Usability Evaluation of a Mobile Application for Foot Health Monitoring of Smart Insoles: A Mixed Methods Study

Stefan Resch
University of Cadiz
Cadiz, Spain
stefan.resch@alum.uca.es
Frankfurt UAS
Frankfurt am Main, Germany

Lucian Zimmermann
Frankfurt UAS
Frankfurt am Main, Germany
lucian.zimmermann@stud.fra-uas.de

Alexander Kunkel
Frankfurt UAS
Frankfurt am Main, Germany
alexander.kunkel2@stud.fra-uas.de

Dennis Gerstung
Frankfurt UAS
Frankfurt am Main, Germany
dennis.gerstung@stud.fra-uas.de

Valentin Schwind
Frankfurt UAS
Frankfurt am Main, Germany
valentin.schwind@fb2.fra-uas.de

Diana Völz
Frankfurt UAS
Frankfurt am Main, Germany
voelz@fb2.fra-uas.de

Daniel Sanchez-Morillo
University of Cadiz
Cadiz, Spain
daniel.morillo@uca.es

ABSTRACT

Increasing research on intelligent wearable technologies, such as smart insoles for mobile health (mHealth) monitoring, brings new challenges for user-centered design, particularly in data visualization. Research shows various developments in mobile apps for monitoring foot health with smart insoles, while emerging trends like chatbot interactions and improved analytic visualizations offer new opportunities to enhance user experience. However, the usability of various health data visualizations remains unvalidated. A mixed-method experimental user study with 30 participants was conducted to assess the usability of three prototype mHealth applications for smart insoles: *Analytical*, *Basic*, and *Chatbot* visualizations. Quantitative results showed that *Basic* visualization achieved the highest usability, followed by *Analytical*, and finally *Chatbot*. Qualitative feedback supported these findings but also highlighted the potential of *Chatbot* interactions to enhance data understanding in mHealth apps. We discuss implications for future mHealth applications to monitor foot health and propose design recommendations to improve usability in chatbot interactions.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; **Graphical user interfaces**; **Smartphones**.

KEYWORDS

Smart Insoles; Health Monitoring; mHealth; Mobile Application; Usability

MOBILEHCI Adjunct '24, September 30–October 03, 2024, Melbourne, VIC, Australia

ACM ISBN 979-8-4007-0506-9/24/09
https://doi.org/10.1145/3640471.3680233

ACM Reference Format:
Stefan Resch, Lucian Zimmermann, Alexander Kunkel, Dennis Gerstung, Valentin Schwind, Diana Völz, and Daniel Sanchez-Morillo. 2024. Usability Evaluation of a Mobile Application for Foot Health Monitoring of Smart Insoles: A Mixed Methods Study. In *26th International Conference on Mobile Human-Computer Interaction (MOBILEHCI Adjunct '24), September 30–October 03, 2024, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3640471.3680233

1 INTRODUCTION

The ongoing development of intelligent wearable electronic devices, such as smart insoles, enables new opportunities for mobile health (mHealth) monitoring of patients with foot disorders. Smart insoles allow continuous monitoring of plantar foot pressure distribution and associated health data through integrated sensor technologies. Related work presents several concepts for smart foot wearables, focusing on the development of the wearable system and proposing corresponding apps [5, 15, 31]. For example, Saidani et al. [27] proposed a smart insole system for detecting gait abnormalities, preventing foot ulcers, improving posture and balance, and enhancing physical activity and health. The associated app allows patients to monitor daily activities by visualizing their gait cycle and the pressure distribution. However, these technologies also need a user-friendly interface and comprehensible data visualization in the mobile app for effective user interaction. As an example, Altenhoff et al. [2] conducted a mixed-method field study and investigated the usability of two wearable fitness trackers and their apps through the think-aloud usability test. The results indicated that the design of the app is crucial for the user acceptance and satisfaction of the entire wearable device. Despite the increasing research on the technical aspects of smart insoles, there is still a need to evaluate the usability of foot mHealth apps to ensure successful usage and high user acceptance. This is particularly important for patients with various foot conditions who wear orthopaedic aids on a daily basis, as they often rely on expert feedback and require personalized instructions to manage their conditions effectively. The use of a smartphone app facilitates the tracking of their foot health, provides personalized recommendations, and allows them greater independence in managing their conditions. Therefore, further research is needed to validate and test mobile apps for smart insoles to ensure they meet the needs of all users and provide accessible health data. This leads to the research question (RQ) of how the visualization of foot health data influences usability in a corresponding mobile app for smart insoles. These new insights should help improve user interaction with mHealth apps for smart insoles and provide future directions to overcome social challenges and avoid incomprehensible user interface (UI) design in associated mobile apps.

In this paper, we present the results of a user study conducted with 30 participants to investigate the effect of health data visualization on usability in mobile apps for foot health monitoring. The summative evaluation of three prototypes shows that interaction with a chatbot leads to lower usability compared to conventional or statistical data visualization. Based on these findings, we discuss future design recommendations for mobile app visualizations of foot health data.

2 RELATED WORK

In this section, we analyze relevant research areas related to our study, focusing on the usability of mobile apps, data visualization techniques for health monitoring, and the impact of chatbots on user interaction. Our work aligns with these areas as we aim to develop and evaluate a mobile app for smart insoles with an effective health data visualization.

2.1 Usability in Mobile Applications

In the field of Human-Computer Interaction (HCI), the aspect of usability in mobile interactions has been extensively researched [17, 24, 33, 35]. Particularly for mHealth apps, specific usability design guidelines have been proposed [13, 22, 26]. For example, Nasr et al. [22] developed and validated a comprehensive usability guideline with 16 key aspects and 154 checklist items to ensure user-friendly design for mHealth apps. Besides applying a user-centered design approach during development, it is crucial to evaluate the usability of these applications. Currently, various standardized questionnaires (e.g., the System Usability Scale (SUS) [11], Post-Study System Usability Questionnaire (PSSUQ) [14], and mHealth Apps Usability Questionnaire (MAUQ) [37]) are used for mHealth usability assessment, but there is no general consensus on the best method [21]. In a systematic literature review, Muro-Culebras et al. [21] identified eight relevant questionnaires for mHealth evaluation, highlighting the strengths and weaknesses of each.

The SUS is a quick and effective assessment tool that is widely used due to its comparability with other studies and its proven effectiveness, even though it was not specifically developed for mHealth apps [9, 21]. Nevertheless, SUS has been applied to evaluate the usability of both smart wearable devices [20, 29] and mHealth apps [21, 34]. For example, Wiyanti et al. [34] conducted a mixed-method study to evaluate the usability and usefulness of a mHealth app using the SUS questionnaire and qualitative think-aloud sessions. The results indicated a "good" level of user satisfaction, according to the adjective scale proposed by Bangor et al. [3], providing a benchmark for usability in similar health apps. In addition to the SUS, the think-aloud method by Lewis [6] is frequently used in usability testing of mobile apps to gain valuable insights into user behavior [19, 23]. This method complements the quantitative data obtained from questionnaires by providing a deeper understanding of how users interact with the app.

2.2 Data Visualization for Health Monitoring

Comprehensible health data visualization is challenging and increasingly important for ensuring user-friendly mobile app design [18]. Particularly, in the field of clinical medicine, information visualization is not as advanced as in other scientific disciplines, indicating a need for further evaluation of health data by patients [10]. Therefore, recent research has focused on identifying challenges in data visualization to overcome the lack of user-friendliness and avoid misunderstandings [1, 30]. For example, Alshehhi et al. [1] conducted a user survey to identify the needs and challenges of data visualizations in mHealth apps. They identified three main challenges (lack of interaction with the presented charts, inconsistency in chart presentation, and complexity of the data) and concluded that a new guideline is required. In addition, P. Wang and J. Wang [32] reported that the use of graphics and diagrams is crucial for data communication. They found that simple visualizations ("basic") have less visual impact but effectively capture users' attention, while complex visualizations ("analytical") provide more details and insights. These studies highlight the ongoing efforts, open questions, and challenges to make health data visualization more user-friendly and effective.

2.3 Impact of Chatbots on User Interaction

The emerging trend of integrating interactive chatbot applications into web pages or mobile apps is widespread, and users encounter them daily [7]. In recent years, chatbot applications have been increasingly used in the healthcare sector [16, 36]. Particularly for mHealth apps, the use of chatbots could improve user interaction and data understanding in the future. For example, Haque and Rubya [8] demonstrated that chatbot technology offers great potential for mental healthcare apps to overcome social challenges. Furthermore, Jovanovic et al. [12] analyzed chatbots for diagnosis, prevention, and therapy purposes in healthcare and discussed various design aspects. They concluded that future research should focus on investigating the user experience and its impact on health outcomes. Despite these promising applications, there is currently no evidence of the effective use of chatbots in foot health apps.

2.4 Summary

Previous work shows the SUS questionnaire is often used to assess usability, and that both analytical and basic visualizations of health data are commonly applied. Additionally, chatbots can positively impact user interaction in mHealth apps. However, it is unknown how analytical or basic visualizations of foot health data influence usability in smart insole apps or how much chatbot integration increases user-friendliness. By addressing this RQ, we aim to close the gap in usability challenges for foot patients and contribute to comprehensible visualizations in mHealth apps.

3 METHOD

The objective of this study was to evaluate the impact of different health data visualizations and chatbot integration on the usability of a mobile application for smart insoles. For this purpose, three mockups for mobile apps were developed and tested in an experimental user study. The hypothesis of our study is that chatbot interaction will lead to improved usability and understanding of health data compared to analytical and basic visualizations. We assume that providing textual and contextual information about foot health will support users in analyzing and interpreting the data. Therefore, we conducted a one-factorial within-subject experimental user study to investigate the usability of mobile apps for foot health monitoring. The independent variable PROTOTYPE includes three individual levels: *Analytical*, *Basic*, and *Chatbot* apps. Prototypes were randomized for each participant using a 3 x 3 balanced Latin square design to avoid sequence effects. We used the HCI User Studies Toolkit to plan the study, as recommended by previous work [28].

3.1 Prototypes

For our study, we developed three prototype design variations based on the app concept proposed by Resch et al. [25]. The main functions of the system include an overview of gait parameters, pressure measurement, personalized training instructions, a health knowledge database, and a walking diary. Based on these functions the three prototypes (*Analytical*, *Basic*, and *Chatbot*) were created using proto.io [1], as shown in Figure 1. In the study, the Hypertext Markup Language (HTML) versions of the prototypes were used on an Apple iPhone 14 running iOS 16.0.1. For comparability, the menu structure was kept the same and the content of the displayed health data was also identical, only the visualizations varied. All applications include features for displaying specific measurement data of pressure distribution and gait parameters, along with a walking diary that allows users to track their progress and rate their daily foot health level. In addition, the knowledge base and training mode provide access to information on foot health topics and training methods aimed at both prevention and management of existing problems. The *Basic* prototype is based on information visualization with simple 2D representations (e.g., foot pressure heatmap) and easily understandable text. The concept of the *Analytical* prototype is similar to *Basic* but includes more detailed and complex graphical data representations. This includes 3D visualizations of the pressure heatmap and graphs with multiple measurements and statistics for gait parameters. The *Chatbot* prototype relies on textual visualization, where the user interacts with a chatbot on a conversational basis. This app aims to represent the text input and response interaction with a text-generating artificial intelligence (AI) tool, with collected knowledge about the measured foot data. To enable simple interaction with the prototype it was possible to click on suggested example questions, to get an answer on specific health data. The chatbot interaction goes beyond simple text by providing context-specific answers and follow-up questions to simulate an interactive dialog in the context of health monitoring.

3.2 Procedure

After signing the informed consent, participants filled out a preliminary questionnaire via Google Forms to gather demographic data, health status, and mobile app expertise. Participants then received a verbal introduction to smart insoles and were given an insole prototype for demonstration purposes. Before starting, participants were instructed to think aloud while going through three realistic use scenarios for each application prototype. Participants could ask questions at any time during each condition, such as comprehension questions, which were documented. In addition, a sheet of paper with three pre-defined scenarios, as well as an iPhone with the corresponding prototype, was handed out. Screen recording on the iPhone and external audio recording documented the tasks. After completing the three conditions, participants filled out the SUS questionnaire to assess the usability of the app. This procedure was then repeated for the remaining two prototypes. Conditions and prototypes were randomized for each participant to avoid order effects. At the end, participants completed a post-study questionnaire via Google Forms to gather detailed feedback on the prototypes and user preferences. The total duration of the study was approximately 45 minutes.

3.3 Conditions - Scenarios

Three realistic use scenarios were created to explore the functions of each prototype and assess the clarity of visualized information. These scenarios were developed in team brainstorming sessions based on user requirements identified in previous interviews and analysis. This ensured that the scenarios were relevant and reflected real-life usage. Each scenario was structured like a typical everyday situation that the user is confronted with when interacting with the app. Using these scenarios ensured that each participant tested the primary functionalities of each app. The scenarios were provided to participants in a randomized order for each condition. All three test scenarios were described as short texts, with an additional tip concerning the corresponding functions. They are documented in the following subsections.

3.3.1 Scenario 1 - Walking Diary. The "walking diary" scenario is designed to help patients monitor their daily physical activity and health status. This scenario is important because it allows users to track their walking activity, document pain status, and examine their gait for potential problems. *"Imagine you come home after a long day at work and want to check how much you walked today. Your doctor recommended that you document your pain daily and record how you feel. Assess whether your gait was well balanced today or if there was a misalignment. While conducting this review, please think aloud about what you are doing and what thoughts are going through your mind. Evaluate whether the information is understandable and easy to follow for monitoring your daily health. Tip: Utilize the 'step count' function for this purpose."*

3.3.2 Scenario 2 - Pressure Measurement & Gait Parameters. This scenario aims to support patients in identifying and understanding issues related to their gait (e.g., problems with pressure distribution),

[1]https://proto.io/

Figure 1: Screenshots of analytical, basic, and chatbot mobile apps (from left to right), each showing six interfaces. Top row: home menu, pressure measurement, gait parameters; bottom row: walking diary, knowledge database, training mode.

which can be crucial in diagnosing and resolving potential problems. *"Imagine you recently noticed that your gait feels unusual and you have slight pain in your right leg. You observe that the insole of your right shoe is more worn. You suspect a problem with the insole. Please check the app for information about the insole of your right foot that could explain your discomfort. Also assess whether there are any abnormalities in your gait, particularly in the stance and swing phase. Please share your thoughts and impressions verbally. Tip: Utilize the 'pressure distribution / stability' & 'stance / swing Phase' function for this purpose."*

3.3.3 Scenario 3 - Knowledge Database & Training Mode. The "knowledge database & training mode" is intended to assist patients in finding information and exercises for pain management and relief, and to enable them to take proactive steps for their health. *"Imagine you are suffering from ankle pain and are looking for information about this issue and how to prevent it. You may want to find a suitable exercise to relieve pain. While you are searching for information, think out loud about what you are doing and what is going through your mind. Assess whether the information provided is understandable and easily accessible. Tip: Utilize the menu function: 'knowledge database' & 'training mode' for this purpose."*

3.4 Measures and Data Analysis

In this mixed-method study, we quantitatively measured usability and obtained qualitative feedback using the think-aloud protocol, supplemented by questionnaires. Data analysis methods are described in the following subsections.

3.4.1 Quantitative subjective: Usability. SUS was used to evaluate the usability of each prototype. After interacting with the system, participants rated ten aspects on a Likert scale from 1 (strongly disagree) to 5 (strongly agree) based on their subjective experience. The total score was calculated and converted into a value between 0 and 100, which quantifies the overall user-friendliness of the system. Quantitative data analysis was performed using descriptive statistics and inferential statistics, particularly a one-way repeated-measures (RM) analysis of variance (ANOVA).

3.4.2 Qualitative: Think-Aloud Protocol. The think-aloud method was used to gain deeper insights into cognitive issues while learning to use a mobile computer system, as recommended by Lewis [6]. Participants were asked to verbalize subjective perceptions of thoughts and impressions. Recorded audio data was transcribed using Adobe Premiere Pro 2023 [2] software and independently reviewed by two researchers for accuracy. A third researcher verified the transcribed data for consistency. An inductive thematic analysis was performed on the transcriptions to identify the main themes of the data set [4]. Two researchers conducted the analysis separately by open-coding the anonymized data paragraph by paragraph. After cross-checking, axial coding was performed to identify patterns in the categorized data.

3.4.3 Questionnaires. Questionnaires were used before and after the study to collect qualitative and quantitative data on user preferences. Both questionnaires contain multiple-choice, open-ended, and closed-ended questions. Additionally, 5-item Likert scales were used to rate statements on satisfaction, likability, and familiarity. The preliminary questionnaire included questions on demographic data, health status, existing foot problems, use of orthopedic devices, daily smartphone usage, and expectations for smart insole app features. Furthermore, participants completed a self-assessment on their expertise with new smartphone technologies/apps, knowledge of orthopedic or medical fields, and understanding of data analysis/interpretation. The post-study questionnaire includes questions

[2]https://www.adobe.com/de/products/premiere.html

on the evaluation of individual prototypes and associated visualized features, reasons for satisfaction with the app, as well as personal preferences, possible improvements, and additional feedback. Five-point Likert scales were evaluated by descriptive and inferential statistics, using the non-parametric Friedman rank sum test. Descriptive statistics were also applied to analyze the data from multiple-choice and closed-ended questions. Open-ended questions were evaluated using a qualitative thematic analysis, similar to the think-aloud protocol.

3.5 Participants

This study involved 30 participants (15 female, 15 male), aged between 20 and 59 years (M = 32.93, SD = 13.44), from diverse educational backgrounds and occupations. These included 15 with a high school diploma, five with vocational training, five with a bachelor's degree, two with no formal education, two with a master's degree, and one with a doctoral degree. Participants' occupations represented the fields of education, engineering, healthcare, information technology, journalism, law enforcement, and management. Participants were recruited through our university's e-learning platform, institutional mailing lists, and personal networks. Inclusion criteria were either an acute/past foot condition, the use of orthopedic aids, or previous experience/familiarity with health monitoring smartphone apps. The study received ethical clearance from the German Society for Nursing Science (No. 23-027) and followed the data privacy regulations and hygiene protocols of our institution. All participants completed the study without interruption and were included in the data analysis. A total of 16 participants stated that they used orthopedic aids, including insoles (N = 12), bandages (N = 4), and cast-braces (N = 2). Twelve participants reported foot conditions, including general foot complaints & misalignments (N = 5), ankle joint overpronation (N = 2), flat feet (N = 2), osteoarthritis (N = 1), plantar fasciitis (N = 1), or a broken middle toe (N = 1). Eighteen participants had used health or fitness monitoring apps or devices, including Apple Health, Samsung Health Monitor, Google Fitbit, and Nike Run Club. Six participants used their smartphone for less than two hours a day, twelve participants for 2-4 hours, eleven for 4-6 hours, and one for more than eight hours. For self-assessment, participants rated their skills in using new smartphone app technology on a Likert scale from 1 (not at all familiar) to 5 (extremely familiar). Eighteen rated themselves as "moderately familiar," eight as "somewhat familiar," and four as "extremely familiar." Knowledge in orthopedic or medical health care was rated as "slightly familiar" by nine participants, "not at all familiar" by eight, "somewhat familiar" by six, "moderately familiar" by five, and "extremely familiar" by two. Knowledge of data analysis was rated as "slightly familiar" by fourteen participants, "somewhat familiar" and "moderately familiar" by seven each, and "not at all familiar" by two.

4 RESULTS

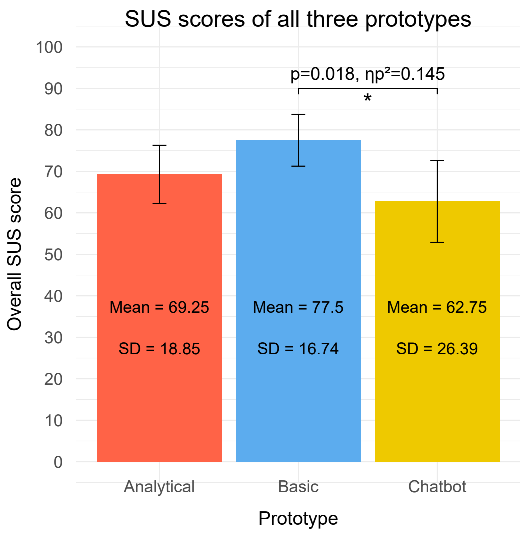
This section presents the statistical analysis of quantitative SUS scores (subsection 4.1) using R, complemented by the qualitative analysis of think-aloud protocols (subsection 4.2), and the analysis of both questionnaires (subsection 4.3).

4.1 Quantitative subjective: Usability

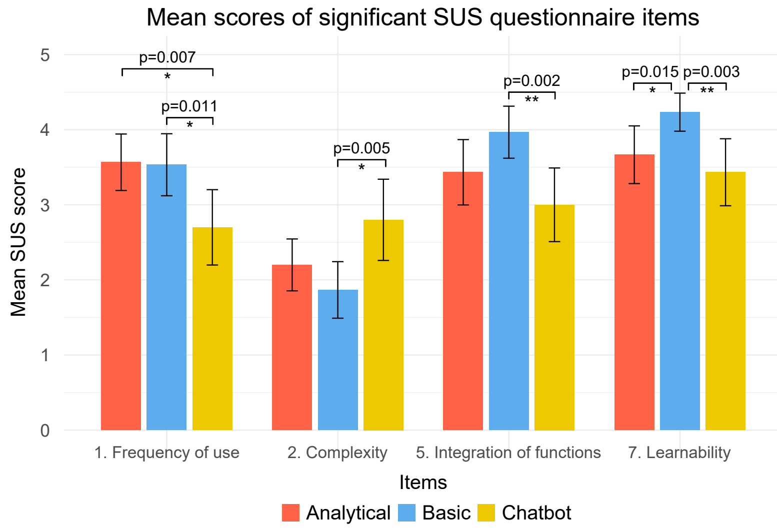
We conducted a one-way RM-ANOVA on the SUS scores to investigate the impact of each Prototype on usability. The Shapiro-Wilk normality test indicated normal distribution for all prototypes ($p \geq 0.05$). One-way RM-ANOVA showed a significant main effect of Prototype on the SUS scores, $F(1.52, 44.16) = 4.937$, $p = 0.018$, $\eta_p^2 = 0.145$ (large effect size). Performing pairwise post-hoc comparisons using Bonferroni-corrected paired t-tests revealed a significant difference between *Chatbot* and *Basic* ($p = 0.013$, $p_{\text{adj.}} = 0.038$). However, no significant differences were found between *Analytical* and *Chatbot* ($p = 0.277$, $p_{\text{adj.}} = 0.277$), as well as between *Analytical* and *Basic* ($p = 0.078$, $p_{\text{adj.}} = 0.157$), see Figure 2a. Additionally, we analyzed the ten SUS questionnaire items to further investigate usability differences between prototypes. The Shapiro-Wilk normality test indicated non-normal distribution across all items (all with $p < 0.05$). Therefore, we performed a non-parametric one-way RM-ANOVA on each of the ten SUS questionnaire items to assess differences between the prototypes. Significant main effects with large effect sizes were identified for five questionnaire items (1, 2, 5, and 7). To determine between which prototypes these effects occur, Bonferroni-corrected pairwise t-tests were performed as post hoc comparisons, which revealed six significant effects (see Figure 2b).

4.2 Qualitative: Think-Aloud Protocol

The *Analytical* app was perceived as "useful and intuitively understandable ... as there were many graphics and diagrams" (P8). Additionally, it was stated that "the detailed information felt like something a lot of other comparable apps would not be actively offering and would fit a niche for those looking to track more than just calories burned or steps taken" (P13). However, participants mentioned that it "requires prior knowledge" (P16, P21). Participants criticized that this app "did not seem to offer any benefits compared to the basic version" (P1). Compared to the other prototypes, the *Basic* app was perceived as "accessible for all" (P16), "more intuitive" (P4), "pleasant to use and user-friendly" (P29), and "easy to use and understand" (P18, P24, P30). Participants stated that this app "combined both approaches and was therefore more easy to understand" (P4) and that it "was exactly as I would expect a basic app of this nature to function" (P13). As a suggestion for improvement, it was mentioned that it still needs "a little more fine-tuning in terms of UI" (P8). The *Chatbot* app was preferred by some participants because "it was the easiest to find more complex things" (P1) and "because of the enabled interaction between the user and the system" (P3). Participants mentioned that this version "... could help people who do not have enough knowledge to interpret the different statistics" (P24) and "... would be the most pleasant for older people" (P28). A positive aspect was that it takes "less effort navigating the app to find something" (P1) and also the "possibility for specific questions and follow-up questions" (P8). One participant explained that it "was simple to use but was quite cumbersome in having to ask every statistic you wanted to view first" (P18). Negative aspects were "too much text" (P7, P25), "missing illustrations" (P15), "not intuitive" (P10), and "no good overview" (P10, P25). Additionally, it was stated that "you had to ask too many questions to get information" (P15), which was perceived as "exhausting" (P9) and "would be annoying on a daily basis" (P20).



(a) Overview of the SUS scores for each prototype



(b) Comparison of significant SUS questionnaire items

Figure 2: Bar charts with descriptive statistics of SUS scores for each prototype (a), and analysis of significant SUS items in (b).

4.3 Questionnaires

In the post-study questionnaire, participants rated their satisfaction with all three apps on a Likert scale from 1 (not satisfied) to 5 (very satisfied). A Friedman rank sum test on non-parametric data showed a significant main effect of PROTOTYPE: $X^2(2) = 15.7$, $p < 0.001$ (small effect size using Kendall's W). Pairwise comparisons using Wilcoxon signed-rank tests showed a significant difference between *Analytical* and *Chatbot* ($p = 0.022$, $p_{\text{adj.}} = 0.044$), and *Basic* and *Chatbot* ($p = 0.001$, $p_{\text{adj.}} = 0.004$). However, no significant effect was found between *Analytical* and *Basic* ($p = 0.069$, $p_{\text{adj.}} = 0.069$). Additionally, 13 participants chose the *Basic* app as their favorite, twelve selected the *Analytical* prototype, and five voted for the *Chatbot* app as their favorite. Regarding the understanding of data visualization, a total of 17 participants rated the *Basic* app as the easiest to understand, seven preferred *Chatbot*, and six chose *Analytical*. On a Likert scale from 1 (extremely unlikely) to 5 (extremely likely), the regular future use of this app for medical needs was rated as "likely" by 17 participants, "neutral" by eight, "extremely likely" by three, and "unlikely" as well as "extremely unlikely" each by one. Sixteen participants rated it as "likely" that they would recommend the favored prototype app to people with orthopedic foot problems. Eight participants indicated "neutral," five rated "extremely likely," and one voted "extremely unlikely."

5 DISCUSSION

We conducted an experimental study with 30 participants to investigate the effect of *Analytical*, *Basic*, and *Chatbot* visualizations on usability in smart insole apps. Results of the SUS questionnaire showed that the *Basic* app leads to the highest usability, followed by *Analytical*, and lastly *Chatbot*. A significant difference in usability was detected between *Basic* and *Chatbot*. The result achieved by *Basic* is *acceptable* (acceptability scale) or grade *B* (grade scale), according to the interpretation proposed by Bangor et al. [3]. In contrast, *Analytical* and *Chatbot* were classified as *marginal* (acceptability scale), which corresponds to a grade of *C* (grade scale). Qualitative results, including the think-aloud protocol and questionnaires, confirmed the results. Based on these findings, we conclude that graphical representations of health data, whether basic charts or analytical, are more user-friendly. Nevertheless, further information and background knowledge are needed to interpret health data effectively. In contrast, interaction with a *Chatbot* enables users to receive specific explanations of health data, which improves understanding. Receiving precise data on request was perceived as positive. Despite this, the interaction negatively impacts user-friendliness due to the lack of graphical visualization. Nevertheless, the implementation of an interactive chatbot in a mHealth app could enhance user engagement. However, further improvements are needed for better integration. We propose combining text-based chatbot interactions with graphical visualizations in smart insole apps to leverage both strengths. This combination could enhance health data understanding and improve overall usability. In general, our study suggests that different visualizations offer varying advantages for different target audiences. *Analytical* visualizations require prior technical knowledge and offer detailed insights for users with expertise, whereas *Basic* visualizations are accessible to a wide audience. Interaction with a *Chatbot* could be beneficial for older users as it helps them understand complex data. A limitation of this study is the controlled setting. Therefore, future research should test the app in a field study to validate its effectiveness.

ACKNOWLEDGMENTS

This research was funded by the Hessian Ministry of Science and Art - HMWK, Germany (FL1, Mittelbau) and supported by the Frankfurt University of Applied Sciences and the University of Cadiz.

REFERENCES

[1] Yasmeen Anjeer Alshehhi, Ben Philip, Mohamed Abdelrazek, and Alessio Bonti. 2023. Needs and Challenges of Personal Data Visualisations in Mobile Health Apps: User Survey. In *2023 IEEE International Conference on Big Data and Smart Computing (BigComp)*. IEEE, Jeju, Republic of Korea, 295–297. https://doi.org/10.1109/BigComp57234.2023.00058

[2] Bliss Altenhoff, Haley Vaigneur, and Kelly Caine. 2015. One Step Forward, Two Steps Back: The Key to Wearables in the Field is the App. In *Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare (ACM Other conferences)*, Bert Arnrich (Ed.). ICST (Institute for Computer Sciences Social-Informatics and Telecommunications Engineering), Brussels,Belgium. https://doi.org/10.4108/icst.pervasivehealth.2015.259049

[3] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 4, 3 (2009), 114–123.

[4] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[5] Diliang Chen, Golnoush Asaeikheybari, Huan Chen, Wenyao Xu, and Ming-Chun Huang. 2021. Ubiquitous Fall Hazard Identification With Smart Insole. *IEEE journal of biomedical and health informatics* 25, 7 (2021), 2768–2776. https://doi.org/10.1109/JBHI.2020.3046701

[6] Clayton, Lewis. 1982. *Using the 'thinking-aloud' method in cognitive interface design.* Technical Report Research Report RC9265. IBM T. J. Watson Research Center, Yorktown Heights, NY. https://cir.nii.ac.jp/crid/1571135649469604736

[7] Asbjørn Følstad, Theo B. Araujo, Effie Lai-Chong Law, Petter Bae Brandtzaeg, Symeon Papadopoulos, Lea Reis, Marcos Báez, Guy Laban, Patrick McAllister, Carolin Ischen, Rebecca Wald, Fabio Catania, Raphael Meyer von Wolff, Sebastian Hobert, and Ewa Luger. 2021. Future directions for chatbot research: an interdisciplinary research agenda. *Computing* 103 (2021), 2915 – 2942. https://api.semanticscholar.org/CorpusID:239048434

[8] M D Romael Haque and Sabirat Rubya. 2023. An Overview of Chatbot-Based Mobile Mental Health Apps: Insights From App Description and User Reviews. *JMIR Mhealth Uhealth* 11 (22 May 2023), e44838. https://doi.org/10.2196/44838

[9] Meng-Hsueh Hsieh, Yu-Ching Chen, and Chun-Heng Ho. 2019. A Usability Evaluation of Diabetes Mobile Applications. In *Human-Computer Interaction. Design Practice in Contemporary Societies (Information Systems and Applications, incl. Internet/Web, and HCI)*, Masaaki Kurosu (Ed.). Springer International Publishing and Imprint: Springer, Cham, 3–15. https://doi.org/10.1007/978-3-030-22636-7{_}1

[10] Institute of Medicine. 2012. *Health IT and Patient Safety: Building Safer Systems for Better Care.* The National Academies Press, Washington, DC. https://doi.org/10.17226/13269

[11] John Brooke. 1996. SUS: a "quick and dirty" usability scale. In *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and A. L. McClelland (Eds.). Taylor and Francis, London.

[12] Mladan Jovanovic, Marcos Baez, and Fabio Casati. 2021. Chatbots as Conversational Healthcare Services. *IEEE Internet Computing* 25, 3 (2021), 44–51. https://doi.org/10.1109/MIC.2020.3037151

[13] Kamran Khowaja and Dena Al-Thani. 2020. New Checklist for the Heuristic Evaluation of mHealth Apps (HE4EH): Development and Usability Study. *JMIR Mhealth Uhealth* 8, 10 (28 Oct 2020), e20353. https://doi.org/10.2196/20353

[14] James R. Lewis. 2002. Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. *International Journal of Human–Computer Interaction* 14, 3-4 (2002), 463–488. https://doi.org/10.1080/10447318.2002.9669130

[15] Feng Lin, Aosen Wang, Yan Zhuang, Machiko R. Tomita, and Wenyao Xu. 2016. Smart Insole: A Wearable Sensor Device for Unobtrusive Gait Monitoring in Daily Life. *IEEE Transactions on Industrial Informatics* 12, 6 (2016), 2281–2291. https://doi.org/10.1109/tii.2016.2585643

[16] Haruhisa Maeda, Sachio Saiki, Masahide Nakamura, and Kiyoshi Yasuda. 2019. Recording Daily Health Status with Chatbot on Mobile Phone - A Preliminary Study. In *2019 Twelfth International Conference on Mobile Computing and Ubiquitous Network (ICMU)*. IEEE, Kathmandu, Nepal, 1–6. https://doi.org/10.23919/ICMU48249.2019.9006645

[17] Nur Syafikin Shaheera Mat Zaini, Syed Nasir Alsagoff Syed Zakaria, and Norshahriah Wahab. 2017. Methods of Evaluating the Usability of Human- Computer Interaction (HCI) Design in Mobile Devices for SAR Operation. In *Advances in Visual Informatics*, Halimah Badioze Zaman, Peter Robinson, Alan F. Smeaton, Timothy K. Shih, Sergio Velastin, Tada Terutoshi, Azizah Jaafar, and Nazlena Mohamad Ali (Eds.). Springer International Publishing, Cham, 714–726.

[18] Jochen Meyer, Anastasia Kazakova, Merlin Büsing, and Susanne Boll. 2016. Visualization of Complex Health Data on Mobile Devices. In *Proceedings of the 2016 ACM Workshop on Multimedia for Personal Health and Health Care* (Amsterdam, The Netherlands) *(MMHealth '16)*. Association for Computing Machinery, New York, NY, USA, 31–34. https://doi.org/10.1145/2985766.2985774

[19] Kristin D. Hansen Morten Hertzum and Hans H.K. Andersen. 2009. Scrutinising usability evaluation: does thinking aloud affect behaviour and mental workload? *Behaviour & Information Technology* 28, 2 (2009), 165–181. https://doi.org/10.1080/01449290701773842

[20] Khadijeh Moulaei, Reza Moulaei, and Kambiz Bahaadinbeigy. 2024. The most used questionnaires for evaluating the usability of robots and smart wearables: A scoping review. *DIGITAL HEALTH* 10 (2024), 20552076241237384. https://doi.org/10.1177/20552076241237384

[21] Antonio Muro-Culebras, Adrian Escriche-Escuder, Jaime Martin-Martin, Cristina Roldán-Jiménez, Irene De-Torres, Maria Ruiz-Muñoz, Manuel Gonzalez-Sanchez, Fermin Mayoral-Cleries, Attila Biró, Wen Tang, Borjanka Nikolova, Alfredo Salvatore, and Antonio Ignacio Cuesta-Vargas. 2021. Tools for Evaluating the Content, Efficacy, and Usability of Mobile Health Apps According to the Consensus-Based Standards for the Selection of Health Measurement Instruments: Systematic Review. *JMIR Mhealth Uhealth* 9, 12 (1 Dec 2021), e15433. https://doi.org/10.2196/15433

[22] Eman Nasr, Wafaa Alsaggaf, and Doaa Sinnari. 2023. Developing Usability Guidelines for mHealth Applications (UGmHA). *Multimodal Technologies and Interaction* 7, 3 (2023). https://doi.org/10.3390/mti7030026

[23] Jakob Nielsen. 1994. *Usability engineering.* Morgan Kaufmann, Burlington, MA, USA.

[24] Lumpapun Punchoojit and Nuttanont Hongwarittorrn. 2017. Usability Studies on Mobile User Interface Design Patterns: A Systematic Literature Review. *Advances in Human-Computer Interaction* 2017, 1 (2017), 6787504. https://doi.org/10.1155/2017/6787504

[25] Stefan Resch, Khalid Zoufal, Imad Akhouaji, Mohamed-Amin Abbou, Valentin Schwind, and Diana Völz. 2023. Augmented Smart Insoles – Prototyping a Mobile Application: Usage Preferences of Healthcare Professionals and People with Foot Deformities. *Current Directions in Biomedical Engineering* 9, 1 (2023), 698–701. https://doi.org/10.1515/cdbme-2023-1175

[26] Bidisha Roy, Mark Call, and Natalie Abts. 2019. Development of Usability Guidelines for Mobile Health Applications. In *HCI International 2019 - Posters*, Constantine Stephanidis (Ed.). Springer International Publishing, Cham, 500–506.

[27] Salma Saidani, Rim Haddad, and Ridha Bouallegue. 2020. A prototype design of a smart shoe insole system for real-time monitoring of patients. In *2020 6th IEEE Congress on Information Science and Technology (CiSt)*. IEEE, Agadir, Morocco, 116–121. https://doi.org/10.1109/cist49399.2021.9357177

[28] Valentin Schwind, Stefan Resch, and Jessica Sehrt. 2023. The HCI User Studies Toolkit: Supporting Study Designing and Planning for Undergraduates and Novice Researchers in Human-Computer Interaction. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (ACM Digital Library)*, Albrecht Schmidt (Ed.). Association for Computing Machinery, New York,NY,United States, 1–7. https://doi.org/10.1145/3544549.3585890

[29] Guna Semjonova, Anna Davidovica, Nikita Kozlovskis, Aleksandrs Okss, and Aleksejs Katashevs. 2022. Smart Textile Sock System for Athletes' Self-Correction during Functional Tasks: Formative Usability Evaluation. *Sensors (Basel, Switzerland)* 22, 13 (2022). https://doi.org/10.3390/s22134779

[30] Ben Shneiderman, Catherine Plaisant, and Bradford W. Hesse. 2013. Improving Healthcare with Interactive Visualization. *Computer* 46, 5 (2013), 58–66. https://doi.org/10.1109/MC.2013.38

[31] Sophini Subramaniam, Sumit Majumder, Abu Ilius Faisal, and M. Jamal Deen. 2022. Insole-Based Systems for Health Monitoring: Current Solutions and Research Challenges. *Sensors* 22, 2 (2022). https://doi.org/10.3390/s22020438

[32] Pengwen Wang and Jiangru Wang. 2021. Research on the Application of Data Visualization in the UI interface Design of Health Apps. In *2021 International Wireless Communications and Mobile Computing (IWCMC)*. IEEE, Harbin City, China, 2013–2019. https://doi.org/10.1109/IWCMC51323.2021.9498648

[33] Pawel Weichbroth. 2020. Usability of Mobile Applications: A Systematic Literature Study. *IEEE Access* 8 (2020), 55563–55577. https://doi.org/10.1109/access.2020.2981892

[34] Zulvi Wiyanti, Fadil Oenzil, Masrul Masrul, and Defriman Djafri. 2021. A Mobile Health App for Premarital Related Word Advice (AmeeSehat): Usability and Usefulness Evaluation Study. *Open Access Macedonian Journal of Medical Sciences* 9, E (2021), 474–480.

[35] Jiabei Wu and Vincent G. Duffy. 2022. Mobile Applications Usability Evaluation: Systematic Review and Reappraisal. In *HCI International 2022 - Late Breaking Papers. Design, User Experience and Interaction (Lecture Notes in Computer Science)*, Masaaki Kurosu, Sakae Yamamoto, Hirohiko Mori, Marcelo M. Soares, Elizabeth Rosenzweig, Aaron Marcus, Pei-Luen Patrick Rau, Don Harris, and Wen-Chin Li (Eds.). Springer International Publishing, Cham, 499–516. https://doi.org/10.1007/978-3-031-17615-9_35

[36] Lu Xu, Leslie Sanders, Kay Li, and James C. L. Chow. 2021. Chatbot for Health Care and Oncology Applications Using Artificial Intelligence and Machine Learning: Systematic Review. *JMIR cancer* 7, 4 (2021), e27850. https://doi.org/10.2196/27850

[37] Leming Zhou, Jie Bao, I Made Agus Setiawan, Andi Saptono, and Bambang Parmanto. 2019. The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study. *JMIR Mhealth Uhealth* 7, 4 (11 Apr 2019), e11500. https://doi.org/10.2196/11500