Original Paper

German Version of the mHealth App Usability Questionnaire in a Cohort of Patients With Cancer: Translation and Validation Study

Preetha Moorthy^{1*}, MSc; Lina Weinert^{2,3*}, MSc; Bendix Christian Harms¹; Carolin Anders^{2*}, MSc; Fabian Siegel^{1*}, MD

¹Department of Biomedical Informatics, Center for Preventive Medicine and Digital Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

²Institute of Medical Informatics, Heidelberg University Hospital, Heidelberg, Germany

³Section for Translational Health Economics, Department for Conservative Dentistry, Heidelberg University Hospital, Heidelberg, Germany *these authors contributed equally

Corresponding Author:

Preetha Moorthy, MSc Department of Biomedical Informatics, Center for Preventive Medicine and Digital Health, Medical Faculty Mannheim Heidelberg University Theodor-Kutzer-Ufer 1-3 Mannheim, 68167 Germany Phone: 49 621 383 8078 Email: preetha.moorthy@medma.uni-heidelberg.de

Abstract

Background: Good usability is important for the adoption and continued use of mobile health (mHealth) apps. In particular, high usability can support intuitive use by patients, which improves compliance and increases the app's effectiveness. However, many usability studies do not use adequate tools to measure perceived usability. The mHealth App Usability Questionnaire (MAUQ) was developed specifically for end users in a medical context. MAUQ is a relatively new but increasingly used questionnaire to evaluate mHealth apps, but it is not yet available in German.

Objective: This study aims to translate MAUQ into German and determine its internal consistency, reliability, and construct validity.

Methods: This validation study was conducted as part of a usability evaluation project for an mHealth app used as a therapy support tool during breast cancer chemotherapy. MAUQ was translated into German through a rigorous forward-backward translation process, ensuring semantic and conceptual equivalence. Patient responses to MAUQ and System Usability Scale (SUS) were analyzed for validation. Descriptive analysis was performed for the MAUQ subscales and SUS standard scores. Significance tests and correlation coefficients assessed the relationship between the SUS and MAUQ results, confirming construct validity. Internal consistency was assessed for item reliability and consistency in measuring the target construct. Free-text questions assessed translation comprehensibility, with responses analyzed descriptively and qualitatively using content analysis.

Results: In this study, 133 participants responded to the questionnaire, and the validation analysis showed substantially positive correlations between the overall MAUQ score and its subscales: ease of use (r=0.56), interface and satisfaction (r=0.75), and usefulness (r=0.83). These findings support the construct validity of MAUQ and emphasize the importance of these subscales in assessing the usability of the Enable app. The correlation coefficients ranging from 0.39 to 0.68 for the items further validate the questionnaire by aligning with the overall score and capturing the intended concept. The high internal consistency reliability of MAUQ (Cronbach α =.81) and its subscales further enhances the instrument's robustness in accurately evaluating the usability of mHealth apps.

Conclusions: We successfully validated the German translation of the MAUQ for stand-alone apps using a standardized approach in a cohort of patients with breast cancer. In our validation study, MAUQ exhibited strong internal consistency reliability (Cronbach α =.81) across its subscales, indicating reliable and consistent measurement. Furthermore, a significant positive correlation (*P*<.001) was found between the subscales and the overall score, supporting their consistent measurement of the intended construct. Therefore, MAUQ can be considered a reliable instrument for assessing the usability of mHealth apps among German-speaking

adults. The availability of the German version of MAUQ will help other researchers in conducting usability studies of mHealth apps in German-speaking cohorts and allow for international comparability of their results.

(JMIR Hum Factors 2023;10:e51090) doi: 10.2196/51090

KEYWORDS

usability; mobile health; mHealth app; questionnaire validation; questionnaire translation; mHealth App Usability Questionnaire; MAUQ; mobile phone

Introduction

Background

The use of digital technology in both routine health care and research continues to rise. The increasing adoption of fitness and medical apps has reached a global market size of US \$43.5 billion in 2022. This increment is projected to experience a compound annual growth rate of 11.6% from 2023 to 2030 [1]. As defined by the World Health Organization (WHO) Global Observatory for eHealth, mobile health (mHealth) encompasses medical and public health practices supported by mobile devices, such as mobile phones, patient monitoring devices, PDAs, and other wireless devices by health care professionals or patients [2]. According to Morse et al [3], mHealth apps refer to software integrated into smartphones with the aim of enhancing health outcomes, health research, and health care services. The continuous increase of reporting, data collection, telemedicine, and emergency medical care using mHealth apps draws attention to the acceptance and user experience of the targeted users. Knowledge about these factors could assist physicians and health care providers in choosing the right mHealth apps for their patients. However, many research studies do not adequately evaluate mHealth interventions nor provide sufficient evidence about the health impact [2-4]. To facilitate such studies, it is essential to establish and reach a consensus on standardized indicators and metrics for monitoring and evaluating purposes. Questionnaires are one of the well-established methods that are widely used in research, clinical trials, and health care settings to gather data and measure various constructs. One of the oldest usability measurement scales, the System Usability Scale (SUS) [5], was developed in 1986 to assess the usability of electronic office systems. Nowadays, SUS is widely used and considered a highly reliable tool for evaluating software, websites, or mobile apps.

When questionnaires need to be used in different cultural and linguistic contexts, it is essential to ensure their equivalence across languages. The identification of the construct to be evaluated using the questionnaire is critical, as it defines the scope of interest and determines the type of measurements that will be obtained. The limited availability of standardized usability questionnaires in languages other than English poses a potential challenge when assessing system usability and user experience among non–English-speaking populations. Only a few usability-related evaluation questionnaires are currently available in the German language such as International Organization for Standardization (ISO) Norm 9241/110 [6], AttrakDiff [7], Scale for Measuring Perceived Website Usability [8], Software Usability Measurement Inventory [9], User

Experience Questionnaire [10], and Mobile Application Rating Scale [11].

Compared with the abovementioned and established usability evaluation questionnaires, the mHealth App Usability Questionnaire (MAUQ) [12] is a relatively new questionnaire in this area. Therefore, it has been used less than other established questionnaires such as SUS [13]. In contrast to the Mobile Application Rating Scale, which is also used to evaluate mHealth apps, MAUQ was developed specifically and initially for the target group of end users in a medical context. However, owing to its specialization for mHealth apps, MAUQ is becoming increasingly popular internationally and has been translated into various languages [14-18]. In the German context, some studies have used self-translated, nonvalidated versions of MAUQ. This shows the need for a validated, translated version of MAUQ for use in German-speaking populations [19-24].

Objectives

This study aimed to achieve a linguistically and conceptually equivalent version of the stand-alone version of the MAUQ instrument in German for the target population.

Methods

Overview of the Questionnaire

MAUQ, developed and validated by Zhou et al [12], assesses the usability of mHealth apps among patients and health care providers. The stand-alone version of MAUQ was used in this study. MAUQ consists of 18 items distributed across 3 dimensions: ease of use (5 items), interface and satisfaction (7 items), and usefulness (6 items). These dimensions capture various aspects related to the usability of the assessed mHealth apps. Using a 7-point Likert scale ranging from 1 (strongly agree) to 7 (strongly disagree), participants rated each item in the questionnaire. Usability was determined by calculating the total and average scores of all statements, where high average scores indicate a high level of usability. The results showed that strong internal consistency was observed in the overall MAUQ for stand-alone apps used by patients (Cronbach α =.914). The respective subscales of MAUQ also exhibited strong internal consistency reliability, as evidenced by Cronbach α coefficients of ease of use (Cronbach α =.847), interface and satisfaction (Cronbach α =.908), and usefulness (Cronbach α =.717).

Participants

Participants were recruited as part of the project, Multicenter Digital Recording of Patient Satisfaction, Quality of Life, and Patient-Reported Adverse Events in Breast Carcinoma in Neoadjuvant, Adjuvant, Follow-Up, and Palliative Care, in

XSL•F() RenderX

short, ENABLE [25], a randomized controlled trial (RCT) that aimed to improve patients' adherence to breast cancer therapy. We followed the recommendations proposed in previous studies [26-28], particularly when conducting the exploratory factor analysis, and used the sample-to-item ratio to determine an appropriate sample size based on the number of items in the study. In accordance with the suggested criterion of maintaining a sample-to-item ratio of no less than 5:1, a preliminary calculation led us to determine an appropriate sample size of 90 before initiating the validation study. It is noteworthy that MAUQ comprises 18 items, and as per the established recommendation, each item necessitates responses from 5 participants [26-28].

Participants were screened to meet the following inclusion criteria: diagnosis of invasive or metastatic breast cancer and planning neoadjuvant, adjuvant, or palliative therapy; minimum age of 18 years; German language skills; and possession of a tablet or smartphone with internet access. Participants were eligible for this study once they were enrolled in the ENABLE study.

Procedure

The aim was to use a standardized approach to translate a usability questionnaire and to ensure that the instrument is equally natural and acceptable and functions effectively across different cultural contexts. An established approach to accomplish this objective is by using forward translations and backward translations. This is a widely used approach to evaluate the comprehensibility of a source text and to identify any errors or uncertainties that may require attention or rectification while finalizing the text to perform translation followed by backward translation [29-31].

Study Context

This validation study was conducted in the context of the ENABLE project, which investigated the use of an mHealth

app as a therapy support tool for patients undergoing chemotherapy for breast cancer. The studied mHealth app, called Enable app, provides the opportunity to conduct reactive patient-reported outcome assessments, such as screening for adverse events, health-related quality of life, and patient satisfaction. In addition, the app offers patients information about their therapy, medication, and common adverse events. The app also visualizes the patient's therapy progress using a progress bar and includes information about upcoming appointments. Both the clinical effectiveness and the usability of this newly developed app were addressed in the ENABLE study, a randomized clinical trial conducted at 3 university hospitals in Germany. The usability of the Enable app was measured through a combination of qualitative measures (semistructured interviews and eye tracking) and quantitative measures (MAUQ and SUS).

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and was approved by the ethics committee of the Heidelberg University Hospital (S-685/2020). Confidentiality and anonymity were ensured throughout the entire study. Study participants provided written informed consent.

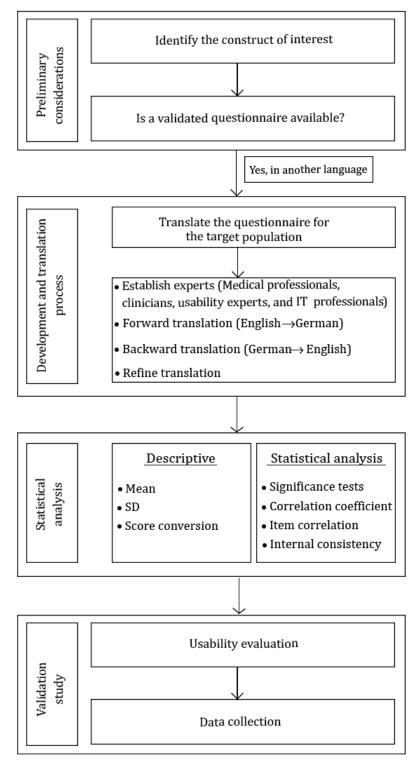
Translation Process

Overview

MAUQ was translated from English to German with the help of translators certified by the ISO 17100:2015 norm. We adapted the WHO guidelines for the translation process (Figure 1): (1) forward translation, (2) expert panel, (3) backward translation, and (4) refine translation [32]. To ensure accuracy and consistency in the translation process from the source to the target language, 2 independent and certified translators performed the translation [33-35]. The different versions of the translation and the final German translation of MAUQ are included in Multimedia Appendices 1 and 2, respectively.



Figure 1. Depiction of the translation process for the purpose of this study.



Forward Translation

During this phase, an experienced and certified translator, proficient in both the source and target languages, carefully translated the questionnaire from the source language, English, to the target language, German. In this way, the linguistic nuances, idiomatic expressions, and cultural references captured can be compared with the original meaning and intent of the questions. This step ensured that the translated questionnaire

was understandable to the target population and maintained the integrity of the original instrument.

Expert Panel

For this phase, a review panel was created, involving skilled experts from different domains, namely, clinicians, medical informaticians, and usability experts. The panel consisted of members who were bilingual and proficient in both the English and the German language. The experts were asked to review all

XSL•F() RenderX

the translations and identify and resolve the concepts of translation that are inadequate. In case of discrepancies between the forward translation and the original, experts could address specific words or expressions in the translated questionnaire and provide alternative suggestions.

Backward Translation

Following the forward translation, a second independent translator who is also ISO certified, experienced, fluent in the target language, and proficient in the source language performed the backward translation. This translator independently translated the questionnaire back to the source language. The purpose of the backward translation was to compare the translated version with the original questionnaire, allowing for an assessment of linguistic accuracy and potential inconsistencies.

Refine Translation

The expert panel thoroughly reviewed the backward-translated version and compared it with the original questionnaire and the forward translation to identify any differences or contrarieties. Irregularities found during the comparison were discussed and reconciled through collaboration among the expert panel and the project stakeholders. This step involved a careful examination of the wording, structure, and semantics of the items. The goal was to ensure that the final translated version maintains semantic and conceptual equivalence with the original questionnaire.

Data Collection

Data collection for this study was conducted from May 2021 to October 2022. MAUQ was administered at 2 different time points: once at 4 weeks after the individual study started and again after 20 weeks. All patients enrolled in the ENABLE project at the Department of Gynecology and Obstetrics, Heidelberg University Hospital, received a printed version of MAUQ and SUS (10 wk after the individual study started) via postal mail. The original, translated surveys were complemented with additional questions developed by the authors (eg, regarding sociodemographic data, use of other mHealth apps, history of smartphone ownership, and understandability of the newly translated MAUQ). Patients were asked to return the completed questionnaires via postal mail.

Survey data were collected and managed using REDCap (Research Electronic Data Capture; Vanderbilt University) electronic data capture tools [36] hosted at the Heidelberg University Hospital. All data were exported from REDCap to R statistical software (version 4.0.4; R Foundation for Statistical Computing). All data were checked for plausibility and analyzed by the study team members.

Statistical Analysis

Descriptive Analysis

An initial descriptive analysis was conducted to examine participant demographics and assess the performance of MAUQ and SUS, by calculating means and SDs for individual items and the overall questionnaire. In this study, MAUQ included 18 items, which were categorized into 3 subscales: ease of use (items 1-5), interface and satisfaction (items 6-12), and usefulness (items 13-18). We calculated the scores for each subscale and the total score for MAUQ, with all scores ranging from 0 to 100 [12]. In addition, we used a standard score conversion procedure for the SUS questionnaire to convert participants' responses into scores ranging from 0 to 100 [37].

Significance Tests

We used the significance tests to assess the statistical significance of differences or relationships within SUS, MAUQ, and MAUQ's subscales. These tests help determine whether the observed findings reflect actual effects or relationships within the population rather than random chance. In addition, the discriminative ability assesses the questionnaire's capacity to differentiate between varying levels of attributes, capturing subtle distinctions in responses for meaningful comparisons and conclusions [38,39]. As such, identifying significant differences between the highest and lowest quintile means is crucial.

Correlation Coefficient and Internal Consistency

The correlation coefficients were computed to assess the relationships among MAUQ SUS, and the subscales of MAUQ to determine the construct validity [37,40,41]. These coefficients quantify the degree of association among these scores, with values close to 1 or -1 indicating a strong correlation. In addition, Kendall rank correlation coefficient was used to examine the relationship between the individual MAUQ items and the overall score. Furthermore, internal consistency was assessed to reflect on the interrelatedness of MAUQ items, indicating questionnaire reliability and validity. Cronbach α measured the internal consistency, with values between 0.7 and 0.8 considered acceptable and values around 0.9 considered excellent. High values signify strong item consistency and enhanced questionnaire validity [42-44].

Comprehensibility

At the end of the translated MAUQ survey, additional questions were added regarding the understandability of the translation ("Was the text in the above questionnaire easy to understand? Yes/No. If not, why not? Do you have any further feedback on the above questionnaire you would like to share with us?"). These questions were added to ensure that participants felt invited to share thoughts or feedback about how to improve the understandability of the translated survey. Any feedback entered in these free-text fields was exported from REDCap to Microsoft Excel and analyzed descriptively by LW and CA. In addition, content analysis was performed to understand the rates of occurrence of the identified themes in the free-text fields.

Results

Overview

After the exclusion of incomplete questionnaires, 133 questionnaires were included in the analysis. For the purpose of this validation study, completed MAUQs from both data collection time points were considered. Overall, 75.9% (101/133) of the participants returned the SUS questionnaire. Tables 1 and 2 show the characteristics of the participants.

XSL•FO RenderX

Table 1. Sociodemographic data of participants (N=133).

Characteristics	Participants, n (%)	
Sex		
Female	133 (100)	
Age group (y)		
<30	2 (1.5)	
30-40	16 (12)	
41-50	48 (36.1)	
51-60	42 (31.6)	
61-70	18 (13.5)	
71-80	7 (5.3)	
Education		
Academic degree	48 (36.1)	
High school education	17 (12.8)	
Lower or intermediate secondary school	62 (46.6)	
Prefer not to say	6 (4.5)	
Employment		
Employed	79 (59.4)	
Unemployed	25 (18.8)	
Studying or vocational training	1 (0.8)	
Retired	18 (13.5)	
Prefer not to say	10 (7.5)	

Table 2. Additional participant characteristics regarding smartphone and app use.

Characteristics	Values	
Length of smartphone ownership (y; N=133), mean (SD)	11 (4.96)	
Use of other mHealth ^a apps (N=133), n (%)		
Yes	44 (33.1)	
No	83 (62.4)	
Prefer not to say	6 (4.5)	
Use of wearables (N=133), n (%)		
Yes	50 (37.6)	
No	77 (57.9)	
Prefer not to say	6 (4.5)	
Frequency of Enable app use (n=101), n (%)		
Daily or several days a week	48 (47.5)	
Once a week	46 (45.5)	
Once a month or less	6 (5.9)	
Prefer not to say	1 (0.9)	

^amHealth: mobile health.

XSL•FO RenderX

Descriptive Statistics

In this study, descriptive statistics were evaluated for SUS and MAUQ to assess the usability of the Enable app. The

https://humanfactors.jmir.org/2023/1/e51090

participants' high level of perceived usability, as indicated by the SUS score with a mean of 88.3 (SD 9.9), noticeably outperformed the average score of 68. Similarly, MAUQ yielded a mean score of 85.89 (SD 11.45), further affirming the

favorable perception of usability. As questionnaire results were comparable at the 2 different data collection time points and to achieve a high validation sample, it was decided to combine the results of the 2 time points for the purpose of this validation study.

The normal distribution of the data was assessed using the Shapiro-Wilk test. The normalization analysis using the Shapiro-Wilk test indicated that none of the items in the questionnaire followed a normal distribution, as evidenced by the small P values (P<.001) obtained, indicating a low likelihood of these items conforming to normality (Table 3). In addition, a visual examination of the histograms further confirms this deviation from normality, as the Shapiro-Wilk test may sometimes overestimate the departure from a normal distribution. Furthermore, these findings suggest that nonparametric statistical tests should be used for further analysis. The results of the normalization analysis are shown in Table 4.

 Table 3. Means and SDs of the subscales of the mHealth App Usability Questionnaire.

Subscale	Score, mean (SD)
Ease of use	6.445 (0.749)
Interface and satisfaction	6.223 (0.847)
Usefulness	5.405 (1.144)

Table 4.	Results of the	normalization	analysis	using the	Shapiro-	Wilk test.
----------	----------------	---------------	----------	-----------	----------	------------

Item	Shapiro-Wilk test P value	Score, mean (SD)	
1	<.001	6.704 (0.773)	
2	<.001	6.76 (0.723)	
3	<.001	6.6 (0.823)	
4	<.001	5.912 (1.426)	
5	<.001	6.248 (1.175)	
6	<.001	6.24 (1.103)	
7	<.001	6.232 (1.086)	
8	<.001	5.784 (1.543)	
9	<.001	5.992 (1.292)	
10	<.001	6.544 (0.875)	
11	<.001	6.456 (1.096)	
12	<.001	6.312 (1.066)	
13	<.001	5.968 (1.373)	
14	<.001	5.472 (1.532)	
15	<.001	5.408 (1.498)	
16	<.001	5.32 (1.511)	
17	<.001	4.848 (1.778)	
18	<.001	5.416 (1.52)	

Significance Tests

The Wilcoxon rank sum test was conducted to assess the significance of differences in the scores of MAUQ and SUS. Significant differences were found in the mean scores between the overall scores of SUS and MAUQ. In addition, significant differences were also observed in the mean scores for the ease of use and usefulness subscales of MAUQ, whereas no significant difference was found for the interface and satisfaction subscales. We evaluated the discriminative ability of the questionnaire items; the Wilcoxon rank sum test was used to compare the highest-scoring and lowest-scoring quintiles. The analysis revealed significant differences in means for all items,

indicating their ability to discriminate between the different levels of usability.

Correlation Coefficient

This study investigated the association between the subscales and overall scores of MAUQ using Kendall rank correlation coefficient. The results revealed a moderate positive correlation between the ease of use subscale and the overall scores (r=0.56; P<.001), indicating that as the perceived ease of use increased, so did the overall scores. Moreover, a strong positive correlation was observed between the interface and satisfaction subscale and the overall scores (r=0.75; P<.001), suggesting that high levels of satisfaction with the interface were associated with high overall scores. In addition, a high positive correlation was

found between the usefulness subscale and the overall scores (r=0.83; P<.001), indicating that the great perceived usefulness of the app was linked to high overall scores. These findings support the construct validity of MAUQ and highlight the importance of these subscales in assessing the usability of the Enable app.

The total item correlation was also computed for the German version of MAUQ (G-MAUQ) using Kendall rank correlation

coefficient, with a predefined threshold value of 0.4. Our results showed that the correlation coefficients ranged from 0.39 to 0.68. These values indicate moderate to strong associations between the items and the overall score of G-MAUQ. This suggests that the items in the questionnaire collectively contribute to the measurement of the construct assessed by the questionnaire. Correlation coefficients of the overall score and subscales of MAUQ are shown in Table 5.

Table 5. Correlation coefficients of the overall score and subscales of the mHealth App Usability Questionnaire.

Subscale	Ease of use	Interface and satisfaction	Usefulness	Overall score
Ease of use	1	0.4176	0.4496	0.5649
Interface and satisfaction	0.4176	1	0.5879	0.7475
Usefulness	0.4496	0.5879	1	0.8295
Overall score	0.5649	0.7475	0.8295	1

Internal Consistency

The intersubscale internal consistency reliability of MAUQ was evaluated using Cronbach α coefficient (Table 6). The obtained Cronbach α value of .81 suggests satisfactory internal consistency reliability among the subscales. This indicates that the items within MAUQ consistently measure related aspects of usability. Furthermore, the internal consistency reliability

within each subscale was assessed using Cronbach α coefficient. The ease of use subscale demonstrated a value of .79, indicating good internal consistency. Similarly, the interface and satisfaction subscale exhibited a value of .85, and the usefulness subscale showed a value of .84, both indicating strong internal consistency within their respective subscales. These results suggest that the items within each subscale are measuring a similar construct consistently.

Table 6. Internal consistency reliability of the mHealth App Usability Questionnaire (MAUQ).

Subscales of MAUQ	Cronbach α
Ease of use	.7857
Interface and satisfaction	.8497
Usefulness	.8375
Overall score	.8102

Comprehensibility

In the translated MAUQ, 95.5% (127/133) of the participants answered the additional yes or no question about understandability, and 95.3% (121/127) replied that the survey was easy to understand. In total, 51 comments were obtained in the free-text fields. Many comments (23/25, 92%) referred to the ENABLE study itself or technical problems experienced in the Enable app. Thus, these comments were excluded for the purpose of this analysis. In the following step, the 55% (28/51) remaining comments were analyzed and categorized into 6 groups. Most comments (14/28, 50%) comprised short, positive statements, such as "good," "questionnaire was quick and easy," and "everything is comprehensible." Another common group of comments (5/28, 18%) covered the wish to add more free-text space in the questionnaire to enable participants to make suggestions for improvement. Some comments (3/28, 11%) described having difficulties with understanding the Likert scale or the impression that questions were very similar and hard to differentiate from each other (3/28, 11%). Overall, 7% (2/28) of the participants noted that they experienced difficulty with understanding the questionnaire owing to their low command of the German language, and 4% (1/28) of the participants

described that question 8 included a term that they could not understand ("The app adequately acknowledged...").

Discussion

Principal Findings

We conducted a translation and validation study of G-MAUQ in a cohort of 133 German-speaking patients with breast cancer. The determination of an appropriate sample size for questionnaire validation lacks universally prescribed guidelines. However, it is generally recommended to use a large sample size to achieve a high respondent-to-question ratio to enhance the statistical robustness of the analysis. In our study, we adhered to the recommendation of maintaining a ratio of at least 5 participants per statement (MAUQ items=18) to ensure an adequate sample size for the questionnaire validation [26-28]. Importantly, our achieved ratio surpassed this threshold, meeting the recommended criterion for a sufficient sample size.

Data were collected in the context of an RCT studying the use of an mHealth app as a support tool during the course of chemotherapy. In our validation study, we observed a positive correlation between the subscales and the overall score of G-MAUQ. However, our findings suggest that the discrepancy

in scores compared with the original validation study [12] could be attributed to the differences in participant characteristics. The previous study recruited participants primarily from the University of Pittsburgh, with a limitation being that approximately one-third of their participants were students. In contrast, our validation study included actual patients with breast cancer who were enrolled in the RCT of the ENABLE project. The contrasting health statuses of healthy participants in the previous study and patients with chronic illness in our study could potentially influence the obtained scores.

On the basis of our results from the statistical analysis, we observed that the correlations support the validity of G-MAUQ in capturing the intended concept and provide evidence of the items' alignment with the overall scoring of the instrument. Correspondingly, the high internal consistency reliability observed across the subscales of G-MAUQ strengthens the confidence in its ability to accurately measure usability in the context of mHealth apps. These findings support the reliability and validity of G-MAUQ as a tool for assessing the different dimensions of usability in this population. This is consistent with the findings from the Chinese and Malay version of MAUQ [15,16].

During the translation process, it was observed that certain words in German did not align perfectly with the original English word and its intended meaning. In addition, in German, some words can have multiple meanings depending on the context. This highlighted the influence of cultural factors on translation, similar to a previous study [16]. To ensure a precise and accurate understanding of the German words within the usability context, modifications were made to the wording. These adjustments aimed to clarify the purpose and meaning of the questionnaire items. For example, consider item 9 within the questionnaire, which incorporates the phrase, "social settings." The term, "social," in this context presents 2 potential translations in German: "soziales Umfeld," signifying the social environment encompassing friends, family, and even unfamiliar individuals within one's social sphere, and "gesellschaftliches Umfeld," which refers to the societal environment, including factors such as a person's upbringing, education, and care. These divergent interpretations of the term, "social," signify a notable variance in how it is perceived and understood. After careful deliberation, we opted for "gesellschatliches Umfeld" (societal environment) as it conveys a more comprehensive and contextually fitting interpretation, closely aligning with the intended meaning of "social settings" in item 9.

Similarly, in item 11, which states, "I would use the app again," the word "use" in the German language carries various connotations and interpretations. Initially, we translated "use" as "verwenden," implying reuse, such as using the app in a different context or situation. However, this translation did not precisely capture the intended sense of the English phrase with its specific context. Therefore, following a thorough examination and expert review by native speakers, we chose to use the literal translation of the term, "use" (nutzen), signifying the act of using the app once more in its original context. This highlighted the influence of cultural factors on translation, similar to the studies by Zhou et al [16]. Furthermore, the results of our qualitative assessment indicated that the translated MAUQ was easily comprehensible based on the modifications implemented.

Regarding the feedback about questionnaire understandability overall, and language. participants perceived the understandability to be adequate, and the questionnaire was easy to complete for most (121/127, 95.3%). Thus, we do not plan to make any further changes to the translation. However, participants provided 51 additional comments in the free-text fields, giving feedback about the overall study, study team members, and the app. In these comments, 18% (5/28) voiced the desire to have more space for individual feedback about the app in the usability questionnaire. Hence, we recommend future users of G-MAUQ to also provide a free-text field when administering the survey, prompting the participants to provide valuable, individual, additional feedback.

Although it may be tempting for researchers to develop new, study-specific questionnaires for their studies, using a validated questionnaire holds several advantages. First, developing a new questionnaire requires a lot of resources and takes time [45]. Second, owing to the extensive validation processes, the validity of established questionnaires is high, which makes the results more trustworthy. Another aspect is that results derived from validated questionnaires can be more easily compared with results from other studies on similar topics. This also applies to validated translations of existing questionnaires. Using these validated translations ensures the comparability of research findings across different cultural contexts and languages [46]. To the best of the authors' knowledge, there are 3 other validated translations of MAUQ available [15-18]. The availability of these validated translations will be helpful in conducting population-specific and methodologically sound studies of mHealth usability.

Strengths and Limitations

Our study followed a structured approach of WHO's Back-Translation Guidelines [32] to validate a self-developed German translation of MAUQ. Data were collected within a large research project, and a sufficient number of participants completed the survey, allowing for sound statistical analysis. In addition, individual feedback about the understandability and wording of the questionnaire was collected. This allowed us to assess the quality of the translation from the perspective of laypersons. To demonstrate the external validity of our findings, we also recommend that future studies should investigate whether the translated questionnaire can be used effectively in the context of other mHealth apps.

In an effort to include feedback from as many participants as possible, even those with low technical capabilities or who were experiencing difficulties with using the Enable app, we decided to collect data through mailed questionnaires. This approach was both time-consuming for the study team and could have introduced mistakes owing to the necessary manual data entry. Another effort that was made to increase the study sample was the combination of questionnaires from 2 different time points. This was only possible because the results from the 2 different time points did not differ in a noteworthy manner. However, this combination could have introduced a potential bias in our validation.

https://humanfactors.jmir.org/2023/1/e51090

xsl•F(

Owing to the study being conducted in the context of a large breast cancer trial, our study sample included only female participants. Although this is an important limitation to note, we do not consider the sex of the participants to play an influential role in the validity and understandability of the questionnaire. This was shown by previous studies, which concluded that there are no significant differences between female and male study participants regarding the perceived usability of a system [47-49].

However, 36.1% (48/133) of the participants in our sample held an academic degree. This is above average compared with the share of academics in the German population overall (24%) [50]. In addition, 1.5% (2/133) of the participants stated that they had difficulties in understanding the questionnaire owing to their low command of the German language. Future research projects should make additional efforts to include participants from these traditionally underrepresented groups (low educational backgrounds and nonnative speakers) in their samples.

Conclusions

We successfully validated G-MAUQ using a standardized approach in a cohort of 133 patients with breast cancer. Similar to the original version, G-MAUQ revealed good reliability and validity in this study. Our validation study demonstrated robust and satisfactory internal consistency reliability among the subscales of MAUQ, with a Cronbach α coefficient of .81, indicating strong and satisfactory reliability. In addition, we observed a significant positive correlation between the subscales and the overall score of MAUQ. These results indicate a high degree of internal consistency and support the construct validity of MAUQ. Hence, it can be used as a reliable tool to evaluate the usability of mHealth apps among German-speaking adults. The availability of G-MAUQ will help other researchers in conducting usability studies of mHealth apps in German-speaking cohorts and allow for international comparability of their results. Further research is recommended to study the validity of the translated questionnaire in other user groups and in other contexts for mHealth apps.

Acknowledgments

This study was funded by the Baden-Wuerttemberg (Germany) Ministry of Science, Research, and the Arts under the reference number, 42-04HV.MED (19)/15/1, as part of the project, Zentrum für innovative Versorgung (ZIV). In addition, for the publication fee, the authors acknowledge financial support from Deutsche Forschungsgemeinschaft within the funding program, Open Access Publikationskosten, and from Heidelberg University. The authors would like to thank all the patients in this study for their participation and valuable input. The authors would like to thank the ENABLE study team (Markus Wallwiener, Thomas Deutsch, Thao Vy-Le) for their support in the execution of and participant recruitment for this study. The authors would also like to thank Thomas Ganslandt for the support and initial planning of this study. The authors also thank Julia Müller and Laura Svensson for their support in the study's execution.

Authors' Contributions

PM, LW, and CA drafted and prepared the original manuscript. PM and LW developed the concept and design for the study. LW and CA were responsible for the data collection. PM and BCH conducted the statistical analysis, and FS supported it. FS contributed to the overall study implementation. All the authors provided substantial comments, contributed to the study, and approved the submitted version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Translations of the mHealth App Usability Questionnaire in German and English. [XLSX File (Microsoft Excel File), 12 KB-Multimedia Appendix 1]

Multimedia Appendix 2

German version of the mHealth App Usability Questionnaire. [PDF File (Adobe PDF File), 277 KB-Multimedia Appendix 2]

References

- mHealth apps market size, share and trends analysis report by type (fitness, medical), by region (North America, Europe, Asia Pacific, Latin America, Middle East and Africa), and segment forecasts, 2022-2030. Research and Markets. 2022 Jan. URL: <u>https://www.researchandmarkets.com/reports/4396364/mhealth-apps-market-size-share-and-trends</u> [accessed 2023-05-17]
- 2. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. World Health Organization. Switzerland: World Health Organization; 2011. URL: <u>https://iris.who.int/handle/10665/44607</u> [accessed 2023-05-17]

- Morse SS, Murugiah MK, Soh YC, Wong TW, Ming LC. Mobile health applications for pediatric care: review and comparison. Ther Innov Regul Sci 2018 May;52(3):383-391 [FREE Full text] [doi: 10.1177/2168479017725557] [Medline: 29714538]
- 4. Mechael P, Batavia H, Kaonga N, Searle S, Kwan A, Goldberger A, et al. Barriers and gaps affecting mHealth in low and middle income countries : policy white paper. The Earth Institute, Columbia University. 2010 May. URL: <u>http://www.globalproblems-globalsolutions-files.org/pdfs/mHealth_Barriers_White_Paper.pdf</u> [accessed 2023-05-17]
- Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996.
 Prümper J. Der benutzungsfragebogen ISONORM 9241/10: ergebnisse zur reliabilität und validität. In: Liskowsky R,
- Velichkovsky BM, Wünschmann W, editors. Software-Ergonomie '97. Berlin, Germany: Vieweg+Teubner Verlag; 1997.
 Hassenzahl M, Platz A, Burmester M, Lehner K. Hedonic and ergonomic quality aspects determine a software's appeal.
- Hassenzahl M, Platz A, Burnester M, Lenner K. Hedonic and ergonomic quality aspects determine a software's appeal. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems. 2000 Presented at: CHI00: Human Factors in Computing Systems; April 1-6, 2000; The Hague, The Netherlands [doi: <u>10.1145/332040.332432</u>]
- 8. Flavián C, Guinalíu M, Gurrea R. The role played by perceived usability, satisfaction and consumer trust on website loyalty. Inf Manag 2006 Jan;43(1):1-14 [FREE Full text] [doi: 10.1016/j.im.2005.01.002]
- 9. Kirakowski J, Corbett M. SUMI: the Software Usability Measurement Inventory. Br J Educ Technol 1993 Sep;24(3):210-212 [doi: 10.1111/j.1467-8535.1993.tb00076.x]
- Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. In: Proceeding of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society. 2008 Presented at: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society; November 20-21, 2008; Graz, Austria [doi: 10.1007/978-3-540-89350-9_6]
- Messner EM, Terhorst Y, Barke A, Baumeister H, Stoyanov S, Hides L, et al. The German version of the mobile app rating scale (MARS-G): development and validation study. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e14479 [FREE Full text] [doi: 10.2196/14479] [Medline: 32217504]
- 12. Zhou L, Bao J, Setiawan IM, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500 [FREE Full text] [doi: 10.2196/11500] [Medline: 30973342]
- 13. Hajesmaeel-Gohari S, Khordastan F, Fatehi F, Samzadeh H, Bahaadinbeigy K. The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. BMC Med Inform Decis Mak 2022 Jan 27;22(1):22 [FREE Full text] [doi: 10.1186/s12911-022-01764-2] [Medline: 35081953]
- 14. Alanzi TM. Users' satisfaction levels about mHealth applications in post-Covid-19 times in Saudi Arabia. PLoS One 2022 May 4;17(5):e0267002 [FREE Full text] [doi: 10.1371/journal.pone.0267002] [Medline: 35507590]
- 15. Mustafa N, Safii NS, Jaffar A, Sani NS, Mohamad MI, Abd Rahman AH, et al. Malay version of the mHealth app usability questionnaire (M-MAUQ): translation, adaptation, and validation study. JMIR Mhealth Uhealth 2021 Feb 04;9(2):e24457 [FREE Full text] [doi: 10.2196/24457] [Medline: 33538704]
- Zhao S, Cao Y, Cao H, Liu K, Lv X, Zhang J, et al. Chinese version of the mHealth app usability questionnaire: cross-cultural adaptation and validation. Front Psychol 2022 Feb 2;13:813309 [FREE Full text] [doi: 10.3389/fpsyg.2022.813309] [Medline: 35185732]
- 17. Quifer-Rada P, Aguilar-Camprubí L, Gómez-Sebastià I, Padró-Arocas A, Mena-Tudela D. Spanish version of the mHealth app usability questionnaire (MAUQ) and adaptation to breastfeeding support apps. Int J Med Inform 2023 Jun;174:105062 [FREE Full text] [doi: 10.1016/j.ijmedinf.2023.105062] [Medline: 37037124]
- Shan Y, Ji M, Xie W, Li R, Qian X, Zhang X, et al. Chinese version of the mobile health app usability questionnaire: translation, adaptation, and validation study. JMIR Form Res 2022 Jul 06;6(7):e37933 [FREE Full text] [doi: 10.2196/37933] [Medline: 35793132]
- Pape M, Färber T, Seiferth C, Roth T, Schroeder S, Wolstein J, et al. A tailored gender-sensitive mHealth weight loss intervention (I-GENDO): development and process evaluation. JMIR Form Res 2022 Oct 27;6(10):e38480 [FREE Full text] [doi: 10.2196/38480] [Medline: 36301614]
- 20. Schillings C, Meissner D, Erb B, Schultchen D, Bendig E, Pollatos O. A chatbot-based intervention with ELME to improve stress and health-related parameters in a stressed sample: study protocol of a randomised controlled trial. Front Digit Health 2023 Mar 1;5:1046202 [FREE Full text] [doi: 10.3389/fdgth.2023.1046202] [Medline: 36937250]
- 21. Mähs M, Pithan JS, Bergmann I, Gabrys L, Graf J, Hölzemann A, et al. Activity tracker-based intervention to increase physical activity in patients with type 2 diabetes and healthy individuals: study protocol for a randomized controlled trial. Trials 2022 Jul 30;23(1):617 [FREE Full text] [doi: 10.1186/s13063-022-06550-z] [Medline: 35907864]
- 22. Weber F, Müller C, Bahns C, Kopkow C, Färber F, Gellert P, et al. Smartphone-assisted training with education for patients with hip and/or knee osteoarthritis (SmArt-E): study protocol for a multicentre pragmatic randomized controlled trial. BMC Musculoskelet Disord 2023 Mar 23;24(1):221 [FREE Full text] [doi: 10.1186/s12891-023-06255-7] [Medline: 36959595]
- 23. Thielmann EA, Müller B, Guethlin C, Gerlach F, Müller A, Nohl-Deryk P, et al. Usability of a communication tool used by general practitioners and public health workers during the COVID-19 pandemic: a proof-of concept-study. Eur J Gen Pract 2021;27(1):358 [FREE Full text]

- 24. Meryk A, Kropshofer G, Hetzer B, Riedl D, Lehmann J, Rumpold G, et al. Implementation of daily patient-reported outcome measurements to support children with cancer. Pediatr Blood Cancer 2021 Nov;68(11):e29279 [FREE Full text] [doi: 10.1002/pbc.29279] [Medline: 34383360]
- 25. Anders C, Weinert L, Moorthy P, Svensson L, Knaup P, Heinze O. Therapieunterstützung beim mammakarzinom: eye-tracking im rahmen einer mixed-methods-evaluation zur untersuchung der usability und feasibility einer smartphone-app. German Medical Science. 2022 Sep 30. URL: <u>https://www.egms.de/static/en/meetings/dkvf2022/22dkvf090.shtml</u> [accessed 2023-07-04]
- 26. Gorsuch RL. Factor Analysis. New York, NY: Psychology Press; 1983.
- 27. Hatcher L, O'Rourke N. A Step-by-Step Approach to Using SAS for Factor Analysis and Structural Equation Modeling, Second Edition. Cary, NC: SAS Institute; Mar 23, 2013.
- 28. Suhr DD. Exploratory Or Confirmatory Factor Analysis?. Cary, NC: SAS Institute; 2006.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3186-3191 [FREE Full text] [doi: <u>10.1097/00007632-200012150-00014</u>] [Medline: <u>11124735</u>]
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993 Dec;46(12):1417-1432 [FREE Full text] [doi: 10.1016/0895-4356(93)90142-n] [Medline: 8263569]
- Sperber AD. Translation and validation of study instruments for cross-cultural research. Gastroenterology 2004 Jan;126(1 Suppl 1):S124-S128 [FREE Full text] [doi: 10.1053/j.gastro.2003.10.016] [Medline: 14978648]
- 32. Process of translation and adaptation of instruments. World Health Organization. 2016. URL: <u>https://www.who.int/</u> <u>substance_abuse/research_tools/translation/en/</u> [accessed 2023-05-24]
- Acquadro C, Conway K, Hareendran A, Aaronson N, European Regulatory IssuesQuality of Life Assessment (ERIQA) Group. Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. Value Health 2008;11(3):509-521 [FREE Full text] [doi: 10.1111/j.1524-4733.2007.00292.x] [Medline: 18179659]
- 34. Harborth D, Pape S. German translation of the unified theory of acceptance and use of technology 2 (UTAUT2) questionnaire. SSRN J 2018 Mar 28:1-12 [doi: 10.2139/ssrn.3147708]
- 35. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. J Eval Clin Pract 2011 Apr;17(2):268-274 [FREE Full text] [doi: 10.1111/j.1365-2753.2010.01434.x] [Medline: 20874835]
- 36. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
- 37. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research. Amsterdam, The Netherlands: Elsevier Science; Jul 12, 2016.
- 38. Elliott DB, Bullimore MA. Assessing the reliability, discriminative ability, and validity of disability glare tests. Invest Ophthalmol Vis Sci 1993 Jan;34(1):108-119 [Medline: <u>8425818</u>]
- Rachakonda T, Jeffe DB, Shin JJ, Mankarious L, Fanning RJ, Lesperance MM, et al. Validity, discriminative ability, and reliability of the hearing-related quality of life questionnaire for adolescents. Laryngoscope 2014 Feb;124(2):570-578 [FREE Full text] [doi: 10.1002/lary.24336] [Medline: 23900836]
- 40. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
- 41. Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. Malawi Med J 2012 Sep;24(3):69-71 [FREE Full text] [Medline: 23638278]
- 42. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951 Sep;16(3):297-334 [FREE Full text] [doi: 10.1007/bf02310555]
- 43. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55 [FREE Full text] [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
- 44. Landauer TK. Behavioral research methods in human-computer interaction. In: Handbook of Human-Computer Interaction (Second Edition). Amsterdam, The Netherlands: Elsevier Science; 1997.
- 45. Mohamad Marzuki MF, Yaacob NA, Yaacob NM. Translation, cross-cultural adaptation, and validation of the Malay version of the system usability scale questionnaire for the assessment of mobile apps. JMIR Hum Factors 2018 May 14;5(2):e10308 [FREE Full text] [doi: 10.2196/10308] [Medline: 29759955]
- 46. Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. Saudi J Anaesth 2017 May;11(Suppl 1):S80-S89 [FREE Full text] [doi: 10.4103/sja.SJA 203 17] [Medline: 28616007]
- 47. Lim KC, Selamat A, Ghani NA, Zabil MH, Alias RA, Puteh F. Pre-processing of gender-based comparative usability performance data in Mobile Augmented Reality English language teaching. In: Proceeding of the IEEE Conference on e-Learning, e-Management and e-Services (IC3e). 2017 Presented at: IEEE Conference on e-Learning, e-Management and e-Services (IC3e); November 16-17, 2017; Miri, Malaysia [doi: 10.1109/ic3e.2017.8409246]

- 48. Ibili E, Billinghurst M. Assessing the relationship between cognitive load and the usability of a mobile augmented reality tutorial system: a study of gender effects. Int J Assess Tool Educ 2019 Jul;6(3):378-395 [doi: 10.21449/ijate.594749]
- 49. Weichbroth P. An empirical study on the impact of gender on mobile applications usability. IEEE Access 2022 Nov 4;10:119419-119436 [doi: 10.1109/access.2022.3219421]
- 50. Bocksch R. Deutschland hat vergleichsweise wenige Uni-Absolvent:innen. Statista. 2021 Oct 8. URL: <u>https://de.statista.com/infografik/25934/anteil-der-hochschulabsolventinnen-in-ausgewaehlten-laendern/</u> <u>#:~:text=Hochschule&text=Nur%20etwa%2024%20Prozent%20der,eine%20Berufsbildung%20durchlaufen%20zu%20haben</u> [accessed 2023-05-24]

Abbreviations

G-MAUQ: German version of mHealth App Usability Questionnaire
ISO: International Organization for Standardization
MAUQ: mHealth App Usability Questionnaire
mHealth: mobile health
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
SUS: System Usability Scale
WHO: World Health Organization

Edited by C Jacob; submitted 20.07.23; peer-reviewed by K Wenderott, K Becerro de Bengoa Losa, D Chrimes; comments to author 20.08.23; revised version received 01.09.23; accepted 02.09.23; published 01.11.23

<u>Please cite as:</u> Moorthy P, Weinert L, Harms BC, Anders C, Siegel F German Version of the mHealth App Usability Questionnaire in a Cohort of Patients With Cancer: Translation and Validation Study JMIR Hum Factors 2023;10:e51090 URL: <u>https://humanfactors.jmir.org/2023/1/e51090</u> doi: <u>10.2196/51090</u> PMID: <u>37910144</u>

©Preetha Moorthy, Lina Weinert, Bendix Christian Harms, Carolin Anders, Fabian Siegel. Originally published in JMIR Human Factors (https://humanfactors.jmir.org), 01.11.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on https://humanfactors.jmir.org, as well as this copyright and license information must be included.

