contrast perceived user difficulties with actual behavioral patterns identified through Hotjar.
- To identify usability pain points and potential improvements, ensuring that UX recommendations align with actual user needs.

5.3.2 Participant Selection

Participants were recruited through the questionnaire, which included an opt-in question at the end regarding participation in a second phase of the study. Those who agreed were contacted via email to schedule a time based on availability.

Despite this approach, participation remained limited. In the first round, four individuals confirmed, but only two were able to attend due to last-minute constraints. In the second round, two more users expressed interest and were scheduled accordingly.

Given the reduced number of participants, the originally planned focus groups were reformulated as structured interviews. While a predefined question script was followed (Appendix C), participants were explicitly encouraged to interact, share experiences, and build upon each other's input — fostering moments of idea complementarity. This hybrid approach allowed for both individual reflection and spontaneous exchange, maintaining methodological structure while enriching the qualitative depth.

5.3.3 Methodology of Implementation

The structured interviews were conducted after the completion of the Hotjar data collection, allowing the discussion to be informed by real behavioral insights. The process followed these steps:

1. Data Collection

   Notes and recordings were used (with consent) to capture user responses for later analysis.

2. Structure

   Moderator-Led Discussion: The sessions will be moderated by the researcher, who will guide the conversation and ensure balanced participation from all attendees.

   Open-Ended Questions: The discussion will explore topics such as ease of navigation, mobile interface usability, and expectations for improvement.

   Observation of Non-Verbal Cues: Attention will be given to body language, hesitation, and emotional responses to better understand user frustrations and satisfaction levels.

5.3.4 Advantages of Using Structured Interviews

Structured interviews offered several methodological advantages in this context:

- **Depth of Insight:** The format allowed users to reflect in detail on their experiences without the influence of group dynamics.
- **Consistency:** The original discussion guide prepared for the focus groups was reused, ensuring uniformity across sessions.
- **Adaptability:** This approach was better suited to the limited availability of participants.

5.3.5 Limitations and Challenges

Despite their strengths, the structured interviews presented some limitations:

- **Limited Sample Size:** The small number of participants limits the generalizability of findings.
- **Hybrid Format Complexity:** Although the sessions were structured, the group setting allowed for interaction between participants. While this enriched the data through complementary perspectives, it also introduced variability in how strictly the script was followed.
- **Scheduling Constraints:** Coordinating availability was challenging, and some expected participants were unable to attend due to last-minute issues.

5.3.6 Conclusion

The structured interviews provided a valuable qualitative complement to the data collected via the questionnaire and Hotjar. Initially planned as focus groups, the sessions were adapted due to limited participant availability. However, the group format was partially preserved by conducting interviews with two participants at a time, following a predefined script.

Although the sessions maintained a structured format, participants were explicitly encouraged to interact and comment on each other's ideas. This created moments of natural exchange and idea complementarity, enriching the discussion without compromising methodological consistency.

These interviews allowed for a deeper exploration of user expectations, frustrations, and suggestions, thereby enhancing the contextual understanding of behavioral and questionnaire data. As such, they played a key role in triangulating findings and shaping actionable UX recommendations for the mobile version of ZeroZero.pt.

5.4 Conclusion

This chapter presented the design and operationalization of the data collection methods used in this study. A mixed-methods approach was adopted to explore different dimensions of user experience

on the ZeroZero mobile platform, combining a structured questionnaire, behavioral analysis via Hotjar, and qualitative insights obtained through structured interviews.

Each method was selected to complement the others, acknowledging that no single technique could fully capture the complexity of UX. The questionnaire provided structured, self-reported data; Hotjar offered evidence of actual user behavior; and the interviews enabled a deeper exploration of user motivations, frustrations, and expectations.

To ensure the quality of the questionnaire, a pilot testing phase was conducted with internal collaborators from ZeroZero, focusing exclusively on refining question clarity, structure, and terminology. This step improved the reliability and user-friendliness of the instrument, without influencing the actual data collection or the validity of the responses gathered.

Hotjar, while valuable for behavioral validation, is restricted to selected pages and lacks access to subjective user intent. The originally planned focus groups had to be reformulated as structured interviews due to limited participant availability. Although some sessions included two participants simultaneously, the focus remained on individual responses, with occasional complementary exchanges enriching the discussion.

Despite these constraints, the triangulation of data collection methods strengthened the study's reliability and depth. The methodological choices reflect both practical limitations and a commitment to capturing diverse perspectives on mobile UX. The following chapter will analyze the collected data, identify usage patterns and pain points, and propose actionable recommendations to improve the mobile experience of ZeroZero.pt.

Chapter 6

Results

This chapter provides an overview of the data collection process, followed by a detailed analysis of both quantitative and qualitative data. It also synthesizes the findings, compares them to the initial hypotheses, and proposes actionable recommendations for improving the mobile user experience.

6.1 Overview of the Data Collection Process

This section outlines the different data collection methods employed in the study. It provides insights into the questionnaire process, stuctured interviews, and behavioral data collected through Hotjar, which served as the foundation for the analysis.

6.1.1 Questionnaire

The questionnaire aimed to gather quantitative data on user perceptions of the mobile version of the ZeroZero platform. It was designed to collect insights on usability aspects such as ease of navigation, design aesthetics, and functional performance.

To ensure relevance and depth, two separate questionnaires were deployed, each targeting a specific user group based on login status:

- **Registered Users:** The first questionnaire was made available exclusively to users with an active account and login. It was accessible from April 29, 2025, to May 12, 2025, and collected a total of 85 responses.
- **Non-Registered Users:** The second questionnaire targeted users who were not logged in, beginning on May 27, 2025, and remained open until June 9, 2025. This separation was intentional to prevent duplicated responses and to ensure a clearer distinction between user types.

This dual-approach allowed for a comparative analysis of the experiences reported by both registered and non-registered users. Nevertheless, it also introduces certain limitations related to the data collection timeline and the available analytical tools, which are further discussed in the limitations section (6.6).

6.1.2 Hotjar

Behavioral data from Hotjar was used to analyze user interactions with the mobile site. This included tools such as heatmaps, session recordings, and funnels, which provided valuable insights into how users navigate the platform, where they experience friction, and which elements attract the most attention.

The data was primarily collected during the distribution of the first questionnaire (April 29 to May 12, 2025), during which the Hotjar free trial was activated. This allowed for temporary access to advanced features like user funnels and session recordings, enriching the behavioral analysis with qualitative dimensions.

However, the second phase of the questionnaire (aimed at non-registered users) did not benefit from these extended capabilities, as the free trial had expired and no premium subscription was used. Consequently, only standard heatmaps and limited interaction tracking were available during this period, which constitutes a methodological limitation in terms of data consistency and comparability across user groups.

6.1.3 Structured Interviews

To complement the quantitative data, two structured interview sessions were conducted (one for each user group) aiming to gather deeper insights into participants' experiences and perceptions of the ZeroZero mobile platform.

The first session was organized for users who had completed the registered user questionnaire. Although four participants confirmed their attendance, only two ultimately participated in the session.

The second interview session targeted non-registered users, following the completion of the second questionnaire. In this case, two individuals also confirmed attendance, but only one participant was present at the scheduled time.

Despite the small number of participants in each session, the interviews provided valuable qualitative input that helped contextualize and interpret certain patterns observed in the questionnaire data.

6.2 Quantitative Data Analysis of Registered Users (Questionnaire 1)

This section presents the analysis of quantitative data collected from users who completed the platform login process and responded to Questionnaire 1. Descriptive and inferential statistical methods are used to interpret the data, providing a clear overview of user experience across various dimensions.

6.2.1 Data Preparation

Prior to analysis, the dataset was reviewed for quality and consistency. No incomplete responses were found; however, two participants provided open-ended comments for the perceived loading time question that did not correspond to a time-based evaluation. These responses were considered inconsistent with the question's intent and were recoded as system-missing values for that specific variable.

Likert scale responses were converted into numerical values to facilitate statistical analysis. This section describes the preparation procedures and any transformations applied to the data.

6.2.2 Descriptive Statistics

This subsection provides an overview of the participants' demographic and usage characteristics to contextualize the interpretation of usability and performance-related responses.

6.2.2.1 Participant Profile

The questionnaire collected responses from a diverse group of registered users, with varying backgrounds and usage patterns.

As illustrated in Figure 6.1, the majority of participants were between 25 and 54 years old: 25 respondents (29%) belonged to the 25–34 age group, 20 (24%) to the 35–44 age group, and 17 (20%) to the 45–54 group. Younger users accounted for 5 responses (6%) in the "Under 18" group and 17 responses (20%) in the 18–24 category. Only a very small number of participants were aged 65 or older (n = 1; 1%).

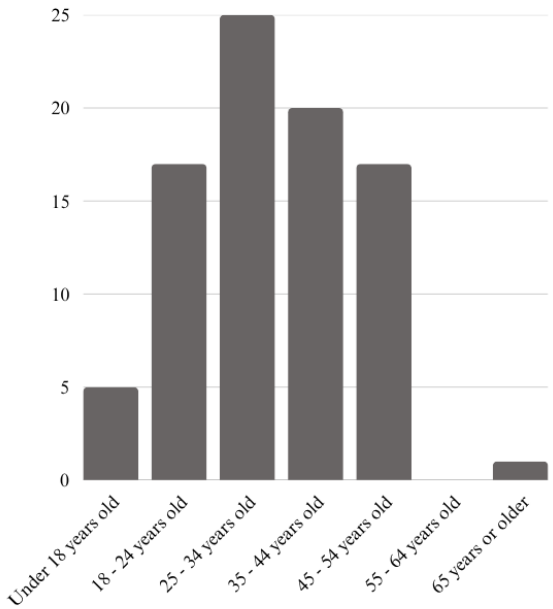


Figure 6.1: Age Distribution of questionnaire participants (Questionnaire 1)

This distribution suggests that the sample is primarily composed of adult users in active professional or academic stages, which aligns with the platform's typical audience. External demographic data supports this interpretation: according to SimilarWeb (May, 2025)[1], the largest age segment of ZeroZero.pt visitors is the 35–44 group (23.77%), closely followed by 25–34 (23.36%).

Regarding usage frequency, the majority of participants reported high levels of engagement with the platform. As illustrated in Figure 6.2, 66 respondents (78%) indicated that they visit ZeroZero.pt "several times a day," followed by 8 participants (9%) who access the platform "about once a day," and 6 (7%) who visit it "several times a week." Very few respondents reported lower usage frequencies such as "about once a week" or "about once a month," and none indicated unfamiliarity with the platform.

This pattern confirms that the sample is predominantly composed of frequent users, which may influence their perceptions of usability through routine and familiarity. As noted in previous usability research, frequent exposure to a digital interface often leads to higher perceived ease of use, even in the presence of suboptimal design features (Nielsen, 1993).

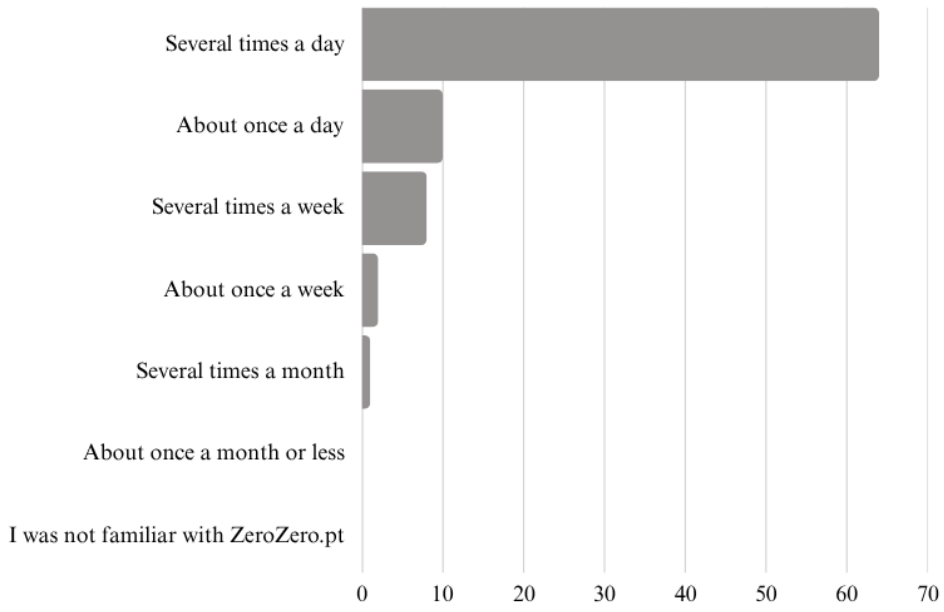


Figure 6.2: Frequency of visits to ZeroZero.pt (Questionnaire 1)

Participants also demonstrated a strong familiarity with ZeroZero.pt. As shown in Figure 6.3, 43 respondents rated their knowledge of the platform at level 5 (maximum), and 37 selected level 4, on a 5-point Likert scale. Only a small minority chose level 3 (n = 5), and no responses were recorded for levels 1 or 2.

These results indicate that the sample is composed predominantly of experienced and regular users, which may influence their usability perceptions positively due to familiarity effects and the development of habitual usage patterns.

[1] https://www.similarweb.com/website/zerozero.pt

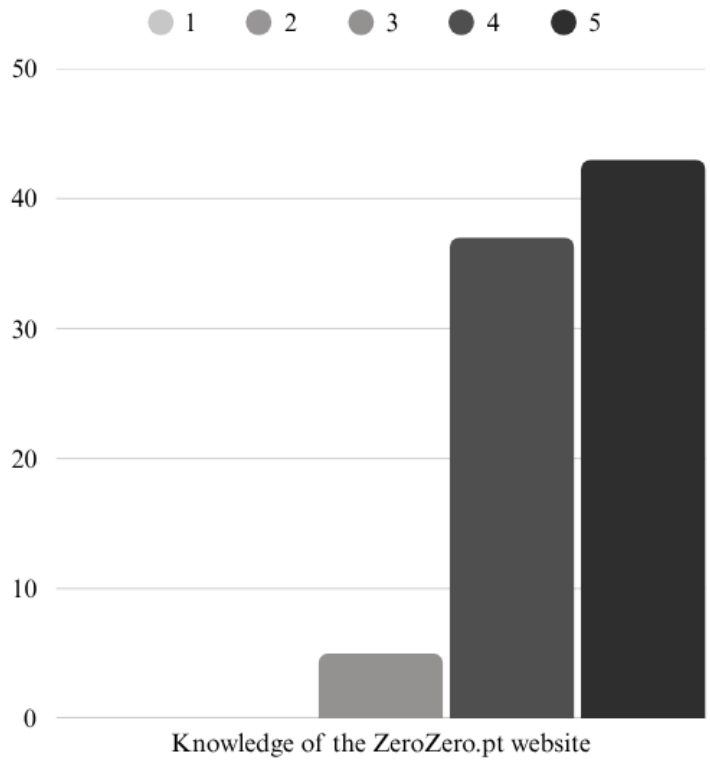


Figure 6.3: Self-reported knowledge of the ZeroZero.pt website (Questionnaire 1)

6.2.3 System Usability Scale (SUS)

The SUS was used to assess participants' overall perception of usability regarding the mobile version of ZeroZero. The analysis yielded a mean SUS score of 81.35 (SD = 17.36), with scores ranging from 37.5 to 100. The median score was 87.5, indicating a tendency towards high perceived usability among users. Based on established benchmarks Brooke, 1996 and Sauro et al., 2012, this average falls into the "Good to Excellent" usability category. The presence of both high and low scores highlights some variability in user experience, which will be further examined in relation to other variables in the following sections.

To visually assess the distribution of SUS scores, a histogram with an overlaid normal curve was generated (Figure 6.4). Although the data distribution is positively skewed, most scores concentrate in the higher range, particularly between 75 and 100. This visual pattern complements the high median and mean values observed and reinforces the overall positive perception of usability.

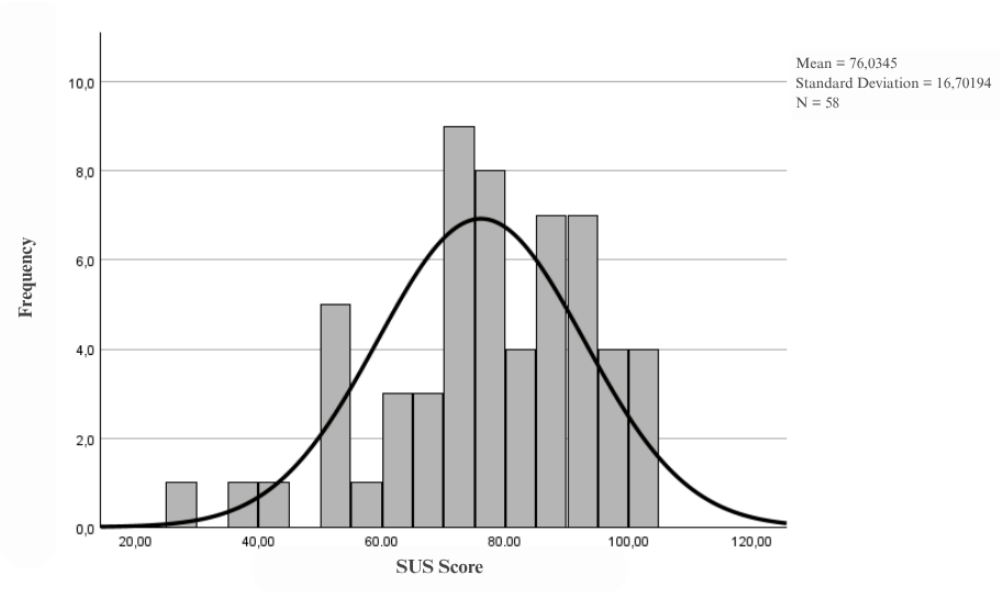


Figure 6.4: Distribution of SUS Scores (Questionnaire 1)

To provide a more interpretable view of the SUS scores, participants' results were categorized using the curved grading scale proposed by Sauro and Lewis Sauro et al., 2012, which converts raw scores into letter grades ranging from A+ to F. This scale is based on percentile ranks derived from a large benchmark dataset of SUS scores, allowing a nuanced understanding of usability perceptions. As shown in Figure 6.5, the distribution was skewed towards the upper end of the scale. A considerable proportion of users received grades in the A range (A+, A, A-), suggesting a strong positive perception of usability. Conversely, only a small number of participants fell into the lower grade categories, such as D or F. This reinforces the interpretation that the mobile version of ZeroZero is generally perceived as highly usable, while still acknowledging that a minority of users reported less favourable experiences.

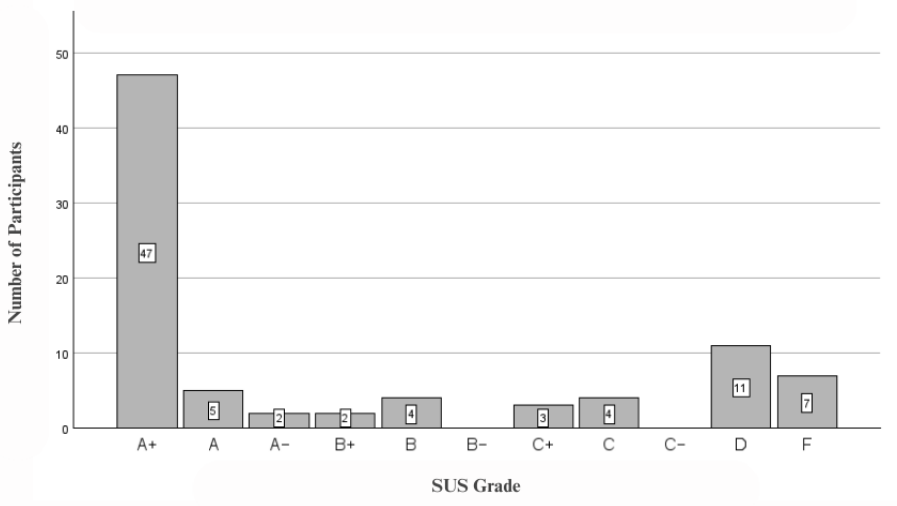


Figure 6.5: Distribution of SUS Grades (Questionnaire 1)

6.2.4 Normality Tests

To determine the appropriate statistical tests, normality was assessed using both the Shapiro-Wilk and Kolmogorov-Smirnov tests. As shown in Table 6.1, all three continuous variables—SUS Score, Task Success Percentage, and Visit Frequency—showed significant deviations from normality ($p < .001$ for all). These results indicate that the assumption of normality required for parametric tests was violated, and therefore non-parametric alternatives were adopted in line with Field, 2018 [2] recommendations for non-normally distributed data.

Table 6.1: Normality Tests for Key Variables (Questionnaire 1)

Variable	Test	Statistic	df	Sig. (p)
SUS Score	Kolmogorov-Smirnov	.148	85	< .001
	Shapiro-Wilk	.886	85	< .001
Task Success %	Kolmogorov-Smirnov	.500	85	< .001
	Shapiro-Wilk	.340	85	< .001
Visit Frequency	Kolmogorov-Smirnov	.444	85	< .001
	Shapiro-Wilk	.567	85	< .001

This was further confirmed by visual inspection. As illustrated in Figure 6.6, the Task Success Percentage distribution is strongly right-skewed, with the majority of participants achieving a 100% success rate. This ceiling effect further supports the rejection of the normality assumption.

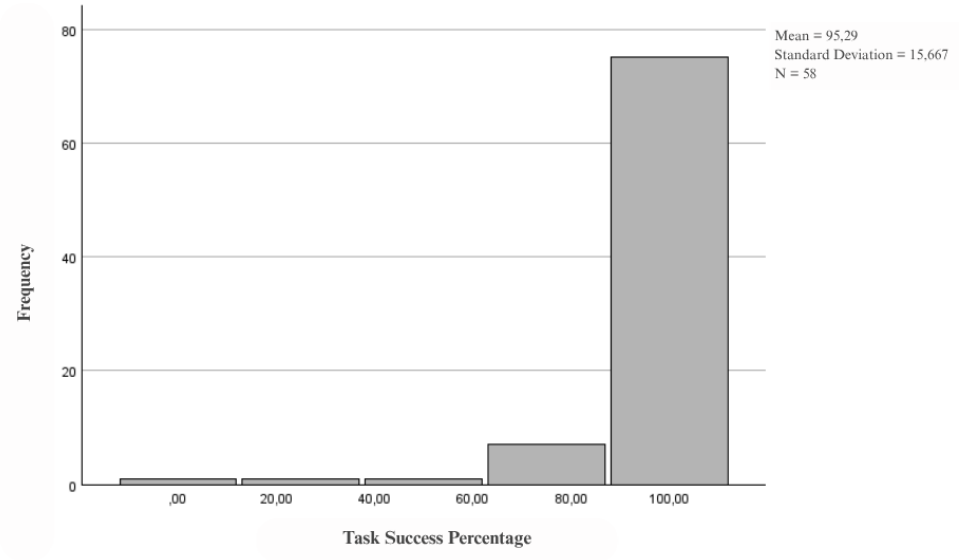


Figure 6.6: Histogram of Task Success Percentage (Questionnaire 1)

[2] See Field, 2018, Chapter 7.2, for a full discussion on the use of non-parametric tests when normality assumptions are not met.

Given these results, parametric tests such as the independent samples t-test, one-way ANOVA, Pearson correlation, and linear regression (which assume normally distributed variables or residuals) were deemed inappropriate. Therefore, non-parametric alternatives were adopted, including the Mann-Whitney U test instead of the t-test, Kruskal-Wallis instead of ANOVA, and Spearman's rank correlation instead of Pearson's correlation.

6.2.5 Inferential Analysis

To further explore the relationships between users' perceptions and their behaviours, a series of inferential statistical tests were conducted using IBM SPSS Statistics (version 29) [3].

In this section, inferential statistical methods are employed to test hypotheses and explore relationships between key variables related to user experience, task performance, and usage patterns. Inferential analysis enables the generalization of sample results to a broader user population.

Prior to conducting inferential tests, the assumption of normality was evaluated using the Shapiro-Wilk test, which revealed significant deviations from normality across all main continuous variables (SUS score, task success percentage, and visit frequency). As a result, non-parametric tests were adopted, given their robustness and suitability for non-normally distributed data.

The selected inferential procedures include Spearman's rank correlation to assess associations between variables, Mann-Whitney U tests to compare differences between two independent groups, and Kruskal-Wallis tests for comparisons involving more than two groups. These tests were applied to examine how usability perceptions, task outcomes, and user characteristics interact in the context of the ZeroZero mobile platform.

6.2.5.1 Comparison of Perceived Usability Based on Task Success (Mann–Whitney U Test)

To examine whether the level of task success influenced users' perceived usability, a binary variable was created from the original percentage of task success. Participants who achieved a full 100% success rate were coded as group 1 ("Complete Success"), while those who completed less than 100% of the tasks were coded as group 0 ("Incomplete Success") [4]. This transformation allowed the use of the Mann–Whitney U test to compare the perceived usability, as measured by the SUS score, between two independent groups, while accounting for the non-normal distribution of the data.

A Mann–Whitney U test was conducted to compare perceived usability (SUS score) between users who completed all tasks successfully and those who did not. Results indicated that the group with full task success (n = 75) had a higher mean rank in SUS scores (44.68) compared to the group without full success (n = 10; mean rank = 30.40). However, this difference was not statistically significant at the conventional threshold of $p < .05$, $U = 249.000$, $Z = -1.728$, $p = .084$ (two-tailed).

[3] https://www.ibm.com/products/spss-statistics

[4] It is important to note that the binary classification of task success was based on a strict threshold: only participants who achieved exactly 100% success were coded as "Complete Success". Any value below this threshold, including 99.99%, was assigned to the "Incomplete Success" group. This decision reflects the operational definition of full task success used in this study and aligns with the binary recoding rules implemented in SPSS.

Although the result does not reach statistical significance, the observed trend suggests that users who completed all tasks may have perceived the interface as more usable. This aligns with expectations in usability research, where successful task completion often correlates with higher subjective evaluations of system quality.

Table 6.2: Results of Mann–Whitney U Test for SUS Score by Task Success (Questionnaire 1)

Group	**N**	**Mean Rank**	**Sum of Ranks**
Incomplete Success (0)	10	30.40	304.00
Complete Success (1)	75	44.68	3351.00
Test Statistic	Mann–Whitney U = 249.000, Z = -1.728, p = .084 (2-tailed)		

Figure 6.7 shows the distribution of SUS scores between users with and without complete task success. Although the difference was not statistically significant, users who completed all tasks (Complete Sucess) tended to report higher perceived usability.

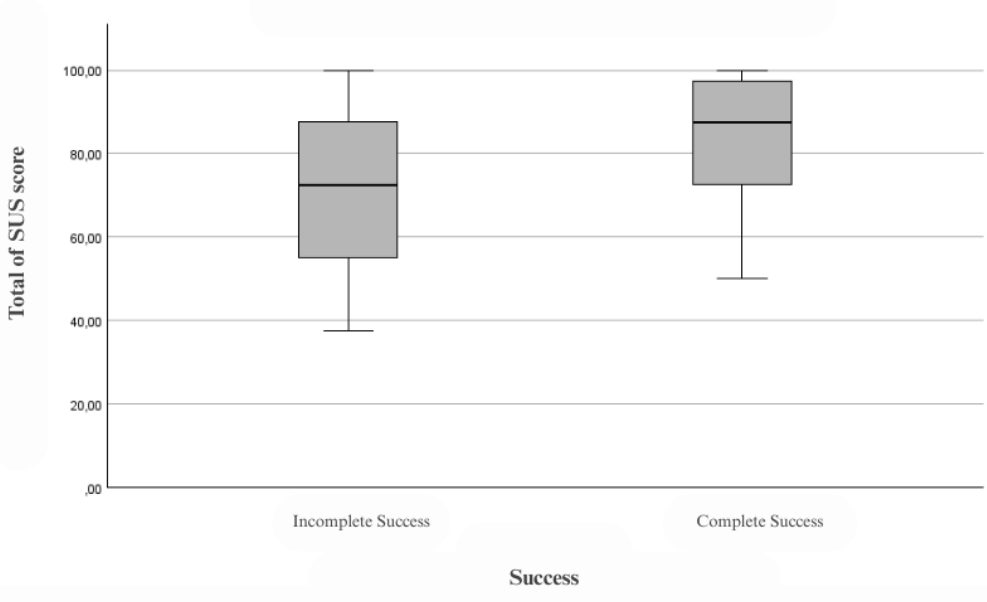


Figure 6.7: Boxplot of SUS scores by task success group (Questionnaire 1)

6.2.5.2 Comparing SUS Scores Based on Platform Familiarity (Kruskal–Wallis Test)

A Kruskal–Wallis test was conducted to examine whether users' perceived usability (SUS score) differed according to their level of knowledge about the ZeroZero platform. The null hypothesis (H_0) stated that the distributions of SUS scores would be equal across the three knowledge levels, while the alternative hypothesis (H_1) proposed that at least one group would differ significantly.

Three groups were compared: users with knowledge levels 3 (n = 5), 4 (n = 37), and 5 (n = 43). The test revealed a difference in mean SUS ranks across groups, with higher knowledge

levels associated with higher perceived usability: level 3 (mean rank = 21.20), level 4 (mean rank = 41.85), and level 5 (mean rank = 46.52). These results are summarised in Table 6.3. However, the result was not statistically significant, $H(2) = 4.908$, $p = .086$. As the p-value did not fall below the conventional alpha level of .05, the null hypothesis could not be rejected. Therefore, no statistically significant difference in perceived usability was found across the three knowledge levels.

Table 6.3: Kruskal–Wallis Test for SUS Score by Knowledge Level (Questionnaire 1)

Knowledge Level	**N**	**Mean Rank**
Level 3	5	21.20
Level 4	37	41.85
Level 5	43	46.52
Test Statistic	$H(2) = 4.908$, $p = .086$	

Although not conclusive, the data suggest a potential positive association between platform familiarity and perceived usability. This trend is also visually supported by the boxplot in Figure 6.8, which shows a consistent increase in the median SUS scores as the level of user knowledge rises.

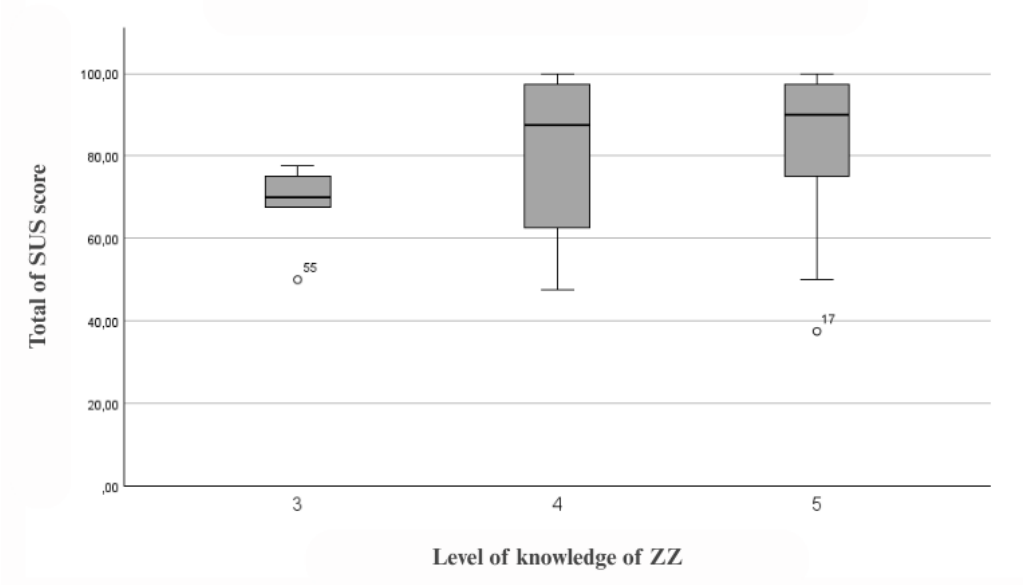


Figure 6.8: Boxplot of SUS scores by knowledge level about the ZeroZero platform

6.2.5.3 Spearman's Correlation Between Frequency of Visit and Perceived Usability

To explore the potential relationship between users' frequency of visiting ZeroZero.pt and their perceived usability of the mobile platform, a non-parametric Spearman correlation analysis was conducted.

The results indicated a statistically significant positive correlation between frequency of visit (coded from 0 to 6, where higher values indicate more frequent usage) and SUS score (r_s = .253, p = .020, N = 85), the results are summarized in table 6.4. This suggests that users who visit ZeroZero more frequently tend to report slightly higher usability perceptions.

While the correlation is relatively weak, the result points to a possible link between familiarity with the platform and more favourable usability perceptions. Given the significance level ($p <$ .05), this association is unlikely to be due to chance.

These findings contribute to understanding the influence of usage patterns on usability evaluation and highlight the potential role of habitual interaction in shaping users' overall experience.

Table 6.4: Spearman's Rank-Order Correlation Between SUS Score and Frequency of Visit to ZeroZero (Questionnaire 1)

	SUS Score	**Frequency of Visit**
SUS Score	1.000	.253*
Frequency of Visit	.253*	1.000
p-value (2-tailed)	.020	.020
N	85	85

*Correlation is significant at the 0.05 level (2-tailed).

6.2.5.4 Association Between Type of Internet Access and Perceived Loading Time: Chi-Square Test of Independence

To explore whether the type of internet access used by participants was associated with their perceived loading times while using ZeroZero's mobile version, a Chi-square test of independence was conducted. Two responses were excluded because participants provided comments that did not correspond to a time-based evaluation (e.g., remarks unrelated to page loading duration [5]), making it impossible to categorize their responses meaningfully. Resulting in a valid sample of $N = 83$.

The variables analyzed included the type of Internet access (Mobile Data vs. Wi-Fi) and the average loading time, recoded into three categories[6]: *< 1 second*, *1–5 seconds*, and *5–10 seconds*.

The Chi-square test revealed a statistically significant association between the type of Internet access and the perceived loading time, $\chi^2(2, N = 83) = 7.373$, p = .025 (see Table 6.5). This suggests that loading time perceptions varied depending on the user's connection type. However, one cell (16.7%) had an expected count below 5, which slightly limits the robustness of the result and should be taken into account when interpreting the outcome.

[5] One participant wrote "I couldn't find the Livescore section" and another noted "There is no Livescore section from the start menu", which did not correspond to a time-based evaluation.

[6] Time categories were defined based on established usability thresholds for system response times, as recommended by Nielsen Norman Group: https://www.nngroup.com/articles/website-response-times/

Table 6.5: Chi-square Test of Independence Between Internet Access Type and Loading Time (Q1)

Test	**Value**	**df**	***p*-value**
Pearson Chi-square	7.373	2	**.025**
Likelihood Ratio	7.070	2	**.029**
Valid cases	83		

As shown in Figure 6.9, Wi-Fi users were more likely to report longer loading times. In the "< 1 second" category, only 3 Wi-Fi users selected this option, compared to 7 mobile data users. Conversely, 37 Wi-Fi users reported loading times in the "1–5 seconds" category (vs. 17 mobile data users), and 15 Wi-Fi users indicated "5–10 seconds" (vs. 4 mobile data users).

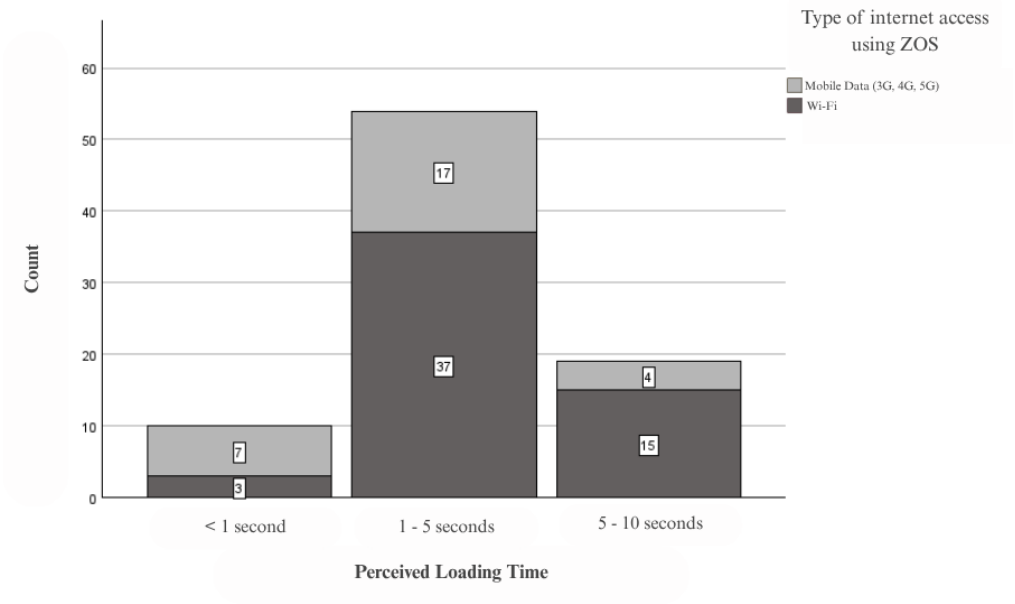


Figure 6.9: Stacked bar chart of perceived loading time by type of internet access (Q1)

Although Wi-Fi was the predominant connection type (66%), as seen in Table 6.6, descriptive results showed that mobile data users more frequently reported faster loading experiences. Specifically, 24.1% of mobile users (n = 7 out of 29) indicated loading times under 1 second, compared to only 5.4% of Wi-Fi users (n = 3 out of 56). This difference in perception was statistically significant, as confirmed by a chi-square test ($\chi^2(2, N = 83) = 7.373$, $p = .025$).

Table 6.6: Distribution of Internet Access Type Among Participants

Type of Internet Access	**Count**	**Percentage**
Mobile Data (3G, 4G, 5G)	29	34%
Wi-Fi	56	66%

6.2.5.5 Correlation Between Perceived System Performance and Usability (Spearman)

Spearman's rank-order correlation was used to examine the relationship between perceived system performance and perceived usability (SUS score). The analysis focused on two indicators of performance perception: experience of content loading speed and perception of unexpected page delays.

The null hypothesis (H_0) stated that there would be no monotonic association between system performance perception and usability. The alternative hypothesis (H_1) proposed the existence of a statistically significant correlation.

A statistically significant, weak positive correlation was found between content loading experience and SUS score, $\rho(85) = .216$, $p = .048$. This indicates that more positive evaluations of loading speed were associated with slightly higher usability scores. No statistically significant correlation was observed between the perception of unexpected delays and the SUS score, $\rho(85) = .196$, $p = .072$. The results are summarised in Table 6.7.

Table 6.7: Spearman's Correlation Between System Performance Perception and SUS Score (Questionnaire 1)

Variable	**ρ (rho)**	**p-value**	**N**
Content loading experience	.216*	.048	85
Perceived unexpected page delay	.196	.072	85

Figure 6.10 displays the distribution of SUS scores across different levels of perceived content loading experience. Each category on the X-axis represents a progressively more negative evaluation, with higher values indicating slower or less satisfactory loading. A visible upward trend can be observed in the median SUS scores as perceived loading experience improves. This visual pattern supports the result obtained from the Spearman's correlation, which indicated a weak but statistically significant positive association between these two variables, $\rho(85) = .216$, $p = .048$.

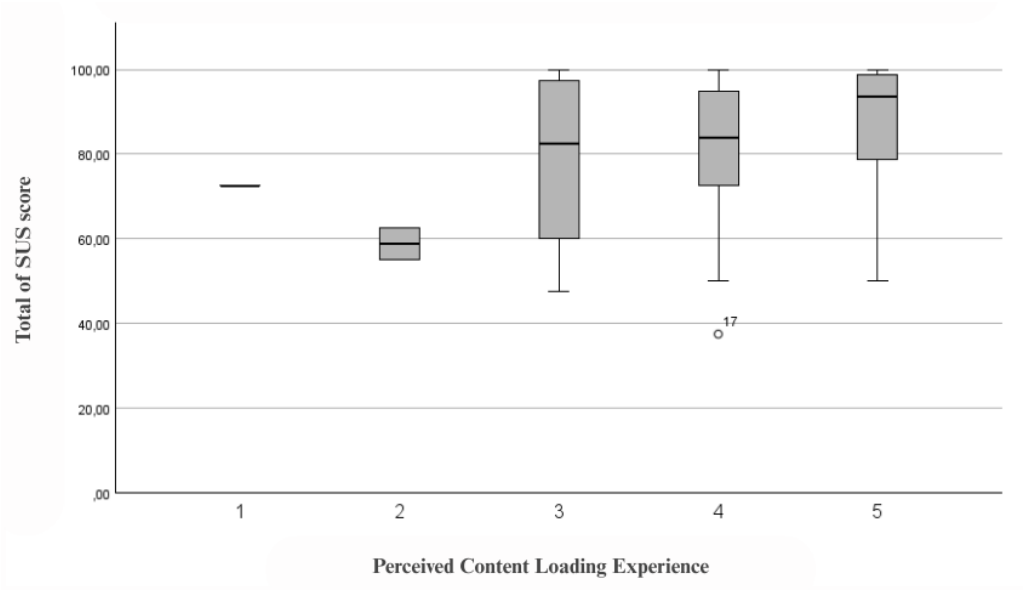


Figure 6.10: Boxplot of SUS scores by perceived content loading experience (Questionnaire 1)

Figure 6.11 shows the distribution of SUS scores according to users' perception of whether any pages took longer than expected to load. The scores are grouped by levels of perceived delay, with higher values corresponding to more negative evaluations. While a slight upward trend in SUS scores appears as delay perception decreases, this pattern was not statistically significant in the correlation analysis, $\rho(85) = .196$, $p = .072$. The boxplot is provided for descriptive purposes only and should not be interpreted as evidence of association.

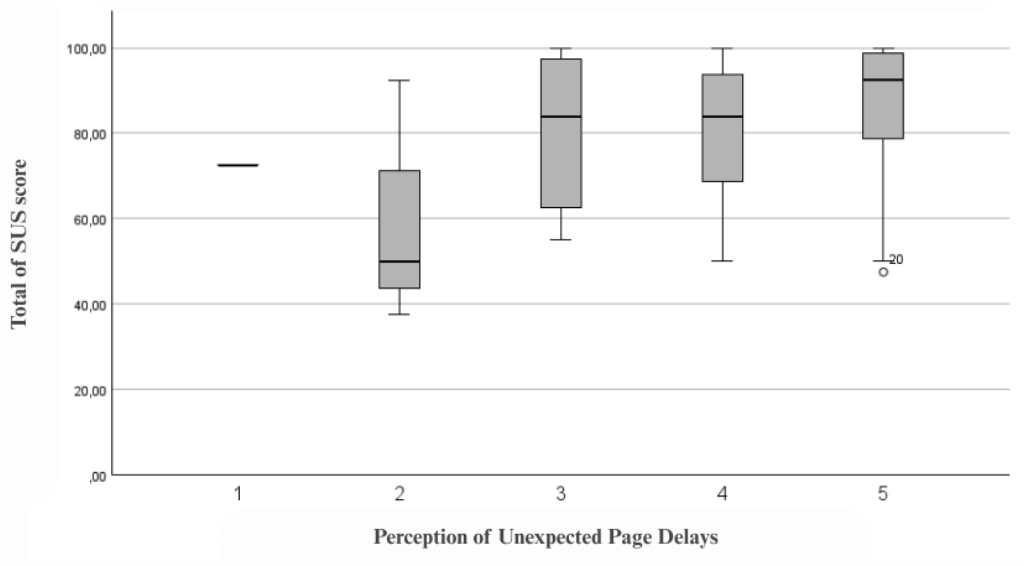


Figure 6.11: Boxplot of SUS scores by perception of unexpected page delays (Questionnaire 1)

6.3 Qualitative Data Analysis of Registered Users (Questionnaire 1)

This section presents the qualitative analysis of data collected from registered users. It includes open-ended questionnaire responses, focus group discussions, and Hotjar session recordings. The analysis aims to uncover deeper insights into user frustrations, preferences, and suggestions for improvement based on their interactions with the mobile version of the platform.

6.3.1 Open-ended Questionnaire Responses

The open-ended responses from were categorized thematically, according to the type of user input:

1. What users would change in the mobile version,
2. Their biggest frustration,
3. Suggested new feature, and
4. Additional comments.

Each question was analyzed independently using a thematic coding framework. The categories were derived inductively based on recurring topics in the responses, allowing a grounded insight into user perspectives. The categories include: Interface/Layout, Ads/Pop-ups, Search, Mobile App, Personalization, Live Updates, Navigation, Advanced Stats, Performance, Content Quality, Suggested Features, User Support & Moderation, No Change and Nothing to add.

1. Question: What would you change in the mobile version?

The majority of responses (21) indicated No Change, suggesting a generally positive perception of the mobile version among registered users. Nonetheless, the remaining responses brought attention to specific aspects of the platform that could benefit from improvement, with several recurring themes emerging.

The most frequently mentioned theme was Interface and Layout Issues (14). Participants described problems with visual coherence, such as poorly adapted screen formats when rotating the phone, a layout that felt outdated or cluttered, and inconsistency between pages. Many users emphasized the need for a cleaner, more modern homepage and a layout that mimics the structure of the desktop version. Others mentioned the difficulty in reading match details or accessing specific data due to how information is currently arranged.

Ads and Pop-ups were also a significant concern, with 10 users reporting that advertising content interfered with the user experience. These users described being frequently interrupted by intrusive pop-ups or banner ads, which not only slowed down the loading of pages but also led to unintended clicks or misnavigation.

Navigation was another prominent issue, referenced in 8 responses. Users highlighted difficulty locating certain competitions, such as youth teams or women's leagues, as well as inefficiencies in the main menu structure. Some also noted that elements like the hamburger menu [7] or the lack of a search history made navigating the site slower and more cumbersome than expected.

The theme of Personalization appeared in 6 responses, where users expressed the desire for a more tailored experience. Suggestions included the ability to pin favorite teams, display frequently accessed content more prominently, or adapt the interface dynamically based on previous user behavior.

Five users pointed to shortcomings related to Live Updates, suggesting the addition of real-time classification tables and live player statistics that reflect in-game activity. This was complemented by 5 users referring to Content Quality issues, ranging from missing information in match pages (e.g., absence of year, incomplete squad data) to requests for broader coverage of local or amateur competitions.

A smaller number of responses suggested the need for a dedicated Mobile App (4), citing limitations of browser-based access. Meanwhile, 3 users mentioned Advanced Stats, asking for richer player-level performance metrics. Another 3 referred to Search limitations, particularly in locating players or clubs with common names.

Only 2 users proposed Suggested Features such as the ability to edit or create team entries without needing administrative approval. Lastly, a single user mentioned Performance problems, specifically citing slow loading speeds during peak usage on Sundays.

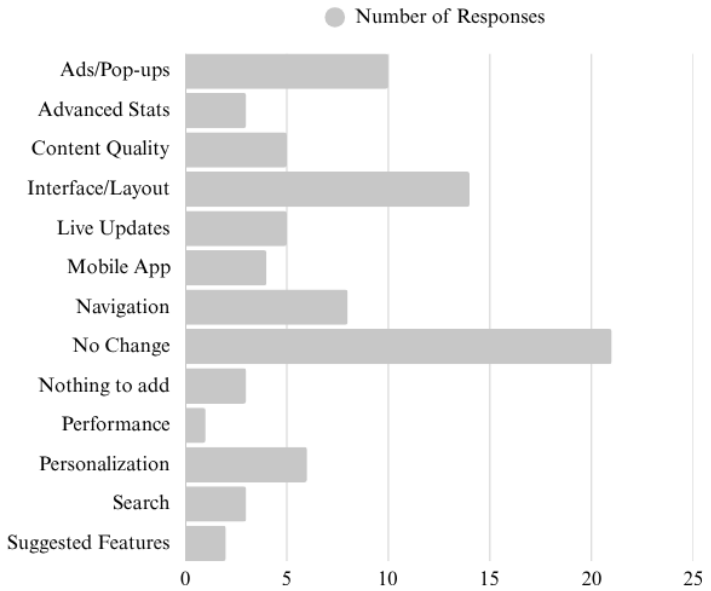


Figure 6.12: Suggested changes to the mobile website (Questionnaire 1)

[7] For an accessible explanation and comparison between the hamburger menu and alternative navigation patterns in mobile UX, see: https://www.weareconflux.com/en/blog/tab-bar-vs-hamburger-menu/

Although a considerable portion of respondents did not suggest any changes, as seen in Figure 6.12, the diversity and specificity of the remaining suggestions highlight meaningful opportunities for improving the mobile user experience.

2. Question: What was your biggest frustration when using the website on mobile?

Although many users (26) responded with "Nothing to add", indicating the absence of a particularly significant frustration, the remaining answers reveal key usability concerns that hinder the overall mobile experience.

A prominent theme was Performance and Speed Issues, with 11 users describing delays in page loading—particularly during high-traffic periods like Sunday afternoons. Some reported the website taking up to 30 seconds to load, while others noted responsiveness issues on specific sections, such as the encyclopedia or matchday content. These concerns extended to live content, where users mentioned delays in the processing of real-time match data.

Another frequent frustration was related to Content Quality (12). Participants highlighted issues such as outdated or missing information, poor journalistic quality of articles, and difficulty in finding older players or competitions. Some users expressed frustration with duplicated entities or content not being accessible without authentication. This category reflects concerns not only with accuracy, but also with the reliability and completeness of the platform's data.

Interface and Layout problems were also raised by 8 users, who pointed out visual inconsistencies, unresponsive menus, and confusion caused by redundant or overly complex structures. Examples included unclear buttons, layout breaking on smaller screens, and difficulties interacting with essential features like the main menu.

Closely related to interface issues were frustrations regarding Search Functionality (9). Users reported that the internal search was unreliable, especially for players or teams with common names. Many claimed they often had to resort to external search engines (e.g., "ZeroZero + player name") to locate specific content.

Issues with Ads and Pop-ups were mentioned by 11 users. These were described as intrusive and sometimes repositioning after the page had loaded, interfering with navigation and content visibility. Others noted that excessive advertising slowed down the site significantly or made it visually overwhelming.

Smaller yet relevant concerns included Navigation challenges (3 users), such as difficulties in finding certain competitions (especially youth or women's) and confusion when clicking links before the page had fully loaded. Additionally, Advanced Stats, User Support & Moderation, and the absence of a dedicated mobile app were each cited by a minority of users, as seen in Figure 6.13.

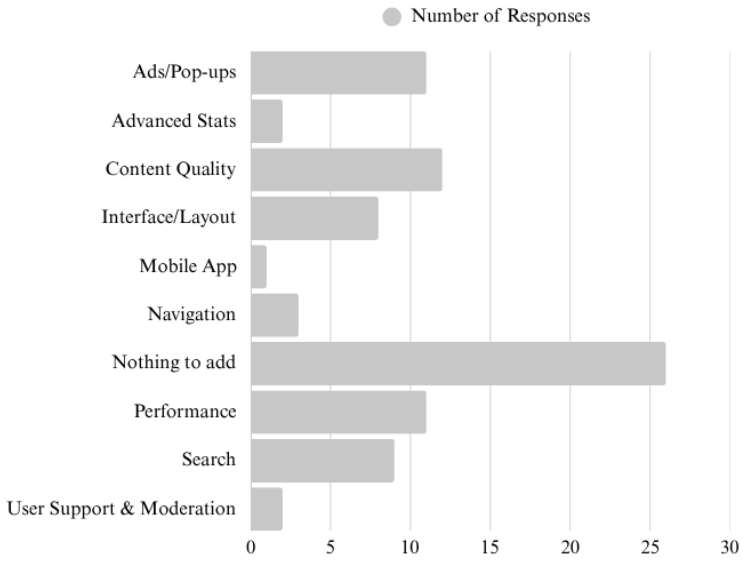


Figure 6.13: Biggest user frustrations with the mobile website (Questionnaire 1)

Taken together, these responses underscore the importance of streamlining navigation, improving performance during peak hours, and refining core functionalities like search and layout—while also ensuring that the content remains accurate, accessible, and free from obtrusive advertising.

3. Question: Do you have any suggestion for a new feature?

Most participants (47) indicated having no suggestions for new features, which may reflect a level of contentment with the current scope of functionalities. However, the responses that did include suggestions pointed to specific areas where users felt additional improvements could enhance their experience.

Among the most mentioned themes was Personalization (5), with users suggesting the ability to customize their homepage, save preferred clubs for faster access, or tailor their experience based on past interactions. This reflects a desire for a more adaptive and user-centric design that responds to individual preferences.

The category of Content Quality (6) also featured prominently. Suggestions included improved editorial standards, better coverage of niche sports such as rugby and amateur leagues, and clearer structures for news articles or match reports. Users also called for expanded databases, for example with historical statistics or additional metadata about players.

Another relevant area was Mobile App (4), where participants suggested creating a dedicated ZeroZero application or improving app-like behaviors for a smoother and more cohesive mobile experience. Related to this, Interface/Layout (4) responses emphasized the importance of clarity and organization, particularly in how information is displayed or sorted on the platform.

Some users advocated for Advanced Stats (4), proposing the inclusion of richer data related to coaching staff, historical performance, and player-specific metrics. Similarly, Live Updates (2)

suggestions pointed to enhancements to the LiveScore feature, such as real-time positioning of the ball or a more responsive update system during matches.

There were also comments under User Support & Moderation (4), where users requested better responsiveness to support tickets, the ability to view comment history, or tools to moderate/block other users. While fewer in number, topics like Search (3), Navigation (1), and Suggested Features (3) also appeared, suggesting refinements in how users access and interact with the site's content.

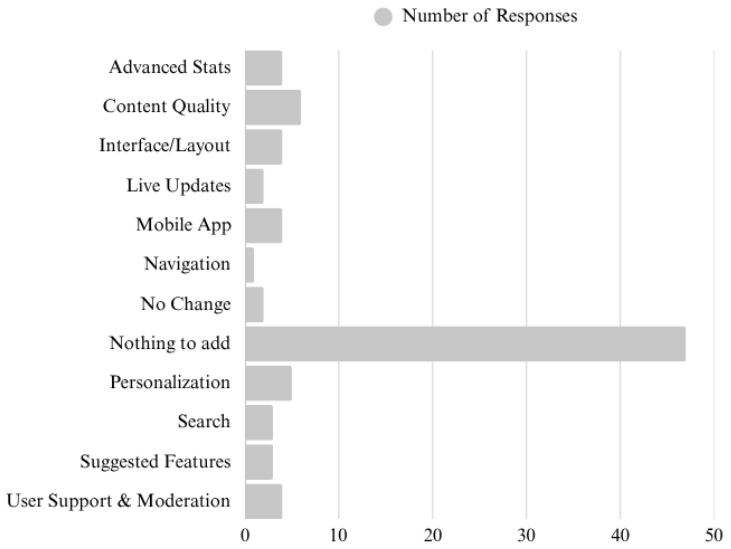


Figure 6.14: Suggested new feature (Questionnaire 1)

Although the majority chose not to suggest any features, as seen in Figure 6.14, the responses collected present several actionable and specific recommendations that can inform future development, particularly in areas tied to customization, content structure, mobile optimization, and statistical depth.

4. Question: Additional Comments?

The vast majority of participants (77 out of 85) reported having no additional comments. This is particularly plausible considering the richness and breadth of the feedback already provided in earlier answers.

Among the few who did share comments, some suggestions highlighted Content Quality (2). Participants noted the presence of grammatical and semantic inaccuracies in published content, such as improper usage of certain expressions, which could affect the perceived professionalism of the platform. One user emphasized that textual refinement would elevate the platform from being seen as merely functional to a more premium and reputable source.

Comments also touched on Interface/Layout (2), with mentions of visual elements such as national flags missing in international club competitions, or usability difficulties when completing

tasks—such as the questionnaire itself—via mobile. These underline opportunities to enhance accessibility and visual clarity across interfaces.

Regarding the Mobile App, a few users (2) expressed a desire for a standalone application, referencing it again in this open comment section. Additionally, a suggestion under Live Updates advocated for the inclusion of real-time player and match data during games, rather than post-match summaries.

A single comment under User Support & Moderation raised concerns about the moderation criteria used to restrict user comments, specifically questioning the lack of transparency or justification behind certain actions.

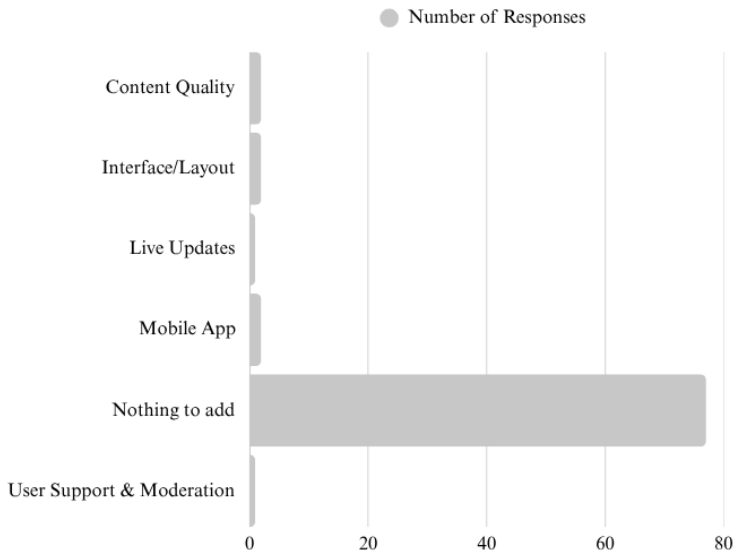


Figure 6.15: Additional Comments (Questionnaire 1)

Despite the limited number of responses, likely due to the optional nature of this question, the few comments offered valuable insights. They reinforce the importance of editorial precision, clear interface design, and transparent moderation practices in fostering user trust and satisfaction.

6.3.2 User Behavior Observations (Hotjar)

To complement the questionnaire data with behavioral insights, user interactions on the ZeroZero mobile platform were tracked using Hotjar. Importantly, session recordings were only triggered for users who accessed the site through the dedicated questionnaire link. This ensured that behavioral data was directly tied to participants in the questionnaire, though it also limited the scope of observation to this specific subset of users.

6.3.2.1 Data Collection Context

User interaction data was collected using Hotjar during the period from **April 29 to May 12th, 2025**, aligned with the first questionnaire phase. Hotjar's *subscription trial version* was used, enabling full access to:

- Heatmaps by device and URL
- User funnels and event-based segmentation
- Detection of rage clicks, U-turns, text input, and page refresh events
- Session-level behavioral filtering (duration, OS, interaction types)

In total, Hotjar recorded **50 user sessions** across multiple task flows linked to the questionnaire. The recording was triggered only when users accessed the platform via the questionnaire-specific link.

6.3.2.2 Heatmap Observations

Interaction patterns were evaluated through click maps, scroll maps, and engagement zones for multiple key pages:

- Homepage (`https://www.zerozero.pt/`)
- Portuguese League page
- Player and team pages (e.g., FC Porto, Viktor Gyökeres)
- News article (latest content about the national team)

The following types of heatmaps were generated and analyzed:

- **Click heatmaps** – to detect zones with high interaction density.
- **Scroll depth heatmaps** – to evaluate how far users navigated through the page content.
- **Engagement heatmaps** – to capture areas of prolonged user attention.
- **Rage click heatmaps** – to identify possible frustration signals.

The homepage heatmaps showed a concentration of interactions in the upper segments, particularly around the navigation bar, featured articles, and usability banner. Scroll depth analysis revealed that while all users viewed the top section, visibility dropped sharply beyond the mid-page, with fewer than 25% reaching the footer. Engagement hotspots aligned with high-click zones, reinforcing the importance of the top-level content.

Similar interaction patterns were observed on the Portuguese League and FC Porto pages, where clicks clustered around tabs, team information sections, and match listings. Scroll data

revealed partial exploration of the content, with significant drop-off after mid-page. Notably, engagement zones were limited mostly to the first screen and navigation tools.

For the Viktor Gyökeres and news article pages, interactions focused on the initial player summary and article header, with diminishing activity further down. Engagement and scroll data both suggest attention was brief and top-heavy.

Rage click heatmaps confirmed the absence of user frustration behaviors across all heatmap analyses.

6.3.2.3 Global Metrics Summary

Table 6.8: Summary of Global Metrics from Hotjar Sessions

Metric	**Value**
Total sessions recorded	50
Average time on page	33 seconds
Total rage clicks	0
U-turns	9
Text input occurrences	3
Page refresh events	8
Drop-off rates (avg.)	3.8%
Devices used (Android/iOS)	40 / 14 sessions

6.3.2.4 Event Funnels Overview

Hotjar funnels were created to trace user paths during task execution. The following events and funnels were monitored:

- Sessions lasting under 50 seconds and under 10 minutes
- Funnel from Portuguese League page to a player (e.g., Viktor Gyökeres)
- Funnels segmented by operating system (Android vs. iOS)
- Events such as rage clicks, U-turns, text entry, and refresh behavior

Each funnel is described individually in Subsubsection 6.3.2.6.

6.3.2.5 Limitations of the Data

Although the trial version allowed access to full features, some limitations must be acknowledged:

- The analysis was limited to sessions where users accessed via the questionnaire link.
- Sample size was conditioned by the number of responses (85) and session triggers (50 recorded sessions).
- External contexts (e.g., network delays, device limitations) were not captured.

6.3.2.6 Analysis of Funnel Recordings

To complement the insights obtained through heatmaps, this subsection presents a series of funnel-based analyses using Hotjar, focused on conversion behavior, interaction patterns, and system-specific usage trends. These funnels were extracted between April 29 and May 12, 2025, and allow a quantitative interpretation of user flow along selected tasks.

The first relevant funnel tracked user engagement on the Liga Portuguesa page, specifically focusing on interactions with the player Viktor Gyökeres. From a total of 25 users who viewed the page for more than 50 seconds, only two proceeded to click on the related player element, corresponding to a conversion rate of 8%.

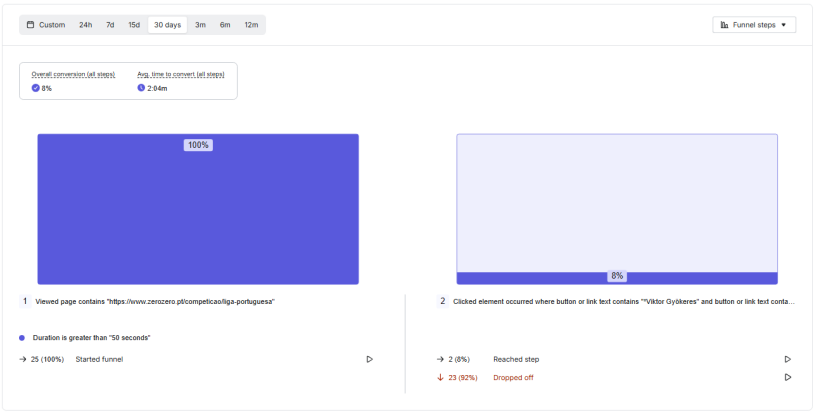


Figure 6.16: Conversion funnel on the Portuguese League page for Viktor Gyökeres

To better understand differences across user devices, the same funnel was replicated while segmenting by operating system. Among 40 users on Android, 10% clicked on the element (4 users), while on iOS, only one out of 14 (7.1%) did so. This reveals a slightly more active behavior from Android users in this conversion path.

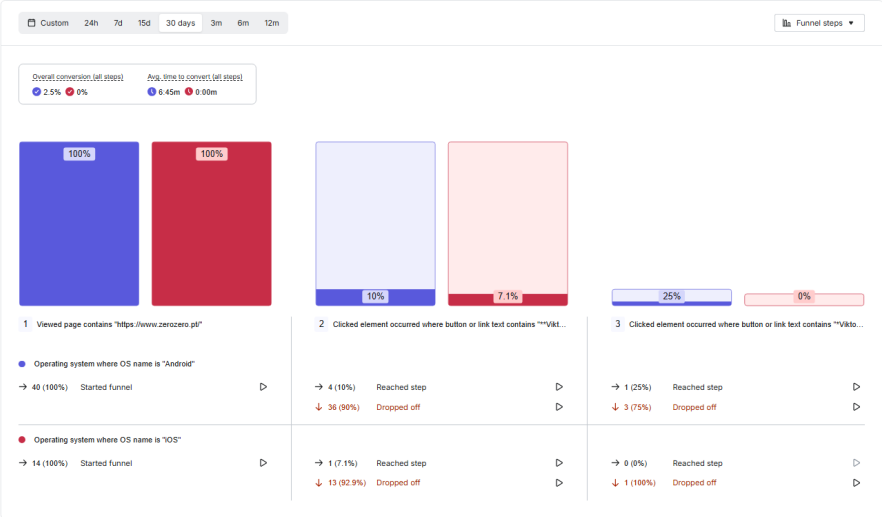


Figure 6.17: Comparative conversion funnel between Android and iOS on the Portuguese League page

Another funnel assessed the historical matchup page between FC Porto and Sporting. Despite the small sample (3 users), one proceeded to the target click, achieving a 33.3% conversion rate.

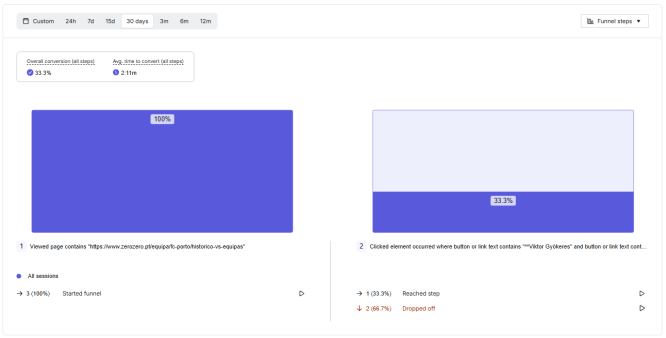


Figure 6.18: Conversion funnel on the FC Porto vs Sporting page (History)

In addition to conversions, user frustration or confusion was explored through behavioral metrics such as U-turns and rage clicks. No rage clicks were recorded across all sessions.

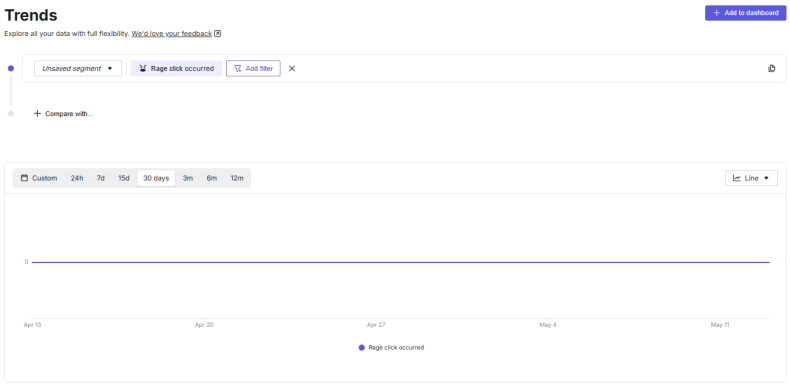


Figure 6.19: No rage clicks during the period analyzed

U-turn events were detected occasionally, indicating users returning rapidly to a previous page. This behavior was plotted globally:

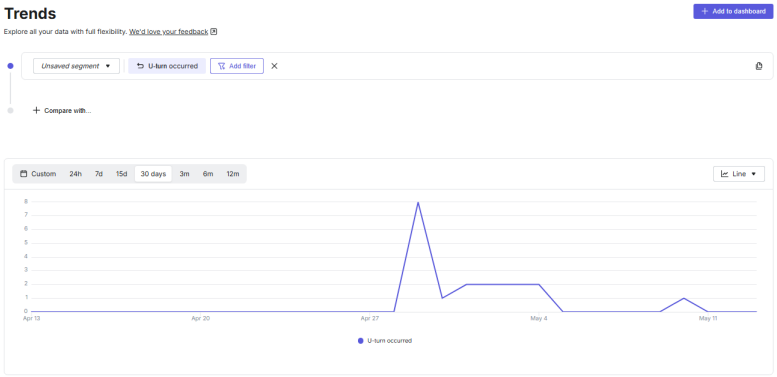


Figure 6.20: General trend of occurrences of U-turns over the analysis period

A device usage breakdown shows that the majority of sessions originated from Android systems (40 vs 14 on iOS). This distribution is illustrated below:

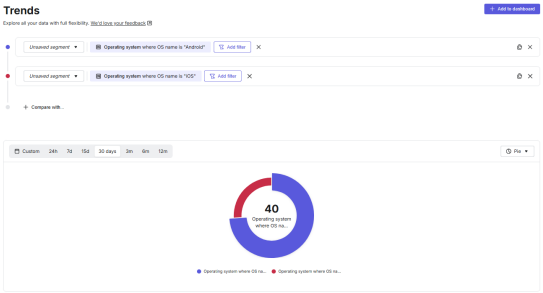


Figure 6.21: Distribution of sessions by operating system

The temporal evolution of device usage is presented in the following figure:

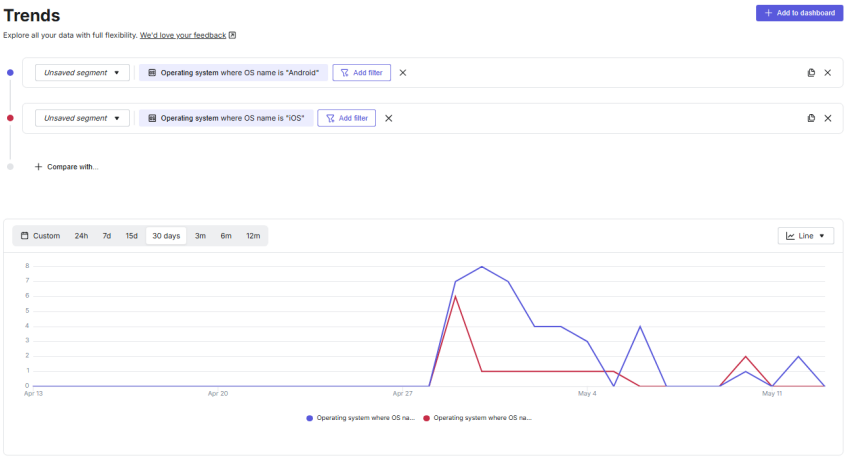


Figure 6.22: Trends in accesses to the site by operating system over time

6.3.3 Thematic Analysis of Structured Interviews (Adapted from Focus Group Plan)

As explained in the methodology chapter, the session initially planned as a focus group was carried out as a structured interview due to low attendance. The session followed the original discussion guide and lasted approximately 37 minutes. Despite the limited number of participants, the conversation offered relevant insights regarding the mobile experience of ZeroZero. The analysis below follows the structure of the original interview script which is available in Appendix C.

1. Icebreaker and General Perception

Participants were asked to compare the ZeroZero platform to a football player. Responses reflected an image of a confident, established platform with room for improvement. This metaphor introduced the session with light-hearted engagement and signaled general familiarity with Zos.

> *"I'd go with Zlatan Ibrahimović, because ZeroZero is a really good site, it's evolved a lot recently, but there are still areas that could improve. That said, according to them, they're the best, just like Zlatan used to claim he was the GOAT* [8]*. So off the top of my head, I'd say Zlatan."*
> *"Mine's not as creative, I was just thinking of Ronaldo, because the site is the most well-known in Portugal in this space. And I also know it's internationally recognized due to its multiple domains. I guess it's not that original."*

2. Usage Habits and Context

Both interviewees use ZeroZero primarily on mobile, mostly to read news, check statistics, or explore player profiles. Mobile access was preferred for speed and convenience. However, they also reported turning to other platforms like Transfermarkt or Wikipedia for easier access to specific content.

[8] GOAT is an abbreviation for Greatest Of All Time

"I use it almost exclusively on my phone... it's easier, faster..."
"Sometimes I'll go straight to Wikipedia for competitions, because even though it's not a dedicated platform, it has a lot of information, and sometimes it's easier to check there than on ZeroZero or other specialized sites."

3. Navigation and Structure

Participants expressed some frustration with content discoverability. Search results often felt poorly ranked or overloaded. The player search feature, in particular, was identified as confusing and inconsistent.

"... sometimes I type "Cristiano" and Cristiano Ronaldo isn't even the first result. You get someone else named Cristiano, and it's hard to find him by last name too. The algorithm should prioritize the more well-known names. Even when the name is simple, you often can't find who you're looking for."

They also suggested improvements such as grouping news by team or topic and offering filters for better organization.

4. Content and Visual Presentation

Comments addressed both the visual clarity and completeness of information. Participants noted occasional bugs in the Playmaker section and felt that some pages were either too dense or not updated quickly enough, especially for match summaries and statistics.

"... sometimes the news pages are too empty. For example, I like to read match stats — like who played well, the referee's performance, etc. Sometimes I go to check and there's nothing there. Then I go back later, and I can't even find the article anymore. It takes too long to update."

They also suggested including more advanced statistics, such as Expected Goals [9] or possession efficiency, to match the standard of competitors like Flashscore or SofaScore.

5. Performance and Frustrations

Participants reported frustration when comments failed to post due to slow loading, as well as when search results took too long to appear. These issues reduced engagement with some features and occasionally led users to abandon their task.

"... I write a comment with a title and everything, then click to submit and it just keeps loading… Eventually it says it was sent, but it's not there. If I go back and try to send it again, sometimes it works, but sometimes it doesn't. If it's a long comment, I usually give up and don't post it."

[9] Expected Goals (xG) is a football statistic that measures the probability of a shot resulting in a goal.

6. Ideal Vision and Suggestions

Both users envisioned a more personalized version of the platform, with content about their preferred teams or competitions appearing first. The idea of a social-media-like homepage with customized feeds was well received.

> *" ... I'd like the whole experience to be more customisable — more like a social network. For example, if I follow a player, club, or competition, I want to see content related to them."*

Additional suggestions included reviving interactive elements such as competitions or quizzes, and enabling user participation in live podcasts.

7. Closing Words

At the end of the session, participants were asked to summarize their experience with the mobile version of ZeroZero.pt in a single word or short phrase. Their responses reflected a generally positive perception, highlighting both intuitiveness and continuity with the desktop experience.

> *"Intuitive."*
> *"As close as possible to the desktop version, which I think is important — not having to re-learn everything."*

Conclusion

The structured interview, despite limited attendance, revealed clear patterns in user preferences, frustrations, and expectations. Major areas for improvement include content discoverability, real-time updates, personalization, and richer statistical features, all of which are consistent with current best practices in mobile UX.

6.4 Quantitative Data Analysis of Non-Registered Users (Questionnaire 2)

This section presents the statistical analysis of data collected from non-registered users who completed Questionnaire 2, without logging into the ZeroZero website.

6.4.1 Data Preparation

Before conducting the analysis, the dataset was reviewed to ensure the quality and interpretability of responses. One case (User 5) was identified as an outlier and excluded from the dataset. This decision was based on multiple indicators of disengagement with the questionnaire. In the open-ended questions, the participant submitted answers that did not reflect meaningful or critical feedback, such as:

- *"What would you change if you could change only one thing in the mobile version?"* — "Sporting"
- *"What was your biggest frustration while using the website?"* — "Benfica"
- *"Suggestion for a new functionality"* — "Porto"
- *"Additional comments"* — "Vitória SC"

Furthermore, all responses to the Likert scale questions (excluding the SUS) were marked with the maximum score ("5"), showing no variation across dimensions. In contrast, all items of the SUS questionnaire were answered with the same neutral value ("3"). This uniform response pattern suggests that the participant did not engage critically with the evaluation process and did not provide usable data. The case was therefore excluded on the basis that it lacked analytical value and could distort the integrity of the findings.

6.4.2 Descriptive Statistics

This subsection provides an overview of the participants' demographic and usage characteristics to contextualize the interpretation of usability and performance-related responses.

6.4.2.1 Participant Profile

Participants were predominantly between 25 and 34 years old (24 respondents), followed by the 45–54 age group (17 respondents) and the 35–44 group (8 respondents). Only 4 participants were between 18 and 24, and a minority were over 55 years old (4 combined). One respondent was under 18.

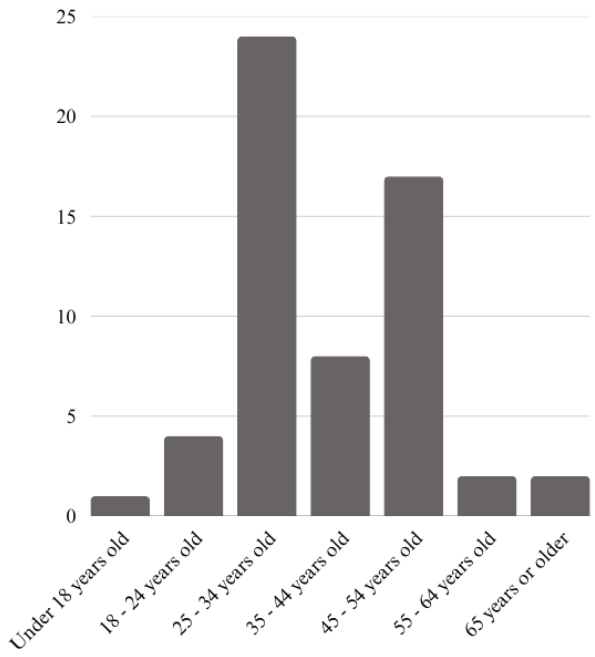


Figure 6.23: Distribution of participants by age group (Questionnaire 2).

This age distribution can be compared with publicly available analytics from SimilarWeb (May, 2025)[10], which report a slightly higher proportion of users in the 35–44 age group (23.77%) than in the 25–34 group (23.36%) among ZeroZero's general mobile audience.

Most participants reported accessing ZeroZero **several times per day** (40), while 10 stated they visit it **about once per day**. Less frequent access was rare, with 6 participants indicating they visited **several times per week**, 1 participant reporting **monthly** usage, and 1 stating they had **never used the platform before**. These results suggest that the majority of respondents were habitual users of ZeroZero.

[10] https://www.similarweb.com/website/zerozero.pt

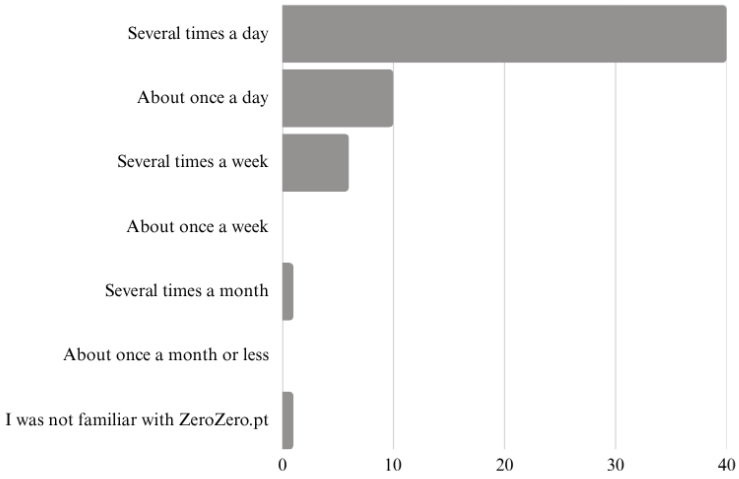


Figure 6.24: Distribution of visit frequency to ZeroZero.pt (Questionnaire 2).

Regarding their level of familiarity with the ZeroZero platform, participants were asked to respond using a five-point Likert scale, where 1 indicated no familiarity and 5 indicated very high familiarity. The majority of respondents demonstrated a high degree of familiarity, with 21 selecting option 5 and 30 selecting option 4. A smaller number of participants chose intermediate or low values, with 6 selecting option 3 and only 1 participant selecting option 1. This distribution suggests that, overall, participants were generally well-acquainted with the platform, which aligns with the previously reported high frequency of use.

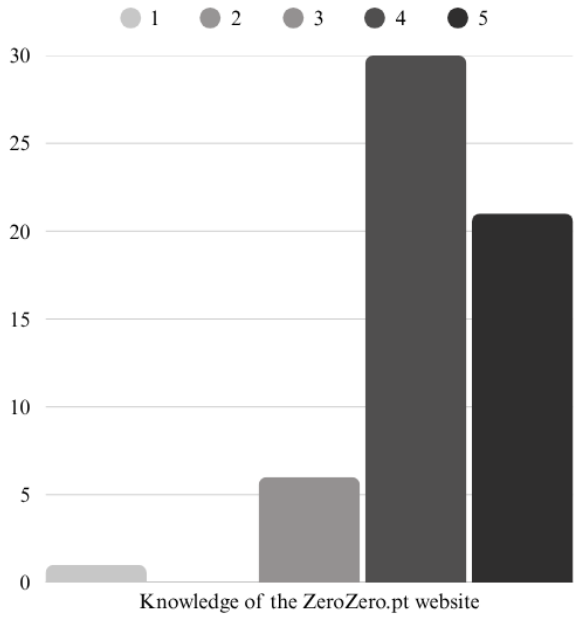


Figure 6.25: Self-reported familiarity with ZeroZero.pt on a 5-point Likert scale (Questionnaire 2).

6.4.3 System Usability Scale (SUS)

A histogram with an overlaid normal curve was generated to visually assess the distribution of SUS scores (Figure 6.26). While the visual distribution appears approximately symmetric, it does not perfectly align with the normal curve. This visual impression is consistent with the results of the Shapiro–Wilk test, which indicated a statistically significant deviation from normality ($p = .017$). Despite the slight negative skew observed, the majority of scores clustered around the mean (M = 76.03, SD = 16.70, N = 58), suggesting that participants generally perceived the system as usable.

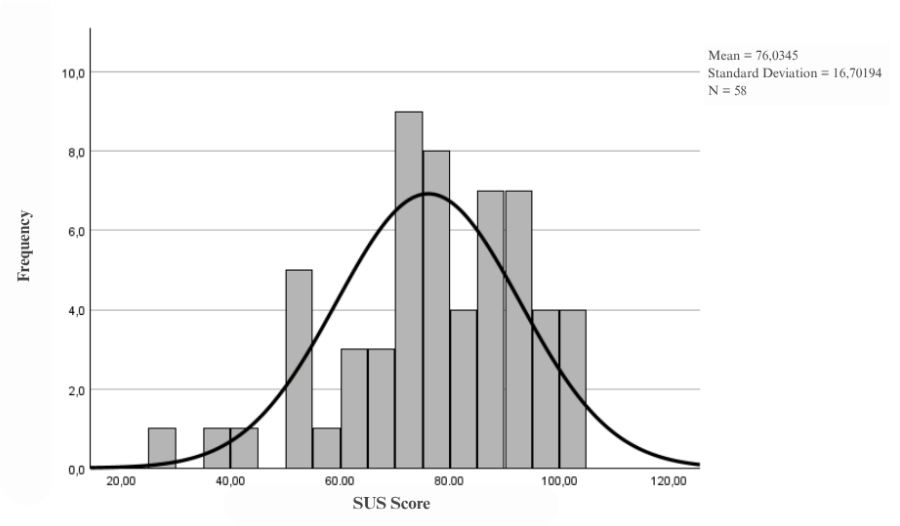


Figure 6.26: Distribution of SUS Score (Questionnaire 2)

To facilitate interpretation, SUS scores were grouped into letter grades, following the original suggestion by Brooke, 1996 that SUS scores can be meaningfully interpreted using school-style grading systems. As illustrated in Figure 6.27, a considerable proportion of participants (37.9%) received an A+ grade, denoting excellent usability. Grades spanned the full range, including lower evaluations such as D and F (each with 10.3%), indicating the presence of diverse user experiences. Although Brooke's original proposal did not specify exact score intervals for each grade, this approach has since been expanded by Sauro et al., 2012, whose curved grading scale was adopted in this study. This distribution suggests that while the overall perception of usability was positive, some users still encountered notable issues during their interaction with the mobile version of ZeroZero.

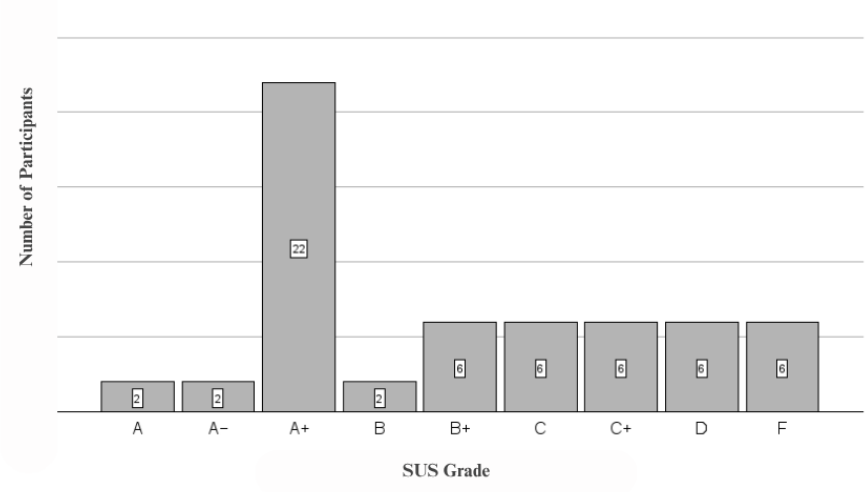


Figure 6.27: Distribution of SUS Grades (Questionnaire 2)

6.4.4 Normality Tests

To determine the appropriate statistical tests, normality was assessed using both the Shapiro–Wilk and Kolmogorov–Smirnov tests, applied to three continuous variables: SUS Score, Task Success Percentage, and Visit Frequency. As shown in Table 6.9, all three variables exhibited significant deviations from a normal distribution. Specifically, for both Task Success Percentage and Visit Frequency, the Shapiro–Wilk and Kolmogorov–Smirnov tests reported $p < .001$, indicating clear violations of the normality assumption. The SUS Score, although closer to a normal distribution, also failed the Shapiro–Wilk test with $p = .017$.

Table 6.9: Normality Tests for Key Variables (N = 58)

Variable	**Test**	**Statistic**	**Sig. (p)**
SUS Score	Kolmogorov–Smirnov	.106	.164
	Shapiro–Wilk	.950	**.017**
Task Success %	Kolmogorov–Smirnov	.492	$< .001$
	Shapiro–Wilk	.437	$< .001$
Visit Frequency (0–6)	Kolmogorov–Smirnov	.382	$< .001$
	Shapiro–Wilk	.564	$< .001$

These results were further confirmed through graphical inspection. The histogram for Task Success Percentage (Figure 6.28) clearly illustrates a ceiling effect, with a strong concentration of responses at 100%. This visual evidence aligns with the statistical rejection of the normality assumption.

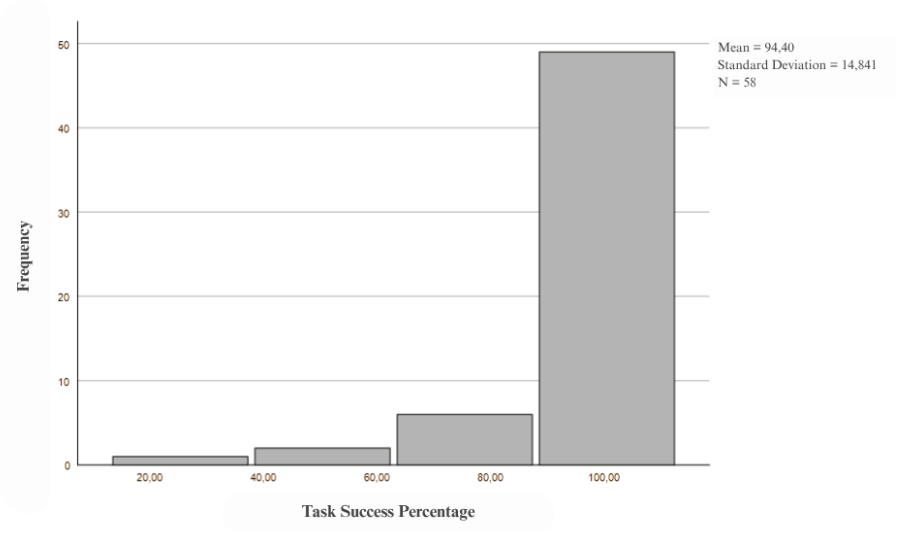


Figure 6.28: Histogram of Task Success Percentage.

Given these results, parametric tests (such as the independent samples t-test, ANOVA, Pearson correlation, or linear regression) were deemed inappropriate, as they assume normally distributed variables or residuals. Therefore, non-parametric alternatives were adopted in line with Field, 2018 recommendations for non-normal data.

6.4.5 Inferential Analysis

Inferential statistical analyses were conducted using IBM SPSS Statistics (version 29), in Section 6.4.4, non-parametric tests were employed. These included the Mann–Whitney U test for comparing two independent groups, the Kruskal–Wallis H test for comparisons involving more than two groups, and Spearman's rank-order correlation to assess associations between ordinal or non-normally distributed continuous variables. The results presented in this section aim to explore potential relationships and differences between variables relevant to the user experience on ZeroZero's mobile platform.

6.4.5.1 Comparison of Perceived Usability Based on Task Success (Mann–Whitney U Test)

To assess whether the level of task success influenced users' perceived usability, a binary variable was derived from the original task success percentage. Participants who achieved a full 100% success rate across the four tasks were grouped as "Complete Success" (coded as 1), while those with any failed task were grouped as "Incomplete Success" (coded as 0).[11]

A Mann–Whitney U test was conducted to compare SUS scores between the two groups. Results showed that participants who completed all tasks successfully (n = 49) had a higher mean rank in SUS scores (30.60) than those who did not (n = 9, mean rank = 23.50). However, the

[11] Only users who successfully completed all four tasks were coded as "Complete Success". Any deviation from 100% was categorized as "Incomplete Success".

difference was not statistically significant, U = 166.500, Z = -1.162, p = .245 (two-tailed).These results are summarized in Table 6.10.

Table 6.10: Results of Mann–Whitney U Test for SUS Score by Task Success (Q2)

Group	**N**	**Mean Rank**	**Sum of Ranks**
Incomplete Success (0)	9	23.50	211.50
Complete Success (1)	49	30.60	1499.50
Test Statistic	U = 166.500, Z = -1.162, p = .245 (2-tailed)		

Although the result was not statistically significant, the observed difference in mean ranks suggests that participants who successfully completed all tasks tended to report slightly higher usability scores. The distribution of SUS scores across both groups is illustrated in Figure 6.29.

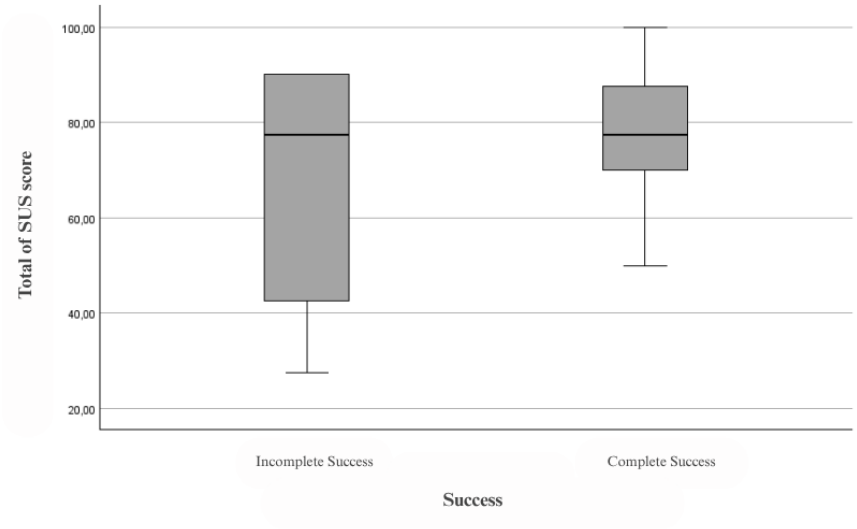


Figure 6.29: Boxplot of SUS scores by task success group (Questionnaire 2)

6.4.5.2 Comparing SUS Scores Based on Platform Familiarity (Kruskal–Wallis Test)

To explore whether users' perceived usability (measured by the SUS Score) differed according to their level of familiarity with the ZeroZero platform, a Kruskal–Wallis test was conducted. The variable representing familiarity was measured on a five-point Likert scale, where 1 indicated no prior knowledge and 5 indicated very high familiarity.

Four groups were included in the analysis (Levels 1, 3, 4, and 5), with the following sample sizes: Level 1 (n = 1), Level 3 (n = 6), Level 4 (n = 30), and Level 5 (n = 21). Table 6.11 summarizes the mean SUS ranks by familiarity level.

The results of the Kruskal–Wallis test indicated no statistically significant difference in SUS scores across the familiarity levels, $H(3)$ = 5.511, p = .138. Although the trend in mean ranks

suggests that users with greater familiarity tended to rate usability more positively, this effect was not strong enough to reach statistical significance.

Table 6.11: Kruskal–Wallis Test for SUS Score by Knowledge Level (Questionnaire 2)

Knowledge Level	**N**	**Mean Rank**
1 (No familiarity)	1	3.00
3	6	20.75
4	30	29.08
5 (Very high familiarity)	21	33.86
Test Statistic	$H(3) = 5.511$, $p = .138$	

Although not conclusive, the distribution of scores (Figure 6.30) visually reinforces the observed trend: participants with higher levels of familiarity with the platform tended to report higher perceived usability scores.

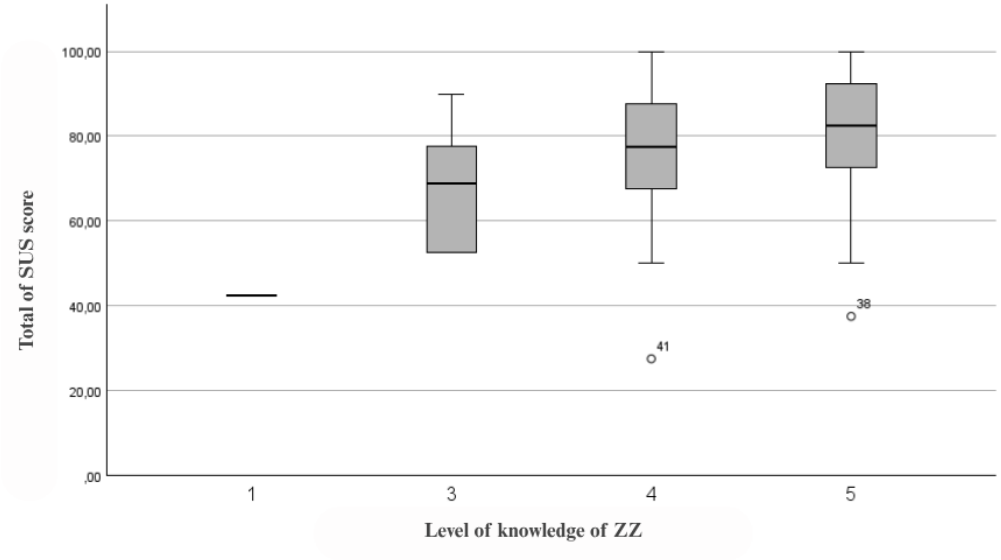


Figure 6.30: Boxplot of SUS scores by knowledge level about the ZeroZero platform (Questionnaire 2).

6.4.5.3 Spearman's Correlation Between Frequency of Visit and Perceived Usability

To explore the relationship between how frequently participants use the ZeroZero platform and their perception of usability, a Spearman's rank-order correlation was conducted. The frequency of visit was measured on a scale from 0 to 6, with higher values indicating more regular usage, and usability was assessed through the SUS Score.

The analysis revealed a weak but statistically significant positive correlation between frequency of visit and SUS score, $\rho = .298$, $p = .023$, $N = 58$ (see Table 6.12). This suggests that users who visit the platform more often tend to report slightly higher perceived usability.

Although the strength of the association is modest, the result indicates that users who visit the platform more frequently tend to report slightly higher usability scores.

Table 6.12: Spearman's Rank-Order Correlation Between SUS Score and Frequency of Visit (Questionnaire 2)

	SUS Score	**Frequency of Visit**
SUS Score	1.000	.298*
Frequency of Visit	.298*	1.000
p-value (2-tailed)	–	.023
N	58	58

*Correlation is significant at the 0.05 level (2-tailed).

6.4.5.4 Association Between Type of Internet Access and Perceived Loading Time: Chi-Square Test of Independence

To investigate whether the type of internet access influenced users' perception of loading speed on the ZeroZero mobile platform, a chi-square test of independence was conducted. The variable *type of internet access* comprised three categories: Mobile Data (3G, 4G, 5G), Wi-Fi, and Wired Connection. The perceived loading time was grouped into three ordinal categories [12]: *< 1 second*, *1–5 seconds*, and *5–10 seconds*.

The test was based on 58 valid responses. The cross-tabulation between the two variables is shown in Table 6.13, and a visual representation is provided in Figure 6.31.

Table 6.13: Cross-tabulation of Internet Access Type and Perceived Loading Time (Questionnaire 2)

Internet Access Type	**< 1s**	**1–5s**	**5–10s**
Mobile Data	5 (18.5%)	16 (59.3%)	6 (22.2%)
Wi-Fi	8 (26.7%)	19 (63.3%)	3 (10.0%)
Wired	0 (0.0%)	0 (0.0%)	1 (100.0%)

The Pearson chi-square test did not reveal a statistically significant association between the type of internet access and the perceived loading time, $\chi^2(4, N = 58) = 6.575$, $p = .160$. Additionally, 4 of the 9 cells (44.4%) had expected counts below 5, suggesting that the assumptions of the test may have been partially violated and that results should be interpreted with caution.

[12] Time categories were defined based on established usability thresholds for system response times, as recommended by Nielsen Norman Group: https://www.nngroup.com/articles/website-response-times/

Despite the absence of statistical significance, some descriptive differences were observed. For instance, Wi-Fi users were proportionally more likely to report loading times under one second compared to mobile data users. Conversely, users on mobile data appeared more frequently in the 5–10 second category. The only participant with a wired connection reported the slowest loading time, although this single case prevents any general conclusion.

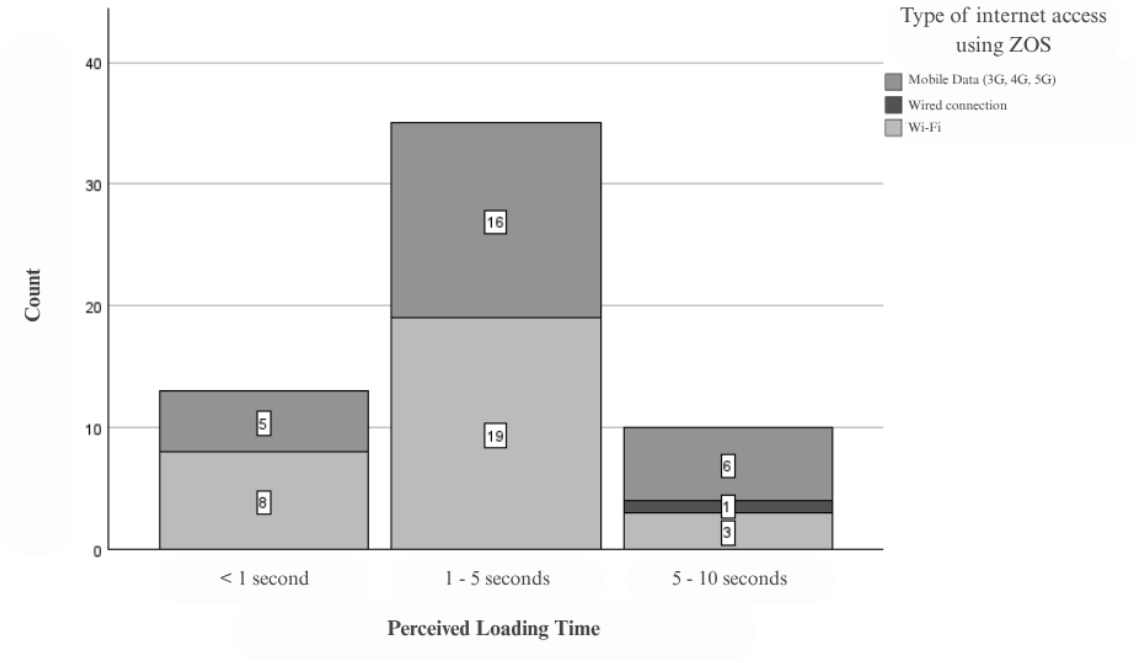


Figure 6.31: Stacked Bar Chart of Perceived Loading Time by Internet Access Type (Questionnaire 2)

In summary, while no statistically significant relationship was established, the observed trends may still offer useful exploratory insights into how perceived performance varies across different types of network access.

6.4.5.5 Correlation Between Perceived System Performance and Usability (Spearman)

To examine the relationship between users' perception of system performance and their perceived usability of ZeroZero's mobile version, a Spearman's rank-order correlation was conducted. Two indicators of performance perception were analyzed: perceived content loading experience and perceived unexpected page delays. The usability perception was measured through the SUS score.

The null hypothesis (H_0) stated that there is no monotonic association between perceived system performance and usability scores. The alternative hypothesis (H_1) proposed that a statistically significant monotonic relationship exists between these variables.

The correlation analysis revealed no statistically significant relationship between SUS score and either of the performance perception variables (see Table 6.14). Specifically, the correlation between content loading experience and SUS score was weak and non-significant, $\rho(58) = .117$,

p = .382. Similarly, no association was observed between perceived unexpected page delays and SUS score, $\rho(58)$ = .016, p = .903.

Table 6.14: Spearman's Correlation Between System Performance Perception and SUS Score (Questionnaire 2)

Variable	**ρ (rho)**	**p-value**	**N**
Content loading experience	.117	.382	58
Perceived unexpected page delay	.016	.903	58

Figures 6.32 and 6.33 provide boxplots illustrating the distribution of SUS scores across the different levels of performance perception. In both cases, no clear trend or pattern is visible, and the data do not suggest a meaningful relationship between system performance perception and usability evaluation.

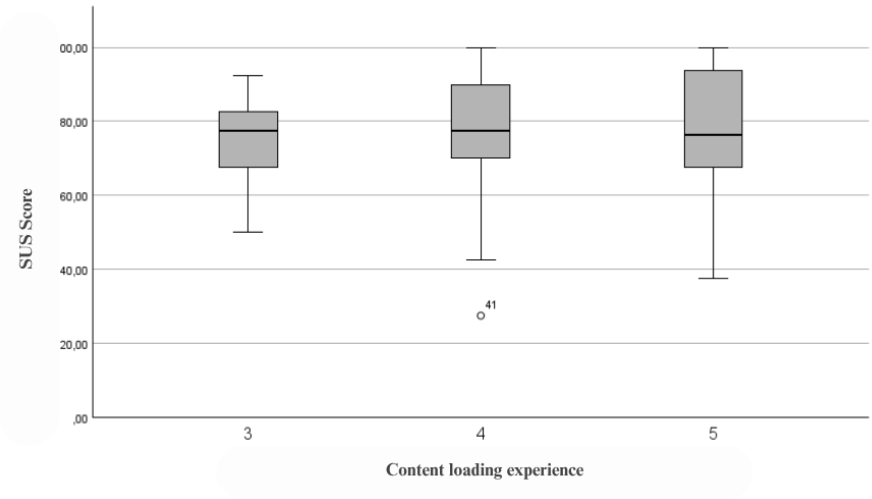


Figure 6.32: Boxplot of SUS scores by perceived content loading experience (Questionnaire 2)

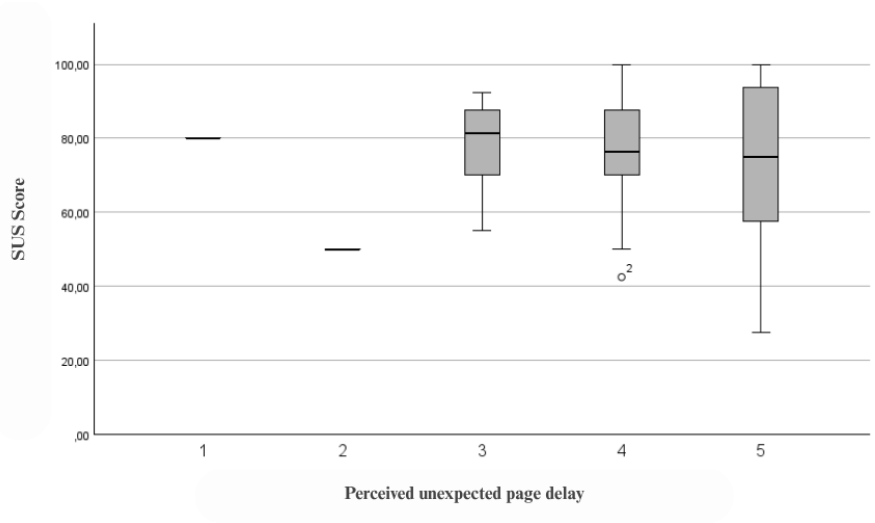


Figure 6.33: Boxplot of SUS scores by perception of unexpected page delays (Questionnaire 2)

6.5 Qualitative Data Analysis of Non-Registered Users (Questionnaire 2)

This section focuses on the qualitative data collected from non-registered users. As with the previous analysis, it covers open-ended questionnaire responses and behavior tracking through Hotjar. However, due to limitations in the available Hotjar features during this phase, the depth of behavioral insights is comparatively reduced.

6.5.1 Open-ended Questionnaire Responses

This subsection presents the qualitative analysis of the collected open responses. The goal of this analysis is to complement the quantitative data by exploring more nuanced perspectives, with a particular focus on identifying recurring usability issues, desired improvements, and overall user sentiment regarding the mobile version of the platform.

Participants were invited to respond to four open-ended questions:

1. What users would change in the mobile version,
2. Their biggest frustration,
3. Suggested new feature, and
4. Additional comments.

As with the previous questionnaire, a thematic coding approach was applied to the responses. Categories emerged inductively from the content, enabling a grounded understanding of user concerns and suggestions. The most frequent themes included: Interface/Layout, Ads/Pop-ups, Search, Mobile App, Personalization, Live Updates, Navigation, Advanced Stats, Performance,

Content Quality, Suggested Features, User Support & Moderation, No Change, and Nothing to add.

Each question is analyzed independently in the following subsections.

1. Question: What would you change in the mobile version?

When asked what they would change in the mobile version, a significant number of respondents (14) indicated "No Change", suggesting a generally favorable perception of the current mobile experience. Nevertheless, a broad range of detailed suggestions from the remaining responses revealed meaningful areas where improvements could enhance usability, clarity, and satisfaction.

The most cited theme was Interface/Layout (12). Users pointed out that the platform presents excessive content and complex visual arrangements that can be overwhelming—especially for less experienced users. Specific comments mentioned the homepage being too long or cluttered, the information density being too high, and elements like the slides and competition listings contributing to confusion. Several suggested simplifying the layout, reducing redundant links, and adopting a more intuitive structure for navigating between leagues or accessing key sections.

Ads and Pop-ups (7) were also a source of frustration. Respondents mentioned that excessive or poorly placed advertising caused involuntary clicks, altered the page layout during loading, and disrupted the reading experience. Some explicitly requested an "ad-free" version of the site, stating that the intrusive nature of ads negatively impacted usability.

Content Quality was referenced in 5 responses. Comments within this theme included a desire for clearer writing, more concise data presentation, and improved separation between content types (e.g., male vs. female football). One participant expressed a preference for written summaries of press conferences rather than only having audio formats such as podcasts.

Performance-related issues (4) were also highlighted, particularly regarding slow loading times during peak periods or when accessing live content. Users pointed out that matchdays—when traffic increases—often led to delayed responses and a slower user experience.

Live Updates appeared in 4 responses as well. Suggestions included making live scores more dynamic and accessible (e.g., updates by league tier or more comprehensive in-match data for lower-profile competitions). These users emphasized the value of receiving real-time content that remains consistent across different league levels.

Navigation and Search each emerged in 3 responses. Navigation-related concerns centered on the complexity of accessing desired competitions or pages, with one user suggesting that common league or club pages should be directly accessible via the main menu without navigating through multiple layers. Search-related feedback called for improved precision and visibility of the search bar to help users find players, matches, or statistics more effectively.

Although mentioned less frequently, Mobile App (3), Advanced Stats (2), and Personalization (1) also appeared. Users who brought up the mobile app often expressed interest in a dedicated application to enhance performance and functionality. Suggestions for advanced statistics included deeper analytical insights such as individual player metrics. The single response on personalization suggested enabling favorites or commonly visited pages for quicker access.

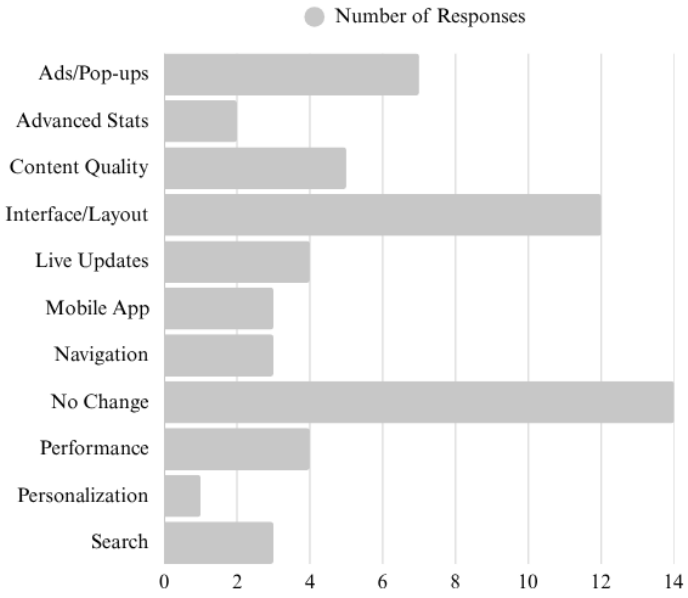


Figure 6.34: Suggested changes to the mobile website (Questionnaire 2)

As illustrated in Figure 6.34, while many users are generally satisfied with the mobile version, the thematic feedback provides concrete directions for optimizing layout, minimizing disruption from advertisements, improving content accessibility, and enhancing performance.

2. Question: What was your biggest frustration when using the website on mobile?

A considerable portion of respondents (14) indicated having no major frustrations while using the mobile version of the platform, reflecting a relatively positive overall experience. Nevertheless, the remaining answers brought to light specific and recurring pain points that negatively impacted user interaction.

The most frequently mentioned theme was Performance (11 responses). Users described slow page loading (especially during high-traffic moments like match days) as a core issue. Others mentioned pages failing to load, system lags, or the need to refresh pages manually to view updated information. This performance bottleneck appeared particularly problematic when users tried to access features such as comment sections, real-time scores, or content-heavy pages.

Ads and Pop-ups were also a significant source of frustration, noted by 10 respondents. Many described intrusive and poorly placed advertising content that disrupted navigation, covered vital areas of the screen, or caused accidental clicks. Some expressed willingness to pay for an ad-free version, citing the impact of excessive ads on readability and overall usability.

Interface and Layout challenges (4) emerged in responses highlighting visual confusion, sliding inconsistencies, or difficulty engaging with slide-based news formats. Users found certain layout decisions (like scrolling behavior or the presence of too many slides) disruptive or unintuitive.

In terms of Navigation (4), users mentioned the complexity of locating specific content such as lower-tier leagues or live scores. The need to click through multiple layers to access basic information was viewed as inefficient, especially by those unfamiliar with the site's structure.

Concerns about Content Quality (4) were also present. Some users criticized vague article headlines that failed to clarify which match or league was being referenced, leading to disinterest after clicking. Others noted that coverage of certain competitions was limited or outdated, particularly for youth teams or niche sports.

Although fewer in number, users also raised frustrations related to Search (2), which they described as imprecise or unhelpful in locating content such as summaries or match statistics. A couple of respondents mentioned issues with User Support and Moderation (2), specifically pointing to what they perceived as inconsistent moderation practices in the comment section. One user described a long-standing ban without a clear justification, contrasting it with other visible rule violations that went unpenalized.

Other isolated points included mentions of the Mobile App (1), with one respondent highlighting the lack of a dedicated app, and Personalization (1), referencing difficulties in tailoring content or distinguishing between male and female football news.

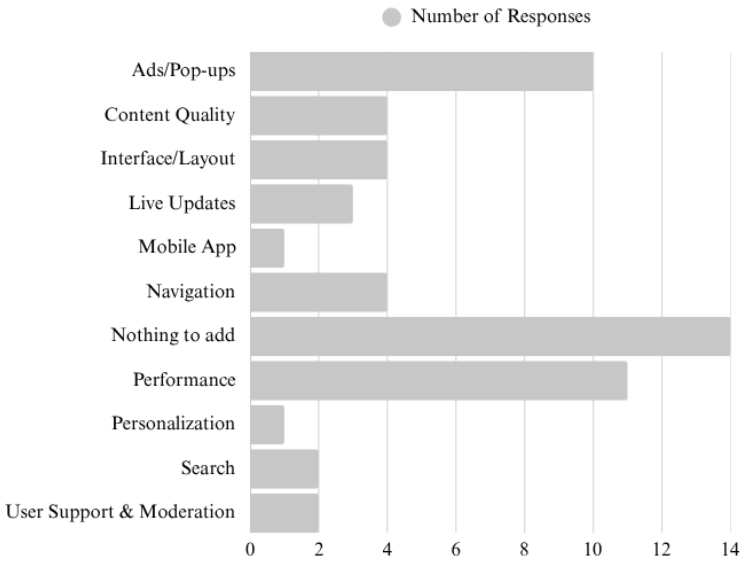


Figure 6.35: Biggest user frustrations with the mobile website (Questionnaire 2)

Despite the variety of issues highlighted, as visualized in Figure 6.35, the data underscores the importance of improving load speeds, minimizing disruptive ads, and refining structural clarity (especially in navigation and real-time interaction) so as to foster a smoother and more engaging mobile experience.

3. Question: Do you have any suggestion for a new feature?

Most participants (37 out of 56) responded with "Nothing to add", suggesting a relatively high level of satisfaction with the existing features of the platform. However, among those who did provide input, several recurring themes emerged—pointing to targeted opportunities for enhancement.

The most frequently mentioned theme was Content Quality (8). Users recommended improvements in how information is structured and presented. Suggestions included the integration of new sports (e.g., tennis, rugby, athletics, motorsports), the addition of contextual metadata in career stats (e.g., number of matches as a substitute), and better differentiation between male and female competitions. Some users also advocated for the inclusion of a dedicated transfer rumours page, clearer categorization of news articles, and enhanced visibility of women's football content.

Other suggestions reflected an interest in deeper data insights, with Advanced Stats (2) cited as a desired area for development. Respondents proposed metrics such as the average number of shots on target converted into goals per team or expanded historical and career-related information for players, coaches, and club directors.

Two users pointed to the need for greater Personalization, such as customizable interfaces with quick-access links to favorite teams, players, and competitions. These users noted that the benefits of being logged in are currently limited and proposed that login functionalities be better leveraged to enhance user experience.

Navigation (2) was another relevant category, with requests for persistent menus across all pages and improved shortcut access to commonly visited areas. A single suggestion under Interface/Layout proposed a more minimalistic homepage design, supported by clearer content segmentation via tags.

Further, a few responses focused on platform accessibility. One user requested a dedicated Mobile App to streamline access, while another emphasized the need for an Ad-Free Version, directly reflecting a concern categorized under Ads/Pop-ups.

Some users also mentioned Suggested Features (2), such as launching a public API for third-party usage or introducing a real-time match clock for games. These responses underscore how certain user groups are eager for technically advanced tools that go beyond basic consumption.

Finally, Search (1) was highlighted by a user who called for a more prominent and accessible search bar to improve navigability and content discovery.

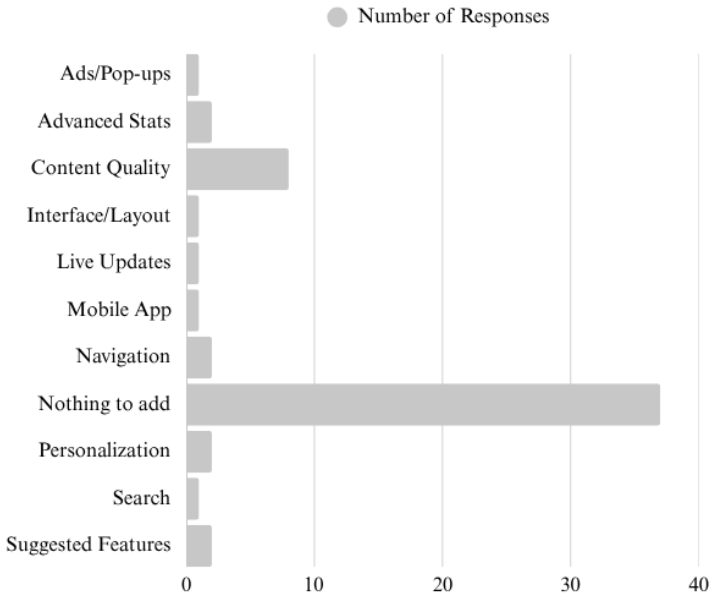


Figure 6.36: Suggested new feature (Questionnaire 2)

Despite the limited number of responses, the suggestions provided were highly specific and actionable, demonstrating that even satisfied users have a clear vision for how the platform can evolve—particularly in terms of information structure, customization, and the introduction of new verticals and features.

4. Question: Additional Comments?

The vast majority of participants (54 out of 58) chose not to leave any additional comments, which may be attributed to the detailed feedback already provided in previous questions.

Among the few who did contribute further thoughts, one user reiterated a strong willingness to financially support an ad-free version of the platform, echoing earlier frustrations with intrusive advertising. A suggestion for a public API was also made (delivered with an enthusiastic tone) highlighting a desire from more technically engaged users to access structured data directly.

Navigation was addressed again, with one participant suggesting that key competitions and teams should be more immediately accessible via icons or shortcuts, rather than being hidden behind submenus. This comment also raised broader concerns about the platform's intuitiveness, describing navigation as cognitively demanding and potentially off-putting for less experienced users. Interestingly, the same comment also touched on the communication strategy of the platform, encouraging the use of alternative channels to social media—especially Twitter/X—due to its perceived toxicity.

Other remarks included appreciation for the project and its openness to user feedback, with one participant explicitly praising ZeroZero for being the only sports platform they consistently

follow. Additionally, a few comments advocated for improved real-time coverage of minority sports and more attention to niche areas such as roller hockey and women's sports.

Figure 6.37: Additional Comments (Questionnaire 2)

Although limited in number, these additional remarks underscore some recurring themes already identified, particularly the need for cleaner navigation, enhanced content accessibility, and support for underrepresented sports. They also reflect a generally positive sentiment among registered users, many of whom demonstrate high levels of engagement and care for the platform's evolution.

6.5.2 User Behavior Observations (Hotjar)

Between 27 May at 10:46 and 05 June at 20:39, a total of 45 user sessions were recorded using Hotjar on the mobile version of ZeroZero. This observation period followed the second questionnaire and aimed to collect behavioral insights to complement the self-reported data.

Although the questionnaire collected 58 valid responses, only 45 sessions were captured through Hotjar. Of these, 36 were conducted on Android devices and 9 on iOS devices. This discrepancy is explained by a technical constraint: session recordings only occurred when users accessed the mobile version of ZeroZero via the specific URL provided within the questionnaire. This link redirected participants to a dedicated observation environment, where Hotjar tracking was active. Consequently, users who completed the tasks by accessing the platform through other means (such as search engines, bookmarks, or direct entry) were not tracked, and their sessions were not recorded. This represents a limitation in the behavioral dataset, which may not fully represent the diversity of task completion strategies among all participants.

Due to the unavailability of premium features in Hotjar during this phase, it was not possible to generate user funnels or trend analysis. The data analysis was therefore limited to session replays and heatmaps.

6.5.2.1 Page-by-Page Interaction Analysis

The recorded sessions correspond to pages that participants were explicitly instructed to visit during the task-oriented phase of the second questionnaire. Therefore, all behaviors must be interpreted in the context of *directed interactions*, not organic or spontaneous navigation patterns.

The homepage heatmap analysis, based on **41 recorded sessions**, revealed a strong concentration of interactions in the *top segments of the page*, especially around the main navigation menu (27 taps), the search bar (9 taps), and the multi-sport selector (8 taps). Most interactions occurred within the first few screen lengths, with limited tap activity observed further down the page.

The **Liga Portugal competition page**, with 19 sessions and 19 taps, showed low interaction with content blocks such as standings, statistics, or FAQs. The majority of taps were located in the upper portion of the page, mainly on the navigation menu.

Team-related pages showed different patterns. The **FC Porto overview page**, accessed via the menu (18 sessions), recorded only 4 taps and an average time on page of 9 seconds. By contrast, the related **Calendar and Results** subpage had 13 sessions and 47 taps, where interactions were mainly concentrated on the match listings.

The **Viktor Gyökeres player profile**, based on 15 sessions and 18 taps, had a longer average session time (55 seconds), with limited but distributed interactions across the stats and content areas.

Finally, the **Portugal national team page** (12 sessions, 4 taps) registered few interactions, with taps focused primarily on the top-level interface elements.

6.5.2.2 Summary of Observed Patterns

Across the observed pages, most interactions occurred in the upper areas of each screen. Taps were generally focused on navigation elements and task-relevant sections, depending on the questionnaire instructions. Pages containing match listings or structured statistical data received a higher number of taps, while more general pages recorded lower interaction levels. These patterns reflect the specific nature of the tasks participants were asked to perform and do not necessarily indicate broader user preferences or usability issues.

6.5.3 Thematic Analysis of Structured Interviews (Adapted from Focus Group Plan)

To complement the quantitative and behavioral data, a structured interview was conducted with one participant who had completed the second questionnaire (targeting non-registered users). Although two individuals had confirmed attendance, only one participated in the session, which lasted approximately 20 minutes. The interview followed a semi-structured script and aimed to explore user experiences with the mobile version of ZeroZero.pt in greater depth.

The session began with an informal ice-breaker to ease the participant into the discussion:

> **Ice-breaker question:** *"If ZeroZero.pt were a football player, who would it be and why?"*
>
> **Participant:** *"It would be Andrea Pirlo, because he was a player who organized the team a lot and always dictated the rhythm of the game."*

This analogy helped set the tone for the conversation, highlighting the participant's perception of ZeroZero as a structured and guiding tool within the sports information space.

The thematic findings below are structured around the six discussion areas covered during the session:

1. **Usage Habits and Context**

   The participant reported using ZeroZero.pt almost exclusively on mobile, describing it as a daily sports resource akin to a "sports newspaper." Football was the primary area of interest, especially for news and occasional data lookups. The user noted that in some sports, the platform lacks real-time updates, prompting occasional use of alternative websites.

2. **Navigation and Information Structure**

   Navigation was described as intuitive and consistent. Familiarity with the site's structure—especially due to long-term use—helped the participant find relevant information quickly. Although some less prominent content areas required more effort to locate, the overall structure was seen as coherent and helpful.

3. **Content Clarity and Visual Experience**

   The participant highlighted issues with slideshow functionality and occasional formatting problems, such as news articles requiring screen rotation to be readable. It was also pointed out a perceived decline in editorial quality, with more frequent spelling and grammar errors. Despite these issues, the mobile layout was considered clean and conducive to fast reading, except for intrusive pop-ups and excessive suggested links at the end of articles.

4. **Performance and Frustration Points**

   No major technical problems were recalled, but the user noted slower live updates as a source of frustration—particularly during live matches. Compared to platforms like SofaScore, ZeroZero was seen as lagging behind in terms of live performance. Mixing videos and written news in the same feed was also flagged as a minor annoyance.

5. **Ideal Vision and Recommendations**

   The participant envisioned a more personalized interface that would prioritize content related to their interests or favorite teams. However, it was also stressed the importance of

maintaining access to diverse content to allow for unexpected discovery. Suggestions included improving coverage of less prominent sports (e.g., beach soccer) and enriching the historical database (Encyclopedia) with more frequent updates.

While SofaScore and Flashscore were mentioned as useful alternatives for live updates, the participant expressed a clear preference for ZeroZero to retain its distinct identity. Its statistical depth and historical content were considered key strengths that differentiate it from other platforms.

6. **Summary Impressions**

When asked to summarize their experience in one phrase, the participant responded: *"A good experience, no doubt."*. This sentiment reflects the overall tone of the interview—constructive and positive, with appreciation for the platform's role and potential.

6.6 Limitations

While the mixed-methods approach adopted in this study provided valuable insights into the mobile user experience of ZeroZero, several limitations must be acknowledged regarding both the quantitative and qualitative datasets.

From a quantitative standpoint, both questionnaires relied on non-probabilistic, voluntary sampling. As such, there is a risk of self-selection bias, particularly favoring users with strong opinions or higher engagement with the platform. Moreover, the deliberate separation between registered and non-registered users, though methodologically justified, introduces a degree of asymmetry: registered users (Questionnaire 1) tend to be more experienced, while non-registered users (Questionnaire 2) may include more casual or first-time visitors. Consequently, each dataset captures only a partial perspective of the broader user base. Additionally, several variables violated assumptions of normality, leading to reliance on non-parametric statistical methods, which are more conservative and may limit the detection of subtle effects.

On the qualitative side, the study combined data from open-ended questionnaire responses, structured interviews, and behavioral analytics via Hotjar. While this triangulation enriched the interpretation of usability challenges, each data source presented its own constraints. Open-ended responses were often brief or vague, limiting their analytical depth. Thematic categorization, conducted by a single researcher, may introduce some degree of subjectivity despite efforts at consistency. The interviews, while richer in nuance, reflected a small, self-selected subset of users and may not be representative. Similarly, Hotjar data, though useful in identifying friction points and navigation bottlenecks, is inherently observational and lacks access to users' intent or motivations, limiting its interpretability without complementary verbal input .Moreover, only users who accessed the platform through the specific link that embedded the tracking were recorded, those who visited the website directly (e.g., via organic search or saved bookmark) were not tracked, potentially leading to an incomplete or biased sample of user behavior.

Taken together, these limitations suggest that while the study offers a multi-layered view of usability and user perception, the findings should be interpreted with caution. Future research could benefit from larger and more diverse samples, collaborative coding of qualitative data, and experimental or longitudinal designs to better isolate causal relationships and evaluate design interventions.

6.7 Summary

This chapter provided a comprehensive analysis of the data collected throughout the study, combining quantitative, qualitative, and behavioral insights to assess the user experience of ZeroZero's mobile platform.

Two questionnaires were administered to distinct user groups (registered and non-registered users) allowing for comparative insights across engagement profiles. Both groups demonstrated high familiarity with the platform and frequent usage patterns, though subtle differences emerged in terms of usability perceptions and reported frustrations.

Quantitatively, SUS scores indicated generally positive evaluations of usability, with mean scores exceeding industry benchmarks for both user groups. However, inferential analyses revealed no statistically significant differences in perceived usability based on task success or familiarity, though some trends suggested that greater engagement and task completion may be associated with more favorable usability ratings. Statistically significant associations were observed between frequency of visits and SUS scores in both questionnaires, suggesting that habitual use may influence perceived ease of use. A significant association between connection type and loading time was found only in Questionnaire 1.

Qualitative findings reinforced and contextualized the quantitative data. Open-ended responses revealed recurring themes such as performance limitations, intrusive advertising, and difficulties in search and navigation. Despite many users expressing satisfaction with the current mobile experience, clear suggestions for improvement emerged, including the need for a mobile app, enhanced personalization, and richer real-time features. Structured interviews echoed these concerns while also highlighting the perceived value of ZeroZero's statistical depth and interface familiarity.

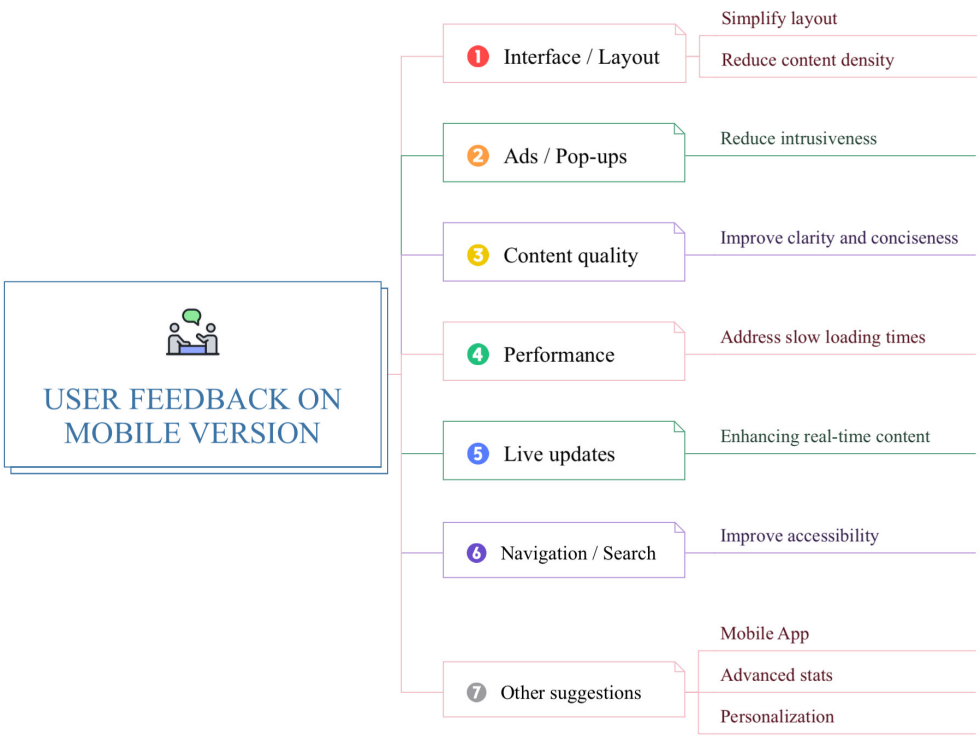


Figure 6.38: Categorization of user feedback on the mobile version of ZeroZero.

Behavioral data collected via Hotjar supported these findings, particularly in identifying top-heavy interaction patterns and drop-offs in engagement beyond initial page sections. While more advanced behavioral tracking was available during the first questionnaire phase, the second phase was limited to basic heatmap analyses due to the absence of premium tracking features.

Finally, this chapter acknowledged several methodological limitations, including sampling constraints, potential biases in self-reported data, and limited representativeness in behavioral tracking. Despite these challenges, the triangulated methodology yielded robust and actionable insights into the strengths and weaknesses of the ZeroZero mobile experience.

Chapter 7

Discussion

This chapter brings together the findings across the five usability dimensions (Learnability, Efficiency, Memorability, Error Prevention & Recovery, and Satisfaction) and evaluates them in light of ISO 9241-11 and Nielsen's heuristics. Empirical results are compared with existing studies to reveal areas where the mobile interface of ZeroZero.pt aligns with or departs from established UX principles. Based on these insights, evidence-based recommendations for design enhancements are presented, and potential directions for further investigation are outlined.

7.1 Overview and Thematic Clustering

Drawing together the quantitative and qualitative evidence from Chapter 6, five usability themes emerged, corresponding to Nielsen, 1993 dimensions of Learnability, Efficiency, Memorability, Error Prevention & Recovery, and Satisfaction. Key data points include:

- **Learnability**: Several open-ended responses from Questionnaire 1 (Fig. 6.12) mentioned experienced search autocomplete failures, some also noted that elements like the hamburger menu or the lack of a search history made navigating the site slower and more cumbersome than expected. Some open-ended responses from Questionnaire 2 (Fig. 6.34) requested enhanced search functionality.
- **Efficiency**: Scroll-depth analysis showed that fewer than 25 % of Hotjar (Sub-subsection 6.3.2.2) users reached the footer, implying that roughly three quarters dropped off before the end of the page. In the open-ended responses from questionnaire 2 , 11 users cited slow page loading as top frustration (Fig. 6.35). Interviewees reported initial confusion with terminology and menu icons, often resorting to an external search engine (e.g., Google) before successfully locating content on the mobile site.
- **Memorability**: A significant positive correlation was found between visit frequency and SUS scores in both Questionnaire 1 ($\rho = .253$, $p = .020$; Table 6.4) and Questionnaire 2 ($\rho = .298$, $p = .023$; Table 6.12). In the open-ended responses from Questionnaire 1 some respondents explicitly asked for "accessibility to favorite teams, for example by adapting the

menu based on the user's recently viewed content" and "direct access to preferred teams." In other words, interviewees struggled to return to a club page they'd just viewed, having to re-navigate or re-search each time. They recommended adding a "Recent" or "Favorites" shortcut in the main menu to surface their most-visited teams immediately.

- **Error Prevention & Recovery**: Hotjar analyses from Questionnaire 1 reported nine U-turns, 8 page refresh events and no rage clicks were reported (as seen in table 6.8. Interviews recounted silent comment-posting errors leading to abandonment (Subsection 6.3.3).

- **Satisfaction**: The mean of SUS score from Questionnaire 1 was 81.35 (Subsection 6.2.3) and 76.03 (SD = 16.70) for Questionnaire 2 (Subsection 6.4.3), both classed as "Good." In the open-ended responses from the questionnaires both mentioned ad frustration (Questionnaire 1: 10 users; Questionnaire 2: 7 users), with one respondent expressing willingness to pay for an ad-free experience. Interviewees praised the site's aesthetic clarity yet criticized intrusive banners, exemplified by the comment:

  > "We were opening a news story or something, and the advertising pop-up would open, and we were stuck there, unable to move either way."

7.2 Identified Usability Issues by UX Nielsen's Usability Attributes

This section presents the key usability challenges identified across the five usability attributes defined by Nielsen: learnability, efficiency, memorability, errors, and satisfaction. Figure 7.1 provides a visual summary of the most prominent issues reported by users and observed in behavioral data. These issues are further detailed in the subsections that follow, and are described in terms of frequency, context, and perceived impact, supported by quantitative and qualitative evidence.

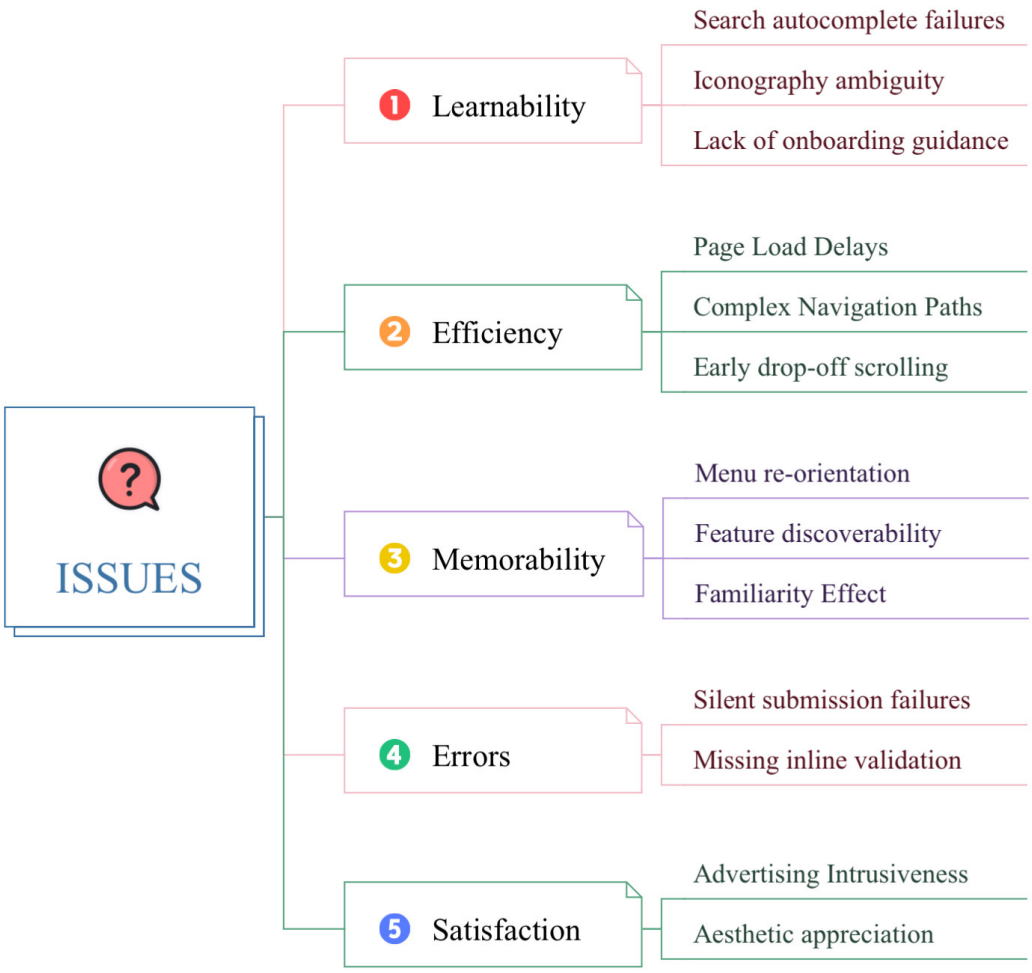


Figure 7.1: Overview of usability issues grouped by UX dimension (based on Nielsen's framework).

7.2.1 Learnability

Users face initial learning barriers in terminology and interface predictability:

- **Search Autocomplete Failures**: Participants in Q1 and Q2 reported that the autocomplete suggestion menu frequently returned irrelevant or no results when typing team, player, or competition names. In interviews, several users admitted bypassing the native search entirely (consulting Google first) before returning to ZeroZero, violating Nielsen, 1993 *Match Between System and the Real World* heuristic.
- **Iconography Ambiguity**: Multiple respondents misinterpreted the "hamburger" menu and other icons, delaying task initiation. Alonso-Virgós et al., 2020 show that misuse of standard icons deeply undermines learnability by stripping users of familiar signifiers .

- **Lack of Onboarding Guidance**: Across both questionnaires, users requested contextual tips or brief tutorials. Nielsen's heuristic on *Help and Documentation* holds that just-in-time assistance is essential for first-time user success (Nielsen, 1993).

7.2.2 Efficiency

Performance bottlenecks and navigation complexity impede task flow:

- **Page Load Delays**: In open-ended responses from questionnaire 2, 11 respondents cited slow page loads (sometimes exceeding 30 s on match days) as their top frustration (Fig. 6.35). According to Nielsen, delays over 1 s break users' flow, while those over 10 s often lead to abandonment[1].
- **Complex Navigation Paths**: Some users noted multi-tap sequences to reach core features like "My Teams," forcing repeated drill-downs through nested menus. This violates Nielsen's "Flexibility and Efficiency of Use" heuristic by burdening users with unnecessary steps (Nielsen, 1993).
- **Early Drop-off in Scrolling**: Hotjar scroll maps showed that fewer than 25% of users reached the footer, indicating that roughly three-quarters abandoned before the end of the page (Sub-subsection 6.3.2.2); visual engagement studies confirm that content below the fold receives substantially less attention (Moran et al., 0000).

7.2.3 Memorability

Returning users encounter reorientation challenges after inactivity:

- **Menu Re-Orientation**: Open responses from Questionnaire 1 explicitly called for "accessibility to favorite teams, for example by adapting the menu based on the user's recently viewed content" and "direct access to preferred teams." Interviewees confirmed they repeatedly re-navigated or re-searched club pages rather than returning quickly, indicating a need for a "Recent" or "Favorites" shortcut to aid recognition and reduce recall burdens (Nielsen, 1993).
- **Feature Discoverability**: In Questionnaire 1, 14 respondents called out interface and layout problems—outdated or cluttered screens made it hard to find match details or the "Match Highlights" section. In Questionnaire 2, 2 users described the internal search as imprecise for locating match summaries or statistics. Interviewees confirmed that buried options and inconsistent layouts forced extra navigation steps, underscoring the need for progressive disclosure to surface secondary features more intuitively (Nielsen, 1993).

[1] Jakob Nielsen, "Website Response Times," Nielsen Norman Group, June 20, 2010, https://www.nngroup.com/articles/website-response-times/.

- **Familiarity Effect**: Significant positive correlations between visit frequency and SUS in Q1 ($\rho = .253$, $p = .020$) and Q2 ($\rho = .298$, $p = .023$) confirm that repeated exposure enhances retention of interface patterns (Sauro et al., 2012).

7.2.4 Error Prevention and Recovery

Deficiencies in feedback and validation hamper error handling:

- **Silent Submission Failures**: Both Q1 and Q2 users, including interviewees, reported comment-posting attempts that yielded no feedback, with form submissions requiring multiple retries. This directly violates the heuristic that systems must help users recognize, diagnose, and recover from errors by providing clear, actionable feedback and guidance (Nielsen, 1993).
- **U-Turns vs. Rage Clicks**: Hotjar recorded nine U-turns and zero rage clicks, indicating users backtracked rather than clicking repeatedly. While Nielsen's heuristic on error prevention is partly met, the lack of recovery support remains problematic (Nielsen, 1993).
- **Missing Inline Validation**: Real-time field validation was absent, forcing repeated input attempts. Nielsen emphasizes prevention of errors before they occur through inline checks (Nielsen, 1993).

7.2.5 Satisfaction

Affective reactions shaped by design elements:

- **Overall Usability (SUS)**: Q1 mean SUS = 81.35 (SD = 17.36), Q2 mean = 76.03 (SD = 16.70), both in the "Good-to-Excellent" range (Brooke, 1996). The lower Q2 score suggests non-registered users perceive slightly lower satisfaction.
- **Advertising Intrusiveness**: Users cited intrusive ads as their main frustration; interviewees similarly criticized banner placement. Excessive monetization elements have been shown to erode user satisfaction (Nielsen; Budiu, 2012).
- **Aesthetic Appreciation**: Despite varied critiques, users praised the platform's typography and visual consistency, underscoring how aesthetic harmony can enhance perceived usability and satisfaction, as Norman demonstrates in his book Emotional Design (Norman, 2005).

7.3 Relation to UX Principles and Guidelines

Each identified issue was further analyzed through the lens of established UX frameworks, including Nielsen's Heuristics [2] and WCAG accessibility principles [3]. This comparison reinforces the

[2] 10 Usability Heuristics for User Interface Design at https://www.nngroup.com/articles/ten-usability-heuristics/

[3] Web Content Accessibility Guidelines (WCAG) 2.1 at https://www.w3.org/TR/WCAG21/

empirical findings and emphasizes where the mobile version of ZeroZero aligns or diverges from industry standards.

7.3.1 Learnability

The mobile version of the ZeroZero website exhibits several learnability challenges when assessed against established UX heuristics and accessibility criteria. First, the complete absence of onboarding cues or progress indicators directly conflicts with Nielsen's *Visibility of System Status* (Heuristic #1), leaving users uncertain about their current position within the site or the outcome of recent interactions. Second, without contextual tooltips or inline instructions, visitors (particularly those with cognitive or learning disabilities) cannot rely on WCAG's 2.1 *Input Assistance* requirement (Success Criterion 3.3.2), hindering accurate interpretation of controls. Finally, the interface fails to introduce functionality incrementally, running counter to the principle of Progressive Disclosure in Material Design, and thereby overwhelming novices with an undifferentiated array of options.

7.3.2 Efficiency

Efficiency suffers similarly when contrasted with core UX and accessibility standards. The site's deep, multi-tap navigation hierarchies undermine Nielsen's *User Control and Freedom* (Heuristic #3), since users cannot easily backtrack or jump to frequently visited pages. Moreover, inconsistent menu layouts and the absence of ARIA landmarks [4] violate WCAG's 2.1 *Navigable* guidelines (Success Criteria 2.4.1 – 2.4.10), making both keyboard and assistive-technology navigation cumbersome. Finally, page load times regularly exceed five seconds—well above the two-second threshold recommended by Material Design for responsive performance—thereby breaking users' sense of fluid interaction.

7.3.3 Memorability

Returning users face steep memorability hurdles because the design forces recall rather than facilitating recognition. Critical actions and controls are hidden behind shifting menu structures, violating Nielsen's *Recognition Rather Than Recall* (Heuristic #6) and imposing undue cognitive load. The inconsistent placement of functions such as "My Teams" further breaches WCAG's 1.2 *Predictable* criterion (Success Criteria 3.2.3 – 3.2.4), as users cannot anticipate where key features will appear. Additionally, variations in iconography and page layouts undermine Material Design's consistency principle, disrupting users' mental models and making reorientation difficult.

7.3.4 Error Prevention and Recovery

Error resilience is likewise inadequate. The lack of real-time input validation or confirmation dialogs runs counter to Nielsen's *Error Prevention* (Heuristic #5), allowing many mistakes to go

[4] ARIA landmarks act as "skip to" links, allowing users to jump directly to specific sections of the page, such as the main content, navigation, or header.

unchecked until form submission. When submissions fail silently, the site violates WCAG's 2.1 *Error Identification* requirement (Success Criterion 3.3.1), leaving users unsure whether an action succeeded or how to correct it. Furthermore, the absence of immediate visual or haptic feedback for failed operations conflicts with Material Design's *Immediate Feedback* principle.

7.3.5 Satisfaction

Finally, overall satisfaction is diminished by avoidable design missteps. Intrusive full-screen advertisements draw attention away from core content, infringing Nielsen's *Aesthetic and Minimalist Design* (Heuristic #8). Although the site's typography and color palette meet WCAG's minimum contrast requirements (Success Criterion 1.4.3), the presence of auto-playing audio advertisements without pause or mute controls violates WCAG's 2.1 *Audio Control* guideline (Success Criterion 1.4.2). Abrupt pop-ups and banners further disrupt focus, running counter to Material Design's recommendation to prioritize content and minimize distractions.

Chapter 8

Recommendations

The following recommendations have been formulated based on direct user feedback gathered during the mobile website evaluation, including analysis of open-ended questionnaire responses and insights from moderated interviews. This multi-method approach ensured that both quantitative metrics and qualitative user voices informed the strategic, tactical, and operational guidance presented herein.

A high-level overview of the most relevant recommendations, structured by the five core usability attributes is presented in Figure 8.1. These insights are derived from recurring user feedback themes and serve as a conceptual anchor for the detailed proposals that follow.

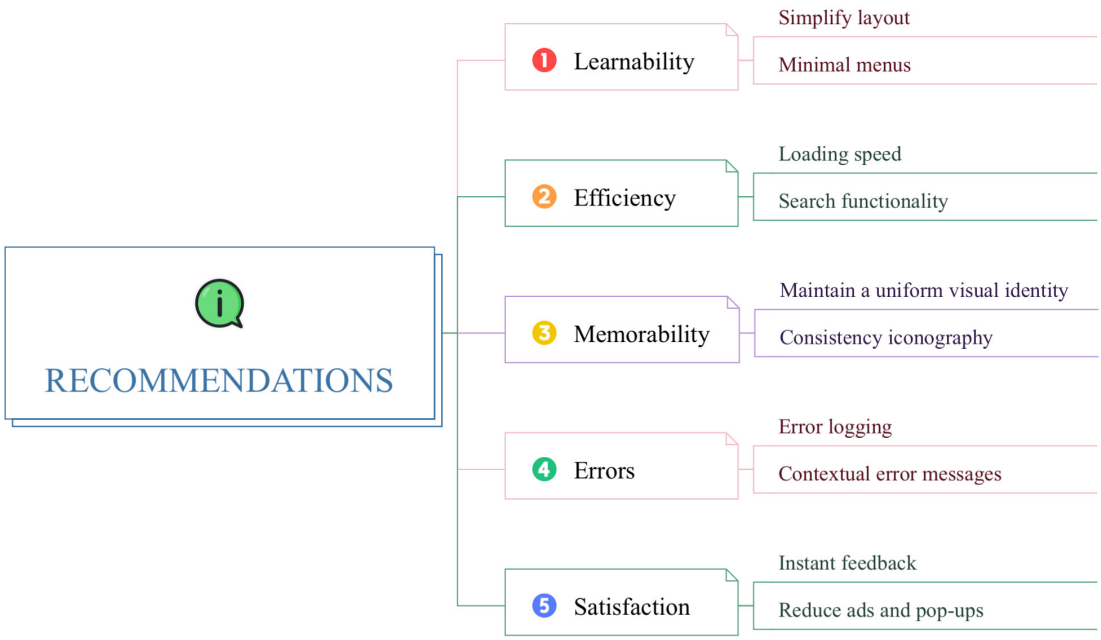


Figure 8.1: Summary of user-driven recommendations structured by usability attribute.

8.1 Strategic, Tactical and Operational Recommendations

For each identified issue, actionable recommendations were developed considering different decision-making levels: strategic, tactical, and operational. These recommendations aim to guide the ZeroZero team in prioritizing and implementing changes in a realistic and effective manner.

A condensed summary of these actions is presented in the Action Sheet in Appendix D. The proposed initiatives cover a twelve-month implementation period and address a total of six critical usability issues, each mapped to a responsible team member, estimated resources, and a time horizon.

8.1.1 Learnability

8.1.1.1 Strategic Recommendations

To improve overall learnability, establish a UX governance framework that embeds user-centered design principles into every product decision. This should include periodic expert reviews of major feature changes and a centralized style guide to ensure consistent terminology, iconography and interaction patterns across the mobile platform. By aligning mobile UI conventions with the desktop site's established mental model, new and returning users will transfer prior knowledge more readily, reducing cognitive load and accelerating task initiation.

Invest in a lightweight, contextual onboarding experience for first-time or infrequent users. Deploy interactive tooltips and short, skippable tutorials that highlight core navigation elements (menu, search, key pages) on initial access. Over time, refine these materials based on mobile web analytics to focus on areas where users most often hesitate or abandon tasks, ensuring that learnability initiatives are both data-driven and adaptable.

8.1.1.2 Tactical Recommendations

Standardize key screens (home, league overview, player profile) with consistent layouts, labels and affordances. For example, always position primary navigation controls (back, menu, search) in the same location and use the same visual weight. This consistency supports rapid visual recognition and muscle memory, reducing the time required to relearn interfaces after app updates.

Simplify on-screen language by using concise, action-oriented labels (e.g., "View Table", "Browse Matches") instead of longer descriptive text. Pair each label with a universally recognized icon to reinforce meaning through dual coding (text + image). Conduct quick validation tests (e.g., five-second glance test) to confirm that users correctly interpret each control without reading extensive copy.

8.1.1.3 Operational Recommendations

Implement progressive disclosure to surface only the most essential options initially, with secondary functions revealed upon demand (e.g., "More Options" or swipe-up panels). This reduces initial complexity, guiding users to complete core tasks before exploring advanced features.

Leverage analytics events (e.g., Hotjar click maps) to identify where users pause or U-turn, and deploy A/B tests to trial alternative labeling or control placements. Roll out changes to small user segments first, and iterate based on task success rate and SUS feedback before full deployment.

8.1.2 Efficiency

8.1.2.1 Strategic Recommendations

Define key user journeys (e.g., finding live scores, checking standings, searching players) and map their end-to-end steps. Use these as primary success metrics in a centralized dashboard to align cross-functional teams on performance goals. Prioritize initiatives that reduce the number of taps or loading screens in these critical flows.

Adopt a mobile-first mindset for all new features, ensuring that every design is optimized for one-hand use and touch interactions. Embed performance budgets into design sprints (for CPU, memory and network calls) to guarantee that UI complexity never compromises speed.

8.1.2.2 Tactical Recommendations

Consolidate search and navigation into a single, omnipresent control: enhance the search bar to support quick filters (e.g., “Teams”, “Players”, “Matches”) and auto-suggestive predictive typing. This reduces context switches between menu navigation and search, shortening time to content.

Optimize paging and data loading by implementing infinite scroll or “load more” patterns only where it adds value (e.g., chronological news feed), and use placeholder skeleton screens for all heavy data views (player stats, historical tables). This gives users immediate visual feedback, decreasing perceived wait time.

8.1.2.3 Operational Recommendations

Implement client-side caching for frequently accessed data (e.g., today’s matches, top leagues) to allow near-instant load on repeat visits. Use stale-while-revalidate strategies so cached data appears immediately while background refreshes ensure up-to-date content.

Integrate granular performance monitoring (e.g., Real User Monitoring via analytics SDK) to capture per-screen load times and error rates. Configure automated alerts when key thresholds (e.g., 2 s load time) are breached, enabling rapid diagnosis and rollback of performance regressions.

8.1.3 Memorability

8.1.3.1 Strategic Recommendations

Create and enforce a mobile style guide that defines color palettes, typography scales, icon sets and interaction patterns for all common components (buttons, lists, cards). This living document

should be version-controlled and integrated into the dev pipeline, ensuring that any new view or feature adheres to the same standards, reinforcing a stable mental model across sessions.

Conduct quarterly "UX health checks" where a small panel of users returns to the app after a 2-week hiatus to complete standard tasks (e.g., find a team, view last match). Measure time-on-task and error rates, and benchmark against initial learnability metrics to detect regressions in memorability.

8.1.3.2 Tactical Recommendations

Introduce persistent "Favorites" functionality at the top of home and league pages, allowing users to pin up to three teams or competitions. This personalized shortcut reduces navigation depth for frequently visited content and reinforces memory by visually surfacing preferred items.

Additionally, provide a dedicated "Favorites" view (accessible via a persistent tab or menu entry) where all starred items are listed on a single screen, allowing the user to review, manage (remove/add) and navigate directly to each favorite.

Use consistent, distinctive icons for top-level actions (search, menu, back) and maintain their order across all screens. Avoid dynamic reordering or hiding of these controls, as fluctuations undermine users' ability to recall control locations.

8.1.3.3 Operational Recommendations

Enable local storage of user settings (e.g., dark mode, preferred league) so that preferences persist across sessions without additional sign-in friction. Confirm that settings survive the mobile updates and reinstalls to preserve the user's mental context.

Deploy push notifications (opt-in) for crucial content (e.g., match start, score updates), using clear, uniform formatting and channel categories (Live, Daily Digest). Well-timed notifications help reorient returning users and maintain awareness without requiring manual navigation.

8.1.4 Errors

8.1.4.1 Strategic Recommendations

Establish a proactive error management framework by integrating comprehensive client-side logging (JavaScript exceptions, network failures) and server-side monitoring. Ensure all captured errors funnel into a central tracking system (e.g., Sentry, DataDog) for trend analysis. Convene a bi-weekly review to triage and prioritize fixes based on frequency and user impact.

Define an "Error Budget" aligned with business goals (e.g., no more than 1% of sessions yield runtime errors) and incorporate it into the product roadmap. Use this budget to balance feature velocity against platform stability, ensuring a reliable baseline user experience.

8.1.4.2 Tactical Recommendations

Implement inline validation and context-aware error messages (e.g., "Unable to load standings—please check your connection or retry") rather than generic pop-up alerts. Provide users with immediate remediation options (Retry, Go Back, Report Issue).

Add an "Undo" affordance where destructive actions occur (e.g., unpinning a favorite, dismissing an in-app alert), allowing users to recover gracefully without restarting tasks or losing context.

8.1.4.3 Operational Recommendations

Roll out client-side feature flags for new or high-risk UI changes. Gradually enable features for small user cohorts, monitor error and performance metrics closely, then proceed to broader releases only once error rates remain within the defined budget.

Configure automated rollback thresholds in the CI/CD pipeline (e.g., if error rate spikes > 0.5% in 10 minutes post-release) to revert to the last known safe version without manual intervention.

8.1.5 Satisfaction

8.1.5.1 Strategic Recommendations

Foster a continuous feedback culture by embedding mobile-site NPS or sentiment prompts at key moments (e.g., post-game summary). Aggregate responses in a real-time dashboard to identify satisfaction trends and correlate them with feature releases or performance changes.

Establish a cross-functional "Voice of the User" council that meets monthly, reviewing feedback data (qualitative comments, questionnaire scores, support tickets) to inform product priorities and ensure that user delight remains central to decision making.

8.1.5.2 Tactical Recommendations

Introduce lightweight gamification elements (such as "Streak Badges" for following daily matches or "Trivia Challenges" with simple quizzes) to deepen engagement and reward exploration of underutilized features (e.g., historical stats, encyclopedia entries).

Enhance visual polish on high-impact screens (home, live match view) by adding subtle motion cues (e.g., header collapse, score update animations) and micro-interactions (e.g., haptic feedback on tap) to create a more delightful and memorable experience.

8.1.5.3 Operational Recommendations

Deploy a closed-loop feedback mechanism: when a user submits a bug report or suggestion, send an automated follow-up email once the issue is addressed, thanking them and outlining the fix. This transparency builds trust and demonstrates responsiveness.

Track post-release satisfaction shifts by mapping questionnaire responses to rollout cohorts. Use cohort analysis to verify that each deployment maintains or improves baseline satisfaction before proceeding with subsequent updates.

Chapter 9

Conclusions and future work

9.1 Meeting Objectives

The first objective was to conduct a comprehensive literature review and state-of-the-art to identify best practices in UX evaluation for mobile platforms. Through the SLR, 429 records were screened and 39 high-quality studies selected, revealing dominant themes such as heuristic evaluation, user-centered design, and advanced methods like eye-tracking and psychophysiological measures. Key gaps were also identified, including the lack of mobile-specific heuristic sets and comparative evidence across methods. These findings informed the selection of mixed UX evaluation techniques: questionnaires (SUS), focus groups, and behavioral analytics to address mobile specific challenges in data-intensive platforms (see Chapter 2 and Chapter 3).

The second objective was to develop and substantiate methodological options for assessing the user experience of mobile platforms. A mixed-methods framework was implemented, combining the System Usability Scale for quantitative benchmarking, session recordings and heatmaps via Hotjar for behavioral validation, and structured interviews for in-depth qualitative insights. This approach allowed triangulation of subjective perceptions, observed behavior, and contextual feedback, yielding a multifaceted evaluation strategy grounded in both academic rigour and practical feasibility (see Chapter 4 and Chapter 5).

The third objective was to identify usability gaps and propose targeted improvements based on empirical findings. Questionnaire data from 85 registered and 58 non-registered users yielded mean SUS scores of 81.4 and 78.2 respectively (both in the "Good to Excellent" range) but also revealed performance-related concerns (e.g., loading delays under mobile data vs. Wi-Fi, $\chi^2(2) = 7.37,\ p = .025$). Open-ended responses highlighted issues with interface consistency, intrusive advertising, search reliability, and difficulty accessing niche competitions. Hotjar analytics corroborated these findings, showing drop-off beyond mid-page scroll depths and sporadic U-turns but no rage clicks. Structured interviews further contextualized these patterns, with participants calling for cleaner layouts, enhanced personalization (e.g., pinned favourites), richer live updates, and incentives to engage in deeper feedback (see Chapter 6 and Chapter 7).

Collectively, this study contributes a validated mixed-methods evaluation framework for mobile UX in data-intensive contexts, and a prioritized set of recommendations, ranging from strategic (redesign of navigation IA and ad placement) to operational (improved search algorithms, streamlined live score updates) tailored to ZeroZero's mobile platform (see Chapter 8).

9.2 Answering the Research Question

This study set out to identify the feature-related design decisions that most impact user experience during the transformation of a desktop website into a mobile version, using the five usability attributes defined by Nielsen (learnability, efficiency, memorability, errors, and satisfaction) as analytical lenses.

Findings indicate that user experience in this context is most influenced by design decisions that enhance simplicity, speed, visual coherence, and contextual feedback. In terms of **learnability**, participants consistently emphasized the importance of a simplified layout and minimal menu structure, which facilitate intuitive navigation for first-time or infrequent users. For **efficiency**, fast loading times and direct access to frequently used sections were highlighted as critical, especially under mobile data conditions.

Regarding **memorability**, maintaining a consistent visual language (including iconography and layout patterns) was found to support task recall and reduce the cognitive effort required in repeated use. The dimension of **errors** was less prominent in observed behavior but emerged in qualitative data as a need for clearer feedback when searches fail or actions produce no visible result. Finally, **satisfaction** was closely tied to the reduction of intrusive elements (such as excessive advertising or pop-ups) and the enhancement of live, up-to-date content.

In short, the most impactful design decisions are those that reduce friction, support user autonomy, and maintain consistency with users' prior experiences on desktop, while adapting to the constraints and expectations of mobile interaction. These conclusions are reflected in the full set of user-driven recommendations developed throughout the study, summarized in Figure 8.1.

9.3 Limitations

Several limitations should be acknowledged. First, the voluntary convenience sample (85 registered; 58 non-registered) and the requirement to access via a questionnaire-specific link introduced selection bias and may not fully represent all user segments. Second, Hotjar data were limited to the trial period and to sessions triggered by the questionnaire link, constraining the behavioral dataset. Third, the structured interviews (initially planned as focus groups) were conducted with only two participants and one participant respectively, reducing the breadth of qualitative perspectives and dynamic group interactions. Fourth, performance measurements were based on self-reported perceptions rather than automated telemetry. Finally, the cross-sectional design captures user perceptions at a single point in time and may not reflect longitudinal changes in behaviour or the impact of future platform updates.

9.4 Future Work

These directions aim to deepen the understanding of mobile UX dynamics in data-intensive platforms and guide iterative improvements, while also informing long-term strategies for user engagement and platform evolution.

Enhanced Qualitative Methods

While diary studies and methods such as contextual inquiry were not feasible in the present project due to time constraints, they remain valuable avenues for future research. Diary studies, in particular, could yield longitudinal insights into user engagement and satisfaction over time, complementing the cross-sectional data collected in this study.

During the research process, organizing participatory activities such as focus groups proved to be challenging—particularly due to reliance on email-based recruitment. User engagement was limited, likely influenced by the absence of immediate incentives. As a result, structured interviews were selected as a more feasible alternative. For future studies requiring greater participant commitment (e.g., diary studies or focus groups), it is recommended that ZOS consider tangible incentives such as merchandise giveaways (e.g., raffles for branded jerseys), podcast guest spots, or exclusive experiences (e.g., attending a match with a ZOS reporter). These strategies could enhance recruitment and retention, while also fostering stronger user-platform relationships.

Quantitative UX Monitoring

In terms of quantitative enhancement, future work could include controlled experiments and automated monitoring. A/B testing of interface iterations (e.g., revised navigation structures or ad placement schemes) would allow for direct comparison of design alternatives based on key performance metrics (e.g., task success rate, time on task, SUS scores). Integration of real user monitoring (RUM) tools would further strengthen evaluation by capturing objective telemetry data—complementing subjective reports and enabling a tighter link between technical performance and perceived usability.

Advanced Instrumentation and Tools

Another promising direction involves the use of eye-tracking technology to collect fine-grained data on visual attention and interaction flows. By identifying which interface elements succeed or fail in capturing user focus, designers can better optimize layout hierarchy and visual cues. When paired with think-aloud protocols, eye tracking also offers rich cognitive insights into user reasoning during task execution.

In addition, future studies could incorporate standardized UX assessment tools such as the *User Experience Questionnaire (UEQ)* Alvarez et al., 2024 and the *AttrakDiff* scale[1]. These instruments enable multi-dimensional evaluation of UX, capturing both pragmatic aspects (e.g.,

[1] AttrakDiff: https://www.attrakdiff.de/sience-en.html

usability, efficiency) and hedonic qualities (e.g., stimulation, identification). Their inclusion would allow more nuanced cross-segment comparisons and benchmarking against industry standards.

Prototyping a Native Mobile Application

One of the most recurrent suggestions across open-ended questionnaire responses and interviews was the development of a dedicated native mobile application. Users explicitly associated a native app with faster performance under mobile data conditions, offline accessibility, and real-time push notifications. Additional expectations included enhanced personalization and reduced dependency on browser environments.

To address these needs, a native app prototype could be designed as a Minimum Viable Product (MVP) focused on three core features:

1. Customizable "Favorites" for teams and competitions;
2. Push notifications for live match events;
3. Offline caching of matchday data.

These features align directly with user priorities such as immediacy, relevance, and reliability.

From a technical standpoint, cross-platform frameworks like Flutter or React Native could be considered to accelerate development and simplify maintenance. Given ZeroZero's robust backend infrastructure, a hybrid model reusing existing APIs (e.g., for live scores, standings) would reduce implementation overhead.

A phased roadmap might include:

- **Phase 1 (0–3 months):** MVP design[2] and technical scoping;
- **Phase 2 (4–6 months):** Development of core features and closed beta testing;
- **Phase 3 (7–9 months):** Public launch and integration of behavioral analytics.

Such a prototype would not only respond to current user demands, but also serve as a testbed for advanced UX patterns (e.g., gesture-based navigation, native transitions), while enabling long-term tracking of user engagement and retention.

Collectively, these future investigations build on the current study's findings and methodological groundwork, offering a foundation for continuous optimization of ZeroZero's mobile user experience.

[2]Minimum Viable Product (MVP): a version of a product with just enough features to satisfy early users and provide feedback for future development.

References

ABULFARAJ, A.; STEELE, A., 2020a. Coherent heuristic evaluation (CoHE): Toward increasing the effectiveness of heuristic evaluation for novice evaluators. In: *Proceedings of the 9th International Conference on Design, User Experience, and Usability*. Vol. 12200 LNCS, pp. 3–20. Available from DOI: 10.1007/978-3-030-49713-2_1.

ABULFARAJ, A.; STEELE, A., 2020b. Detailed Usability Heuristics: A Breakdown of Usability Heuristics to Enhance Comprehension for Novice Evaluators. In: *Proceedings of the 22nd International Conference on HCI International 2020 – Late Breaking Papers*. Vol. 12423 LNCS, pp. 3–18. Available from DOI: 10.1007/978-3-030-60114-0_1.

ADAMS, Anne; HART, Jennefer; IACOVIDES, Ioanna; BEAVERS, Sian; OLIVEIRA, Manuel; MAGROUDI, Maria, 2019. Co-created evaluation: Identifying how games support police learning. *International Journal of Human-Computer Studies*. Vol. 132, pp. 34–44. ISSN 1071-5819. Available from DOI: 10.1016/j.ijhcs.2019.03.009.

ALONSO-VIRGÓS, Lucía; ESPADA, Jordán Pascual; THOMASCHEWSKI, Jörg; CRESPO, Rubén González, 2020. Test usability guidelines and follow conventions. Useful recommendations from web developers. *Computer Standards & Interfaces*. Vol. 70, p. 103423. ISSN 0920-5489. Available from DOI: 10.1016/j.csi.2020.103423.

ALVAREZ, David Anton Corpuz; MONTELOYOLA, Charles Ian Sarmiento; OGAYA, Lex Anilov Tamanio; PASARON, King Ralph Daeniel Ebol; PURIFICACION, Patrick Miguel Garcia; SERRANO, Elcid A., 2024. Development of a Community-Based Disaster Awareness Web-based Application for Human-Computer Interaction: Pulse PH. In: *Proceedings of the 2024 9th International Conference on Intelligent Information Technology*. New York, NY, USA: Association for Computing Machinery, pp. 512–516. ICIIT '24. ISBN 9798400716713. Available from DOI: 10.1145/3654522.3654599. event-place: Ho Chi Minh City, Vietnam.

ANGULO, Cecilio; CHACÓN, Alejandro; PONSA, Pere, 2024. Chapter 4 - Workspace metrics and evaluation. In: ANGULO, Cecilio; CHACÓN, Alejandro; PONSA, Pere (eds.). *Cognitive Assistant Supported Human-Robot Collaboration*. Academic Press, pp. 77–105. ISBN 978-0-443-22135-4. Available from DOI: 10.1016/B978-0-44-322135-4.00013-4.

ASEMI, Adeleh; ASEMI, Asefeh, 2022. A judgment-based model for usability evaluating of interactive systems using fuzzy Multi Factors Evaluation (MFE). *Applied Soft Computing*. Vol. 117, p. 108411. ISSN 1568-4946. Available from DOI: 10.1016/j.asoc.2022.108411.

BACIKOVÁ, M; PORUBÄN, J; SULÍR, M; CHODAREV, S; STEINGARTNER, W; MADEJA, M, 2021. Domain Usability Evaluation. *ELECTRONICS*. Vol. 10, no. 16. ISSN 2079-9292. Available from DOI: 10.3390/electronics10161963.

BITKINA, Olga Vl.; KIM, Hyun K.; PARK, Jaehyun, 2020. Usability and user experience of medical devices: An overview of the current state, analysis methodologies, and future challenges. *International Journal of Industrial Ergonomics*. Vol. 76, p. 102932. ISSN 0169-8141. Available from DOI: 10.1016/j.ergon.2020.102932.

BOKINGKITO, Paul B.; CAPARIDA, Lomesindo T., 2022. Usability evaluation of a real-time water quality monitoring mobile application. *Sixth Information Systems International Conference (ISICO 2021)*. Vol. 197, pp. 642–649. ISSN 1877-0509. Available from DOI: 10.1016/j.procs.2021.12.185.

BROOKE, John, 1996. *SUS: A quick and dirty usability scale* [Usability Evaluation in Industry, edited by P. W. Jordan, B. Thomas, B. A. Weerdmeester, and A. L. McClelland, London: Taylor & Francis]. Available online at: https://hell.meiert.org/core/pdf/sus.pdf.

BURRELL, Gibson; MORGAN, Gareth, 2019. *Sociological Paradigms and Organisational Analysis: Elements of the Sociology of Corporate Life*. 2ª. London: Routledge. Routledge Revivals, no. 2. ISBN 978-1-4724-6247-3. Place: Gibson Burrell, Gareth Morgan.

CAJANDER, Åsa; LARUSDOTTIR, Marta; GEISER, Johannes L., 2022. UX professionals' learning and usage of UX methods in agile. *Information and Software Technology*. Vol. 151, p. 107005. ISSN 0950-5849. Available from DOI: 10.1016/j.infsof.2022.107005.

CAMPOS, Thiago Prado De; DAMASCENO, Eduardo Filgueiras; VALENTIM, Natasha Malveira Costa, 2024. Usability and User Experience Evaluation of Touchable Holographic Solutions: A Systematic Mapping Study. In: *Proceedings of the XXII Brazilian Symposium on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. IHC '23. ISBN 9798400717154. Available from DOI: 10.1145/3638067.3638071. event-place: Maceió, Brazil.

CAPELETI, Bruna Santana; FERREIRA, João Pedro Batista; DOMINGUETE, Gustavo Lopes; PEREIRA, Marluce Rodrigues; FREIRE, André Pimenta, 2020. Heuristic evaluation and user tests of wearable mobile health monitoring applications: what results do different methods yield? In: *Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. IHC '20. ISBN 978-1-4503-8172-7. Available from DOI: 10.1145/3424953.3426493. event-place: Diamantina, Brazil.

CHENG, Shiwei; PING, Qiufeng; WANG, Jialing; CHEN, Yijian, 2022. EasyGaze: Hybrid eye tracking approach for handheld mobile devices. *Special Issue on Intelligent Interaction in Mixed Reality*. Vol. 4, no. 2, pp. 173–188. ISSN 2096-5796. Available from DOI: 10.1016/j.vrih.2021.10.003.

FERNÁNDEZ, Jon; MACÍAS, José A., 2021. Heuristic-Based Usability Evaluation Support: A Systematic Literature Review and Comparative Study. In: *Proceedings of the XXI International Conference on Human Computer Interaction*. New York, NY, USA: Association for Computing Machinery. Interacción '21. ISBN 978-1-4503-7597-9. Available from DOI: 10.1145/3471391.3471395. event-place: Málaga, Spain.

FIELD, Andy, 2018. *Discovering Statistics Using IBM SPSS Statistics*. 5th. London: SAGE Publications.

GARRETT, Jesse James, 2011. *The elements of user experience: user-centered design for the Web and beyond*. 2nd ed. Berkeley, CA: New Riders. Voices that matter. ISBN 978-0-321-68368-7. OCLC: ocn503049598.

GEORGSSON, Mattias; STAGGERS, Nancy; ÅRSAND, Eirik; KUSHNIRUK, Andre, 2019. Employing a user-centered cognitive walkthrough to evaluate a mHealth diabetes self-management application: A case study and beginning method validation. *Journal of Biomedical Informatics*. Vol. 91, p. 103110. ISSN 1532-0464. Available from DOI: 10.1016/j.jbi.2019.103110.

HOLZINGER, Andreas, 2005. Usability engineering methods for software developers. *Commun. ACM*. Vol. 48, no. 1, pp. 71–74. ISSN 0001-0782. Available from DOI: 10.1145/1039539.1039541. Place: New York, NY, USA Publisher: Association for Computing Machinery.

ISO, 2018. *Ergonomics of human-system interaction – Part 11: Usability: Definitions and concepts*. International Organization for Standardization.

JAIN, Paras; GUPTA, Vishan Kumar; TIWARI, Harish; SHUKLA, Anurag; PANDEY, Priyank; GUPTA, Ashutosh, 2023. Human-Computer Interaction: A Systematic Review. In: *2023 International Conference on Advanced Computing & Communication Technologies (ICACCTech)*. Los Alamitos, CA, USA: IEEE Computer Society, pp. 31–36. Available from DOI: 10.1109/ICACCTech61146.2023.00015.

KHAIRAT, Saif; COLEMAN, Cameron; NEWLIN, Thomas; RAND, Victoria; OTTMAR, Paige; BICE, Thomas; CARSON, Shannon S., 2019. A mixed-methods evaluation framework for electronic health records usability studies. *Journal of Biomedical Informatics*. Vol. 94, p. 103175. ISSN 1532-0464. Available from DOI: 10.1016/j.jbi.2019.103175.

KOHAVI, Ron; LONGBOTHAM, Roger; SOMMERFIELD, Dan; HENNE, Randal, 2009. Controlled experiments on the web: Survey and practical guide. *Data Mining and Knowledge Discovery*. Vol. 18, pp. 140–181. Available from DOI: 10.1007/s10618-008-0114-1.

KRAWIEC, Łukasz; DUDYCZ, Helena, 2020. A comparison of heuristics applied for studying the usability of websites. *Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 24th International Conference KES2020*. Vol. 176, pp. 3571–3580. ISSN 1877-0509. Available from DOI: 10.1016/j.procs.2020.09.029.

KUHAR, Maja; MERČUN, Tanja, 2022. Exploring user experience in digital libraries through questionnaire and eye-tracking data. *Library & Information Science Research*. Vol. 44, no. 3, p. 101175. ISSN 0740-8188. Available from DOI: 10.1016/j.lisr.2022.101175.

KUIJK, Jasper van; DAALHUIZEN, Jaap; CHRISTIAANS, Henri, 2019. Drivers of usability in product design practice: Induction of a framework through a case study of three product development projects. *Design Studies*. Vol. 60, pp. 139–179. ISSN 0142-694X. Available from DOI: 10.1016/j.destud.2018.06.002.

MAGABLEH, Aws A.; TAHAT, Khalid S.Al-, 2024. Psychometric evaluation of the Arabic standardized user experience percentile rank questionnaire (Arabic SUPR-Q). *Computers in Human Behavior Reports*. Vol. 13, p. 100376. ISSN 2451-9588. Available from DOI: 10.1016/j.chbr.2024.100376.

MARAMBA, Inocencio; CHATTERJEE, Arunangsu; NEWMAN, Craig, 2019. Methods of usability testing in the development of eHealth applications: A scoping review. *International Journal of Medical Informatics*. Vol. 126, pp. 95–104. ISSN 1386-5056. Available from DOI: 10.1016/j.ijmedinf.2019.03.018.

MARTINELLI, Suéllen; LOPES, Larissa; ZAINA, Luciana, 2024. UX Research practices related to Long-Term UX: A Systematic Literature Review. *Information and Software Technology*. Vol. 170, p. 107431. ISSN 0950-5849. Available from DOI: 10.1016/j.infsof.2024.107431.

MARTINS, Ana Isabel; ANDIAS, Rosa; AZEDO, Diana; BAPTISTA, Frederico; URSINE, Bárbara; SILVA, Anabela G.; ROCHA, Nelson Pacheco, 2024. Heuristic Evaluation of a Web-Based Application to Deliver a Home-Based Personalized Physical Exercise Program for Older Adults. *CENTERIS – International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2023*. Vol. 239, pp. 158–165. ISSN 1877-0509. Available from DOI: 10.1016/j.procs.2024.06.158.

MORAN, Kate; LIU, Feifei, 0000. *UX Metrics and ROI*. 5ª. Fremont, CA, USA: Nielsen Norman Group. Year unknown. Placeholder used for compatibility. Place: Fremont Publisher: Warm Springs Blvd.

NAKAI, Mana; MIYAKE, Shinji; KOBAYASHI, Daiji, 2024. Evaluation of Tourism Website User Experience Based on Psychophysiological Measures. In: *Proceedings of the 2024 13th International Conference on Software and Computer Applications*. New York, NY, USA: Association for Computing Machinery, pp. 146–151. ICSCA '24. ISBN 9798400708329. Available from DOI: 10.1145/3651781.3651803. event-place: Bali Island, Indonesia.

NAMOUN, Abdallah; ALREHAILI, Ahmed; NISA, Zaib Un; ALMOAMARI, Hani; TUFAIL, Ali, 2024. Predicting the usability of mobile applications using AI tools: the rise of large user interface models, opportunities, and challenges. *The 15th International Conference on Ambient Systems, Networks and Technologies Networks (ANT) / The 7th International Conference on Emerging Data and Industry 4.0 (EDI40), April 23-25, 2024, Hasselt University, Belgium*. Vol. 238, pp. 671–682. ISSN 1877-0509. Available from DOI: 10.1016/j.procs.2024.06.076.

NASIR, Muhammad; IKRAM, Naveed; JALIL, Zakia, 2022. Usability inspection: Novice crowd inspectors versus expert. *Journal of Systems and Software*. Vol. 183, p. 111122. ISSN 0164-1212. Available from DOI: 10.1016/j.jss.2021.111122.

NIELSEN, Jakob, 1993. *Usability engineering*. 1st ed. Boston: Academic Press. ISBN 0-12-518405-0.

NIELSEN, Jakob; BUDIU, Raluca, 2012. *Mobile usability*. Vol. 10. 1ª. Berkeley, CA: New Riders. ISBN 978-0-321-88448-0. ISBN 10: 0321884485; ISBN 13: 9780321884480.

NORMAN, Donald A., 2005. *Emotional Design: Why We Love (or Hate) Everyday Things*. New York, NY: Basic Books. ISBN 9780465004171.

NORMAN, Donald A., 2013. *The Design of Everyday Things: Revised and Expanded Edition*. New York, NY: Basic Books. ISBN 9780465050659.

NOWAKOWSKI, Michał, 2020. Analysis and Evaluation of Information Usefulness and User Experience For Content Presentation in Electronic Media. *Procedia Computer Science*. Vol. 176, pp. 3654–3664. ISSN 1877-0509. Available from DOI: https://doi.org/10.1016/j.procs.2020.09.021.

QUIÑONES, Daniela; RUSU, Cristian, 2019. Applying a methodology to develop user eXperience heuristics. *Computer Standards & Interfaces*. Vol. 66, p. 103345. ISSN 0920-5489. Available from DOI: 10.1016/j.csi.2019.04.004.

ROHRER, Christian, 2022. *A Guide To Using User-Experience Research Methods; Content by Poster Design by Kelley Gordon*. Nielsen Norman Group. Available also from: https://www.nngroup.com/articles/which-ux-research-methods/. Design by Kelley Gordon.

SAAVEDRA, M.-J.; RUSU, C.; QUIÑONES, D.; RONCAGLIOLO, S., 2019. A Set of Usability and User eXperience Heuristics for Social Networks. In: *Proceedings of the 11th International Conference on Social Computing and Social Media*. Vol. 11578 LNCS, pp. 128–139. Available from DOI: 10.1007/978-3-030-21902-4_10.

SAHUA, Joe; MOQUILLAZA, Arturo, 2020. A Systematic Review of Usability Evaluation Methods and Tools for ATM Interfaces. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 12202 LNCS, pp. 130–141. Available from DOI: 10.1007/978-3-030-49757-6_9. Type: Conference paper.

SALVADORINHO, Juliana; BASTOS, Tiago; TEIXEIRA, Leonor, 2024. A talent management platform for HR 4.0 challenges: usability evaluation based on testing and the TAM scale. *CENTERIS – International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2023*. Vol. 239, pp. 1328–1335. ISSN 1877-0509. Available from DOI: 10.1016/j.procs.2024.06.303.

SAURO, Jeff; LEWIS, James R., 2012. *Quantifying the User Experience: Practical Statistics for User Research*. Morgan Kaufmann.

SHIBUYA, Yuya; HAMM, Andrea; CERRATTO PARGMAN, Teresa, 2022. Mapping HCI research methods for studying social media interaction: A systematic literature review. *Computers in Human Behavior*. Vol. 129, p. 107131. ISSN 0747-5632. Available from DOI: 10.1016/j.chb.2021.107131.

VERAL, Roberto; MACÍAS, José A., 2019. Supporting user-perceived usability benchmarking through a developed quantitative metric. *International Journal of Human-Computer Studies*. Vol. 122, pp. 184–195. ISSN 1071-5819. Available from DOI: 10.1016/j.ijhcs.2018.09.012.

WEBER, T.; ZOITL, A.; HUBMANN, H., 2019. Usability of development tools: A CASE-Study. In: *Proceedings of the 22nd International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C)*, pp. 228–235. Available from DOI: 10.1109/MODELS-C.2019.00037.

WEICHBROTH, P., 2020. Usability of mobile applications: A systematic literature study. *IEEE Access*. Vol. 8, pp. 55563–55577. Available from DOI: 10.1109/ACCESS.2020.2981892.

YIN, Robert K., 2008. *Case Study Research: Design and Methods*. 4ª. Thousand Oaks, CA: SAGE Publications, Inc. Applied Social Research Methods. ISBN 978-1-4129-6099-1.

YUHUI, Wang; YIYANG, Zhang; YANYING, Chen, 2024. Using animation to develop a usability questionnaire: Scale development and psychological measurements. *International Journal of Human-Computer Studies*. Vol. 187, p. 103258. ISSN 1071-5819. Available from DOI: 10.1016/j.ijhcs.2024.103258.

ZAKI, Tarannum; ISLAM, Muhammad Nazrul, 2021. Neurological and physiological measures to evaluate the usability and user-experience (UX) of information systems: A systematic literature review. *Computer Science Review*. Vol. 40, p. 100375. ISSN 1574-0137. Available from DOI: 10.1016/j.cosrev.2021.100375.

ZHENG, Qingxiao; TANG, Yiliu; LIU, Yiren; LIU, Weizi; HUANG, Yun, 2022. UX Research on Conversational Human-AI Interaction: A Literature Review of the ACM Digital Library. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. CHI '22. ISBN 978-1-4503-9157-3. Available from DOI: 10.1145/3491102.3501855. event-place: New Orleans, LA, USA.

Appendix A

Appendix 1

Appendix 1 presents the systematic literature review documentation used during the research process. Both the original version in Portuguese and the translated version in English are included for transparency and accessibility.

A.1 Systematic Literature Review File

Access the literature review file: You can access it here: Literature Review File.

A.2 Explanation of the SLR Appendices

The appendices contain supplementary materials related to the systematic literature review conducted as part of this dissertation. These materials are organized into several sections, as described below:

A.2.1 Final Queries

This section lists the final search queries used across various academic databases, including Web of Science, IEEE, Scopus, ScienceDirect, and ACM. These queries were formulated to identify relevant literature focusing on the UX within the context of the mobile version of ZeroZero. The detailed structure of the queries ensures a comprehensive and systematic approach to literature retrieval.

A.2.2 Criteria

The *Criteria* section outlines the inclusion and exclusion criteria applied to filter the retrieved studies. These criteria were designed to ensure that only the most relevant and high-quality research was included in the review. They address aspects such as publication type, year, language, and specific relevance to the research topic.

A.2.3 Identification

This section describes the identification process, detailing how studies were selected from the initial pool of results. It includes an overview of the steps taken to refine the dataset, from an initial keyword-based search to the removal of duplicates and irrelevant results.

A.2.4 NºResults

The *NºResults* section provides a summary of the number of results retrieved using the search queries in each academic database. It reflects the effectiveness of the queries and the overall scope of the systematic search.

A.2.5 Queries in Specific Databases

Two subsections are dedicated to queries executed in specific databases:

- **QS5 (WoS, IEEE, Scopus, SD[1]):** This subsection details the search queries applied to Web of Science, IEEE Xplorer, Scopus, and Science Direct, highlighting their structure and scope.
- **ACM QS3:** This subsection focuses on the queries executed in the ACM Digital Library, providing insight into how they were adapted for this particular database.

A.2.6 Eligible Results

The final section, *Eligible Results*, summarizes the studies that were deemed eligible after the screening and refinement process. In addition to listing the eligible studies, this section also presents the QA conducted for each study. The QA evaluates the methodological rigor and relevance of the selected studies, ensuring the dataset used for further analysis is robust and reliable.

[1] SD: Science Direct

Appendix B

Appendix 2

B.1 Questionnaire Materials

Appendix 2 presents the materials related to the questionnaire, including the design process of the informed consent and the questionnaire itself.

B.1.1 Informed Consent File for the Questionnaire

The informed consent form that was shown to participants at the beginning of the questionnaire. This version was written in Portuguese, in accordance with the language of the majority of ZeroZero.pt users.

Participants were required to read the information before proceeding. By continuing with the questionnaire, they confirmed their understanding and agreement with the terms outlined. This passive consent model was clearly indicated at the end of the consent form, aligning with standard practices for anonymous academic surveys.

B.1.1.1 Portuguese Version

Access to the Portuguese version of the Informed Consent File: You can access it here: Informed Consent File (Portuguese).

B.1.1.2 English Translation

Access to the English translation of the Informed Consent File: You can access it here: Informed Consent File (English).

B.1.2 Questionnaire File

Access to the Questionnaire File: You can access it here: Questionnaire File.

B.1.3 Questionnaire Dashboard

B.1.3.1 Dashboard with Questionnaire 1 Results

Access to the Questionnaire 1 Dashboard: You can access it here: Questionnaire 1 Dahsboard.

B.1.3.2 Dashboard with Questionnaire 2 Results

Access to the Questionnaire 2 Dashboard: You can access it here: Questionnaire 2 Dahsboard.

Appendix C

Appendix 3

C.1 Structured Interview Materials (Adapted from Focus Group Plan)

This appendix includes the informed consent form, the interview script, and the anonymised transcript of the structured interview conducted as part of the study. Portuguese and English versions are provided where applicable.

C.1.1 Informed Consent File for Structured Interviews

This document detailed the study's objectives, participants' rights, confidentiality terms, and procedures for data collection and usage, ensuring that all participants gave informed and voluntary consent prior to participating in the structured interview.

Access to the informed consent: You can access it here: Informed Consent.

C.1.2 Structured Interview Script

The following script was used in both structured interviews. It is the same guide originally designed for the focus group and is presented below in both the original Portuguese and an English translation.

C.1.2.1 Portuguese Version

Access to the Portuguese version of the script: You can access it here: Portuguese Script.

C.1.2.2 English Translation

Access to the English version of the script: You can access it here: English Script.

C.1.3 Structured Interview Transcript

The original transcripts were produced in Portuguese to capture participants' responses in their native language. These transcripts were later translated into English, with careful attention to

preserving the original meaning and tone of the statements. Minor adjustments were made for clarity and readability, while maintaining fidelity to the participants' perspectives.

C.1.3.1 Interview 1

Portuguese Version

Access to the Portuguese version of the transcript: You can access it here: 1st Structured Interview Transcript (Portuguese).

English Translation

Access to the English translation of the transcript: You can access it here: 1st Structured Interview Transcript (English).

C.1.3.2 Interview 2

Portuguese Version

Access to the Portuguese version of the second transcript: You can access it here: 2nd Structured Interview Transcript (Portuguese).

English Translation

Access to the English translation of the second transcript: You can access it here: 2nd Structured Interview Transcript (English).

Appendix D

Appendix 4

D.1 Action Sheet

Access to the Action Sheet: You can access it here: Action Sheet.

A formal organizational chart is not available, after consultation with company thesis advisor, departments have been structured as follows:

- **Helpdesk**: Provides user support and resolves technical issues reported.
- **Development**: Handles coding, integration, and performance optimization tasks; includes UX specialists for mobile layout prototyping and user-testing refinements.
- **Commercial Relations**: Manages partnerships, client management, sponsorships and advertising integrations.
- **Journalism**: Produces and curates sports news and match reports.
- **Content**: Manages multimedia content production (audio, video), ensures editorial consistency.