ALCANCE

DOI: https://doi.org/10.14210/alcance.v31n2(mai/ago)52-66
ISSN - 1983-716X

PERCEPTION OF PRIVACY AND RELIABILITY IN THE USE OF M-BANKING SERVICES: DEVELOPMENT OF A MEASUREMENT SCALE USING MULTIDIMENSIONAL ITEM RESPONSE THEORY

PERCEPÇÃO DE PRIVACIDADE E CONFIABILIDADE NO USO DOS SERVIÇOS DE M-BANKING: DESENVOLVIMENTO DE UMA ESCALA DE MENSURAÇÃO UTILIZANDO A TEORIA DA RESPOSTA DO ITEM MULTIDIMENSIONAL

PERCEPCIÓN DE PRIVACIDAD Y CONFIABILIDAD EN EL USO DE SERVICIOS DE BANCA MÓVIL: DESARROLLO DE UNA ESCALA DE MEDICIÓN QUE UTILIZA LA TEORÍA DE RESPUESTA AL ÍTEM MULTIDIMENSIONAL

ABSTRACT

Objective: To measure the perception of privacy and data security and reliability in m-banking.

Design/methodology/approach: A set of 26 items was developed and applied to a sample of 292 Brazilian m-banking users. Data were analyzed using the multidimensional item response theory and the compensatory gradual response model. This model allows us to analyze whether there are compensatory aspects between elements of different dimensions of a multidimensional construct, as is the case of perceived quality. This is a quantitative and cross-cutting study of individual users in the Brazilian context. Results: As a result, it was identified that m-banking users have a high degree of perceived reliability when using the banking application. Usage-oriented issues, such as service provided quickly/immediately, information provided without errors and delivery of the functionalities promised by the applications, show that users' perceive the applications' ability to perform the service in a reliable and satisfactory manner is perceived. However, the degree of perception of privacy/security of personal information is still low, indicating that it needs to be treated as a priority by financial institutions.

Originality/value: The article presents the application of a statistical technique that has been gaining ground in publications in several areas but has still not been explored in the area of administration – multidimensional item response theory. In the article, we present the potential of this technique for analyzing multidimensional latent traits, as is the case of perceived quality in the context of m-banking.

Keywords: M-banking. Perception. Privacy. Reliability.

Monique Wagner de Souza
PhD
Faculdade CESUSC - Brazil
moniquewsouza@gmail.com

Rafael Tezza
PhD
Universidade do Estado de Santa Catarina – Brazil
rafael.tezza@udesc.br

Submitted in: 05/30/2024
Accepted in: 11/29/2024

Special Issue: Quantitative Methods for Decision-Making in Organizations: Contributions by Miguel Angel Verdinelli

How to cite: de Souza, M. W, & Tezza, R. (2024). Percepção de privacidade e confiabilidade no uso dos serviços de M-banking: desenvolvimento de uma escala de mensuração utilizando a teoria da resposta do item multidimensional. *Revista Alcance (online)*, 31(2), 52-66. Doi: https://doi.org/10.14210/alcance.v31n2(mai/jun).52-66

OPEN ACCESS

INDEXED BY redalyc.org

 DOI: https://doi.org/10.14210/alcance.v31n2(mai/ago).52-66

RESUMO

Objetivo: Mensurar a percepção de privacidade/segurança dos dados pessoais e a confiabilidade no uso do m-banking.

Design / metodologia / abordagem: Foi desenvolvido um conjunto de 26 itens e aplicado a uma amostra de 292 usuários brasileiros de m-banking. Os dados foram analisados por meio da Teoria da Resposta ao Item Multidimensional, utilizando o modelo compensatório de Resposta Gradual. Esse modelo permite analisar a existência, ou não, de aspectos compensatórios entre elementos de diferentes dimensões de um constructo multidimensional, como é o caso da qualidade percebida. Trata-se de um estudo quantitativo e transversal no contexto brasileiro e de usuários pessoa física.

Resultados: Como resultado, identificou-se que os usuários de m-banking possuem um elevado grau de percepção de confiabilidade no uso do aplicativo bancário. Questões voltadas ao uso, como serviço fornecido de maneira rápida/imediata, informações prestadas sem erros e entrega das funcionalidades prometidas pelos aplicativos, mostram que a capacidade dos aplicativos de executar o serviço de maneira confiável e satisfatória é percebida pelos usuários. Contudo, o grau de percepção de privacidade/segurança das informações pessoais ainda é baixo, indicando que precisam ser tratadas como prioridade pelas instituições financeiras.

Originalidade / valor: O artigo apresenta a aplicação de uma técnica estatística que vem ganhando espaço nas publicações em diversas áreas, mas ainda pouco explorada na área de administração – Teoria da Resposta ao Item Multidimensional. No artigo apresentamos a potencialidade desta técnica para análise de traços latentes multidimensionais, como é o caso da qualidade percebida no contexto de m-banking.

Palavras-chave: M-banking. Percepção. Privacidade. Confiabilidade.

RESUMEN

Objetivo: Medir la percepción de privacidad/seguridad de los datos personales y la confiabilidad en el uso de la banca móvil.

Diseño/metodología/enfoque: Se desarrolló y aplicó un conjunto de 26 ítems a una muestra de 292 Usuarios brasileños de banca móvil. Los datos fueron analizados utilizando la Teoría de Respuesta al Ítem Multidimensional, utilizando el modelo compensatorio de Respuesta Gradual. Este modelo nos permite analizar la existencia, o no, de aspectos compensatorios entre elementos de diferentes dimensiones de un constructo multidimensional, como es el caso de la calidad percibida. Se trata de un estudio cuantitativo y transversal en el contexto brasileño y de usuarios individuales.

Recomendaciones: Como resultado, se identificó que los usuarios de banca móvil tienen un alto grado de confiabilidad percibida al utilizar la aplicación bancaria. Las cuestiones orientadas al uso, como el servicio proporcionado rápida/inmediatamente, la información proporcionada sin errores y la entrega de las funcionalidades prometidas por las aplicaciones, muestran que los usuarios perciben la capacidad de las aplicaciones para realizar el servicio de manera confiable y satisfactoria. Sin embargo, el grado de percepción de privacidad/seguridad de la información personal sigue siendo bajo, lo que indica que las instituciones financieras deben tratarla como una prioridad.

Originalidad / valor: El artículo presenta la aplicación de una técnica estadística que ha ido ganando terreno en publicaciones de diversas áreas, pero aún es poco explorada en el área de la administración – Teoría de Respuesta al Ítem Multidimensional. En el artículo presentamos el potencial de esta técnica para analizar rasgos latentes multidimensionales, como es el caso de la calidad percibida en el contexto de la banca móvil.

Palabras clave: Banca móvil. Percepción. Privacidad. Fiabilidad.

INTRODUCTION

The methods of service provision have evolved due to advancements in information and communication technology and the proliferation of mobile internet. Consequently, organizations have adopted various innovative channels to rea-

 DOI: https://doi.org/10.14210/alcance.v31n2(mai/ago).52-66

ch their consumers (Jebarajakirthy & Shankar, 2021). High investments in information technology have enabled banking institutions to adapt to this new reality, remain competitive, and address challenges posed by new technology companies entering the market to maximize service offerings. These companies, referred to as "financial technology" or fintechs, represent a combination of financial services and information technology (Cantú & Ulloa, 2020).

Measures taken to combat the COVID-19 pandemic led to the temporary closure of many financial establishments, rendering it impossible to provide banking services at physical branches. As a result, digital channels became more prevalent. According to a survey conducted by the Brazilian Federation of Banks (Febraban, 2023), nearly 8 out of 10 banking transactions are now carried out using banking apps or online platforms. Information technology has facilitated digital services, enabling financial institutions to reach locations that were previously difficult to access, thereby expanding their customer base and providing individuals with easier access to banking services. The adoption of mobile banking (m-banking) has significantly impacted the volume of banking transactions in Brazil. In 2014, mobile banking accounted for 10% of transactions, representing BRL 4.7 billion. By 2022, this share had risen to 66%, meaning that more than half of all banking transactions in the country are now conducted via m-banking (Febraban, 2023).

As the customer base grows, so does the need to meet customer expectations by ensuring quality service amid rapid transformations. In 2020, banking institutions allocated more than BRL 25 million to technology investments (Febraban, 2020). A significant portion of this investment has been directed toward analytics and big data, reflecting the strategic importance of data and the implications of Brazil's General Data Protection Law (LGPD), which came into effect in September 2020. This regulation emphasizes the need for heightened attention to user data privacy and security. In this context, understanding aspects related to quality, privacy, reliability, and perceived security is essential for financial institutions to deliver quality services from the customer's perspective.

The user's perception of the quality of m-banking applications is a latent variable – difficult to measure directly – but can be assessed through customer surveys using item response theory (IRT). IRT is capable of measuring latent traits through a set of items and creating a comparative scale.

Although m-banking is an increasingly important platform for delivering banking services, this topic still needs to be better explored (Shankar, Datta, & Jebarajakirthy, 2019). Few studies have investigated service quality within this context, and none have developed a specific scale to measure dimensions such as reliability, privacy, and security (Arcand, Promtep, Brum, & Rajaobelina, 2017; Jun & Palacius, 2016).

This gap in the literature regarding the measurement of m-banking service quality is also highlighted by Shankar et al. (2019), who note that the development of a scale in this area is still in its early stages. The authors emphasize the importance of further investigation into the specificities of mobile technologies. Thus, this study aims to measure aspects of privacy and reliability in m-banking using IRT. An instrument was developed and administered to a sample of 292 Brazilian m-banking users.

This article is structured into five sections. Following this introduction, the literature review explores perceived quality in the context of m-banking and the foundations of IRT. The methodological procedures are then described, followed by the presentation of empirical results. Finally, the article concludes with a discussion of the findings.

THEORETICAL BACKGROUND

Advances in information technology and internet infrastructure have provided electronic channels and platforms to deliver services effectively and efficiently (Shankar et al., 2019). E-service, which involves offering services through the internet, leverages information and communication technologies to facilitate transactions across various domains.

The quality of e-service differs from the quality perceived in physical environments due to specific factors such as system availability, ubiquity, and interactivity, which may vary depending on the context. In the financial sector, mobile technology has enabled banks to offer more convenient, accessible, interactive, and high-value services without restrictions in terms of time or location. In today's competitive financial landscape, banks must prioritize excellent customer service and tailor solutions to meet the needs of both existing and potential clients (Legass & Durmus, 2024). M-banking has emerged as one of the most promising services, characterized by high ubiquity and independence from physical locations (Arcand et al., 2017; Koenig-Lewis, Palmer, & Moll, 2010).

Given its specific characteristics, m-banking enables banking institutions to reduce transaction costs (Legass & Durmus, 2024) while expanding their customer base. This dynamic reinforces the importance of understanding how customers perceive the quality of the services provided. According to Parasuraman, Zeithaml, and Malhotra (2005), organizations must proactively understand customer demands to ensure satisfaction.

Lin (2013) defines banking service quality as an overall judgment of the quality and excellence of the content delivered in the m-banking context. However, there is no consensus on the quality dimensions specific to mobile banking, with variations across studies.

Shankar et al. (2019) examined the measurement of m-banking service quality using three generic scales (SERVQUAL, SERVPERF, and E-S-QUAL) to determine which is most suitable for this context. The SERVPERF and SERVQUAL instruments included tangibility, reliability, promptness, security, and empathy, while E-S-QUAL focused on efficiency, system availability, compliance, and privacy. None of these scales were deemed fully adequate for capturing the specificities of m-banking. Issues such as mobility and the risk of financial loss heighten concerns about privacy and security, as remote connectivity and potential device loss or theft increase perceived risks in m-banking (Hanafizadeh, Behboudi, Koshksaray, & Tabar, 2014; Arcand et al., 2017). However, reliability, security, empathy, efficiency, compliance, and privacy are crucial dimensions for explaining service quality in m-banking (Shankar et al., 2019). Information privacy is related to the permissions granted to mobile applications to access data on devices, such as location, messages, files, and usage data (Albashrawi & Motiwalla, 2019). According to Kar (2021, p. 1348), "Such privacy concerns could be related to access and use of sensitive personal information within or outside the boundary of the firm for unintended usage."

Shankar et al. (2020) found that previously studied quality dimensions overlooked the unique mobile context of financial transactions and the distinct characteristics of these services, which are not appropriate for assessing m-banking service quality. The authors adopted qualitative methods, including interviews, focus groups, critical incident techniques, and netnography, to assess key dimensions of these services' quality perception. They identified security/privacy, user support, interactivity, efficiency, and content as most relevant for evaluating quality perceptions in m-banking.

Although m-banking studies have explored a variety of quality dimensions, some of them are frequently highlighted as critical, such as privacy and security (Sagib & Zapan, 2014; Nisha, 2016; Stefani, Azevedo, & Duduchi, 2019; Arcand et al., 2017; Shankar et al., 2019). This research aligns with studies that emphasize these dimensions (Parasuraman et al., 2005; Sagib & Zapan, 2014; Nisha, 2016; Stefani, Azevedo, & Duduchi, 2019; Arcand et al., 2017; Shankar et al., 2019).

Security and privacy are particularly crucial in online and mobile financial services (Arcand et al., 2017; Shankar et al., 2019; Legass & Durmus, 2024). The aspect of mobility increases security/privacy risks due to remote connectivity and the potential loss or theft of devices (Hanafizadeh et al., 2014; Arcand et al., 2017).

Transaction security is also considered a significant antecedent of satisfaction in online/electronic environments (Szymanski & Hise, 2000), especially in financial services (Liao & Cheung, 2008). Sagib and Zapan (2014) define security and privacy as the degree to which m-banking services are secure and protect customer information. For Parasuraman et al. (2005)

DOI: https://doi.org/10.14210/alcance.v31n2(mai/ago).52-66

and Huang, Lin, and Fan (2015), these dimensions reflect the degree to which customers believe that applications are protected against intrusions and safeguard personal information. This study defines the security/privacy construct as the degree to which m-banking applications are secure for financial transactions and protect users' personal information without sharing it. This definition builds on studies by Parasuraman et al. (2005), Sagib and Zapan (2014), and Huang et al. (2015).

For this research, the reliability dimension refers to the m-banking application's ability to accurately and timely perform the services it was designed to deliver. Sagib and Zapan (2014) view reliability as the likelihood of services delivering the expected results appropriately and consistently. Similarly, Huang et al. (2015) highlight the importance of reliable and accurate service delivery. Wolfinbarger and Gilly (2003) identify reliability as the most significant predictor of customer satisfaction and service quality, as well as a key driver of repurchase intentions. Thus, the security/privacy and reliability dimensions form the foundation of this study, as illustrated in Table 1.

Table 1

Quality dimensions analyzed and definitions

Quality dimensions	Definitions
Privacy/security	The extent to which the m-banking application is secure for conducting financial transactions while safeguarding users' personal information.
Reliability	The capability of the m-banking application to reliably perform services and fulfill promised functions within the expected timeframe.

Source: Elaborated by the authors based on Parasuraman et al. (2005), Sagib and Zapan (2014), and Huang et al. (2015).

Item Response Theory

The characteristics of individuals that cannot be observed directly but must be inferred through secondary variables related to them are known as latent traits (Andrade, Tavares, & Valle, 2000). Examples of latent traits include proficiency in a given field, attitudes, perceived risk, and quality, among others.

Item response theory (IRT) introduces a new approach to studying psychometric variables. It allows for comparisons between populations subjected to tests with some common items or even between individuals within the same population who have taken different tests (Andrade, Tavares, & Valle, 2000). This comparison is possible because one of the main features of IRT is its focus on analyzing each item individually, ensuring that the response to a specific item is independent of responses to other items (Reckase, 2009). Consequently, conclusions are drawn based on each item's characteristics rather than exclusively on the overall test or research instrument (Araujo, Andrade, & Bortolotti, 2009).

IRT is a set of mathematical models designed to measure latent traits using a set of items, thereby creating a scale that compares the respondent's latent trait level with the difficulty of an item (Hambleton, 2000; Embretson & Reise, 2000). The relationship between the probability of an individual providing a particular response to an item as a function of the item's parameters, along with the respondent's ability is expressed in a dynamic where the higher the ability, the greater the likelihood of correctly answering the item (Andrade et al., 2000). The proficiency of individuals and the difficulty parameters of items are estimated on a standardized latent trait scale with a mean of zero and a standard deviation of 1. Higher scores on this scale indicate a greater degree of the measured latent trait, while lower scores signify a lesser degree (Ayala, 2009). Therefore, the latent trait scale measures an individual's proficiency, with the score achieved representing their level in the studied latent trait.

Using IRT requires defining the appropriate model, which involves considering three factors: the nature of the items (dichotomous or non-dichotomous), the number of latent traits being measured, and the number of populations involved (single or multiple) (Andrade, Tavares, & Valle, 2000). Costa (2011) notes that IRT has been recognized as the most robust theory for achieving reliable measurement results.

 DOI: https://doi.org/10.14210/alcance.v31n2(mai/ago).52-66

As a statistical framework, IRT encompasses models with varying specificities and applications. For instance, one-dimensional models, suitable for data explained by a single dimension, become inadequate for data requiring a multidimensional structure (involving two or more dimensions). In such cases, models from multidimensional item response theory (MIRT) are necessary.

MULTIDIMENSIONAL ITEM RESPONSE THEORY MODELS

MIRT can be considered a special case of multivariate statistical analysis, particularly factor analysis, or an extension of one-dimensional IRT, depending on the study's objectives and the data final structure (Reckase, 2009). According to Reckase (1997), the primary distinction between factor analysis and IRT lies in their interpretation. Exploratory factor analysis has traditionally been employed to identify the smallest possible number of hypothetical variables or factors that explain the relationships among many empirical variables. Essentially, exploratory factor analysis is a data reduction technique (Reckase, 1997), whereas IRT focuses on the interaction between respondents and items.

MIRT consists of various models, and the choice of a specific model depends on the nature of the latent dimensions, the items, and their associations within the test. In addition to these factors, it is crucial to understand the multidimensional structures involved in a test. This includes the relationships between latent dimensions and items (multidimensionality among items and within each item) and the relationship between latent dimensions and respondents' abilities, which can involve compensatory or non-compensatory interactions.

A key aspect of MIRT is understanding the interaction between latent dimensions and respondents' abilities, specifically the compensatory or non-compensatory nature of these interactions. These models are defined by how the coordinates of the latent trait vector interact with item characteristics to determine the probability of a given response to the item (Reckase, 2009; Tezza, Bornia, & Spenassatto, 2016). A model is classified as compensatory when "the probability of correctly answering an item is maintained or increases, even if one skill is low, as long as another skill compensates with a higher value" (Zaffalon et al., 2020, p. 782). In other words, a lower value in one dimension is offset by a higher value in another, hence the term compensatory. Conversely, a non-compensatory model does not allow a high level of ability in one dimension to compensate for a low level in another (Nojosa, 2002).

Reckase (2009) describes the compensatory model as an extension of the two-parameter logistic model (M2PL), represented by Equation 1:

$$P\left(u_i = 1 \middle| \theta_j\right) = \frac{e^{aik\theta jk+di}}{1 + e^{aik\theta jk+di'}} \qquad (1)$$

Where u_i is the response to item i;

aik is the vector of the discrimination parameter of item i in dimension k;

θjk is the vector of the latent trait of individual j in dimension k;

di is a scalar of the difficulty parameter of item i.

The exponent e of Equation 1 is described by Equation 2.

$$a_{i1}\,\theta_{j1} + a_{i2}\,\theta_{j2} + \cdots + a_{im}\,\theta_{jm} + d_i \quad (2)$$

The exponent in Equation 1 is a linear function comprising elements of θ and parameter d as the intercept and the elements of vector a as the discrimination parameters. This linear function defines a straight line in k-dimensional space, producing equiprobability lines with infinite linear combinations that yield the same exponent, thereby resulting in the same probability of success. These diverse combinations give the model its compensatory property.

For each item in the model, a discrimination parameter (a) and a skill parameter (θ) are estimated for each dimension. However, only one difficulty parameter is estimated, encompassing all dimensions. These parameters generate vectors to be analyzed. A noteworthy aspect is

DOI: https://doi.org/10.14210/alcance.v31n2(mai/ago).52-66

that the multidimensional space can undergo rotations to align the θ axis with points of greater significance in the space. These rotations may or may not preserve the initial covariance structure of the θ dimensions. The model can be rotated orthogonally or obliquely, aiming to achieve the best statistical fit with the theoretical framework.

The vector aik defines the discrimination parameters of the model, with its elements corresponding to the slope of the item response surface in the direction of the respective θ axis (Reckase, 2009). This set of vectors can be represented by a general vector, which is the sum of all individual vectors, referred to as the multidimensional discrimination of item i (MDISC). Equation 3 expresses this parameter (Reckase, 1985).

$$A_i \sqrt{\sum_{k=1}^{m} a_{ik}^2} \qquad (3)$$

The d_i parameter of the model refers to the item's difficulty. The interpretation of this parameter d is not the same as the b parameter of one-dimensional models (Reckase, 2009). To facilitate interpretation, it is possible to transform d into a function of b, as represented by Equation 4.

$$B_i = \frac{-d_i}{\sqrt{\sum_{k=1}^{m} a_{ik}^2}} \qquad (4)$$

The B_i measure, represented by Equation 4, indicates an item's difficulty (MDIFF) and corresponds to the vector's location in space. A positive B_i value reflects more difficult items, while a negative B_i value denotes easier items. The item's difficulty level is determined by the distance from the origin to the point of the steepest slope of the vector A_i.

The angular direction of each item relative to the k-dimensional axis illustrates the multidimensional location of the item in the plane and is expressed in Equation 5 (Reckase, 1986). The angle formed by the line connecting the origin of the space to the point of the steepest slope with the k-axis is represented as αik.

$$\cos \alpha_{ik} \frac{a_{ik}}{\sqrt{\sum_{k=1}^{m} a_{ik}^2}} \qquad (5)$$

Most polytomous multidimensional models are based on the compensatory model. Given the construct examined in this research (quality in the use of m-banking from the perspective of privacy and reliability) and the categorization of the polytomous responses from the data collection instrument, this research applied Samejima's (1974) multidimensional graded response model. This model is specifically designed for items with polytomous responses and accounts for the possibility of multidimensionality among them.

METHOD

This research is an exploratory and descriptive study employing the survey method. The target population consisted of Brazilian individuals who use m-banking applications, regardless of whether they hold accounts with traditional banks or fintechs.

A sample of 292 respondents was analyzed, with data collected during the first semester of 2021. The total sample was randomly split into two groups for analysis: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). For the EFA, the polychoric correlation matrix was adopted to account for the nature of the data, and the analysis was performed using the psych package in the R statistical software (R Core Team, 2023). The CFA employed the MLR (maximum likelihood with robust standard errors) estimation method to address data non-normality, and the analysis was conducted with the R lavaan package. The item response theory (IRT) analysis used the R mirt package.

The data collection instrument was structured based on the frameworks provided by Shankar et al. (2019), Sagib, Stefani, Azevedo, and Duduchi (2019), Arcand et al. (2016), Huang et al. (2015), Chen (2012), and Lin (2012), alongside the conceptual model outlined in Table 1.

 DOI: https://doi.org/10.14210/alcance.v31n2(mai/ago).52-66

Respondents evaluated statements related to the quality dimensions "privacy/security" and "reliability" using a Likert scale ranging from "1 – I completely disagree" to "5 – I completely agree." The instrument included items on banking habits. The survey was distributed electronically through social media, with additional dissemination facilitated by university departments offering undergraduate and graduate programs across various fields in different Brazilian states, which shared the survey with their students.

RESULTS AND DISCUSSIONS

The sample comprised 292 respondents, of whom 55% were female and 45% were male. Additionally, 71.6% of the participants were between 21 and 40 years old.

An interesting aspect of the sample relates to the smartphone operating systems used by the respondents. Given the study's focus on information systems, differences in usability and security features across operating systems could influence users' perceptions of system characteristics. Among the respondents, 65.7% used Android devices, while 34% used iOS. Notably, 0.3% of participants did not know which operating system they used.

EXPLORATORY FACTOR ANALYSIS

Before conducting the IRT analysis, Exploratory Factor Analysis (EFA) was performed on the 26 items of the research instrument. Based on the analysis of factor loadings and communalities, two items were excluded. The remaining items are grouped into two distinct dimensions: the first, labeled "privacy and data security," and the second comprising items representing "reliability." Table 2 displays the rotated component matrix.

Table 2

Factor loadings of items

Items	Factor 1	Factor 2	Commonality
1 - M-banking applications protect my banking transactions	0,77		0,67
2 - M-banking applications protect my personal information (document numbers, telephone, address)	0,64		0,46
3 - M-banking applications do not share my personal information (document numbers, telephone, address) with other companies and/or applications	0,59		0,36
4 - M-banking applications protect my bank card information (credit/debit)	0,72		0,60
5 - When I use m-banking, the privacy of my personal information is guaranteed by my cell phone's operating system (Android/iOS)	0,42		0,23
6 - The privacy of my personal information is guaranteed by the bank itself when I use the m-banking application	0,72		0,57
7 - I believe that the m-banking application provides sufficient security for my personal and confidential information to carry out banking transactions	0,86		0,77
8 - I have complete trust in the m-banking application provided by my bank	0,90		0,85
9 - I feel safe when carrying out transactions through the m-banking application	0,77		0,74
10 - I feel safe when providing confidential information (such as passwords) when using the m-banking application	0,73		0,67
11 - In general, I feel safe when using the m-banking application	0,83		0,82
12 - I feel comfortable when making transactions via m-banking applications	0,76		0,76
13 - If m-banking applications promise to complete a transaction within a certain period of time, they deliver on their promise		0,64	0,56
14 - When problems arise during a transaction, m-banking applications provide adequate information to resolve the issue without the need to go to the branch		0,61	0,39
15 - M-banking applications are reliable	0,79		0,84
16 - M-banking applications should provide their services quickly/immediately		0,57	0,53
17 - I usually do not have problems using the m-banking application, as it always completes the transactions performed		0,61	0,54
18 - I believe that m-banking applications execute transactions within a period of time that I consider adequate		0,80	0,76
19 - I think that m-banking applications present error-free banking information (e.g. balances/statements)		0,53	0,45
20 - M-banking apps are honest about their interest rate offers for loans, credit cards, etc		0,60	0,47
21 - If my transaction is not completed, my m-banking app provides quick answers to the problem		0,75	0,58
22 - If an error occurs in a transaction, I can resolve it quickly through the m-banking app itself		0,75	0,57
23 - I feel comfortable using the m-banking app because it provides the services exactly as promised		0,80	0,83
24 - When I make a transaction through the m-banking app, it happens quickly		0,72	0,79

The KMO measure of sampling adequacy exceeded the recommended threshold of 0.70, achieving a value of 0.92. Regarding the factors identified, the first group of items (1 to 12 and item 15) forms a dimension labeled "Privacy and Data Security." This characteristic is considered highly relevant for online and mobile financial institutions (Arcand et al., 2017; Shankar et al., 2019; Çelik & Özköse, 2023). According to Parasuraman et al. (2005) and Huang (2015), security and privacy refer to the degree to which individuals believe the application is safeguarded against breaches and that their personal information is secure. Item 15 exhibited a stronger loading on the dimension "privacy and data security," although it was initially expected to belong to the dimension "reliability." Upon further analysis, it became evident that the item addresses trust in m-banking, a complex construct that significantly influences users' perceived risk when using the application (Koenig-Lewis, Palmer, & Moll, 2010; Chen, 2013). The respondents were likely to associate the notions of reliability and privacy and data security, which led the item to be considered in the "privacy and data security" dimension. Consequently, this item was classified under this category. The items 13, 14, and 16 to 24 were classified in the category "reliability." In the online banking context, reliability is a critical attribute that has been extensively studied by various authors (TsaiTai, Lin & Lin, 2018; Zoghlami, Yahia & Berraies, 2018; Shankar et al., 2019; Sharma & Sharma, 2019; Shankar et al., 2020).

CONFIRMATORY FACTOR ANALYSIS

A confirmatory factor analysis was performed after defining the items that best related to each factor. The construct perceived quality in using m-banking was formed by two dimensions: "privacy" and "reliability," measured by 24 items, all of which were submitted to the CFA, obtaining the results shown in Table 3.

Table 3
Fit indices

Summary of indices in the CFA	
RMSEA	0,08
CFI	0,91
Cronbach's alpha of factors1 and 2	0.94 and 0.88 respectively

A structure can only be validated if it is reliable and the set of indicators produces interpretable, adequate, and accurate measures for the constructs in a structural model (Cronbach, 1957; Malhotra, 2012). Achieving a Cronbach's Alpha above 0.7 for the factors ensures the reliability criterion of the model.

The final model comprised 22 items: 12 items for the "privacy" dimension and 10 for the "reliability" dimension. These two dimensions together explain 62% of the model's total variance. Based on these findings, a multidimensional model was adopted to develop the Privacy and Perceived Reliability Scale for m-banking applications using IRT.

TEORIA DA RESPOSTA AO ITEM MULTIDIMENSIONAL

A construção do modelo do traço latente iniciou-se com a estimação dos parâmetros dos itens pelo modelo de resposta gradual utilizando o pacote *mirt* do *software* R. Os valores dos parâmetros de discriminação (a) para cada dimensão, a discriminação multidimensional (MDISC) de cada item e também os parâmetros de dificuldade multidimensional (MDIFF) são apresentados na tabela 3:

 DOI: https://doi.org/10.14210/alcance.v31n2(mai/ago).52-66

Table 4

Estimation of discrimination and multidimensional difficulty parameters.

Item	a1	a2	MDISC (A)	MDIFF (B1)	MDIFF (B2)	MDIFF (B3)	MDIFF (B4)
01	2,11	1,24	2,45	NA	-2,29	-1,03	0,68
02	1,47	0,81	1,68	-2,60	-1,22	-0,15	1,37
03	1,24	0,57	1,36	-2,17	-0,71	0,44	1,71
04	1,71	1,08	2,02	-2,71	-1,87	-0,70	0,54
05	0,60	0,73	0,95	-2,91	-1,24	0,53	2,24
06	1,66	1,05	1,96	NA	-1,57	-0,50	0,79
07	2,76	1,17	3,00	-2,30	-1,46	-0,70	0,60
08	2,74	1,63	3,19	-2,39	-1,54	-0,72	0,73
09	3,56	2,22	4,20	NA	-1,80	-1,31	0,10
10	2,34	1,48	2,77	-2,35	-1,44	-0,71	0,54
11	3,91	2,31	4,54	NA	-1,95	-1,08	0,24
12	2,71	1,83	3,27	-3,21	-2,04	-1,27	-0,18
13	0,26	1,81	1,83	NA	-2,60	-1,49	-0,24
14	-0,42	1,84	1,89	-1,91	-1,09	-0,16	1,06
15	2,39	2,16	3,22	NA	-2,08	-0,95	0,54
16	0,69	1,28	1,45	NA	-4,33	-2,57	-1,02
17	0,47	1,72	1,78	NA	-2,43	-1,67	-0,04
18	0,01	2,68	2,68	NA	-1,98	-1,49	-0,27
19	0,48	1,68	1,74	-2,87	-2,37	-1,45	-0,21
20	0,01	1,50	1,50	-2,31	-1,19	0,08	1,16
21	-0,96	2,85	3,01	-1,60	-1,05	-0,21	0,68
22	-0,93	2,65	2,81	-1,34	-0,91	-0,02	0,92
23	0,28	2,87	2,88	NA	-1,85	-0,94	0,34
24	0,56	3,02	3,07	NA	-2,55	-1,47	-0,19

Analyzing the values of the discrimination parameters for each item, it is evident that, in the compensatory IRT analysis, the loadings within each dimension align with the findings of the confirmatory analysis, indicating the dimension that provides the most information. Generally, the two dimensions share loadings (secondary loadings) across most items. However, the primary loading for each item can be clearly identified, demonstrating its specific alignment with one dimension. Unlike factor analysis, shared loadings across multiple dimensions in IRT analysis do not pose a problem. Instead, this sharing reflects the compensatory nature of the model.

To exemplify and facilitate the understanding of the analysis of the item parameters, Figure 1 presents the surfaces of the characteristic curves of two items, with the highest and lowest values found in the estimated model, from different angles to facilitate the interpretation. Figure 1 shows the response surfaces for the item and not just the characteristic curves as in the one-dimensional models, which makes the interpretation more complex. The images show the axes θ1 (privacy) and θ2 (reliability) and the probability of response of the individuals (vertical axis). Note that Figures 1a and 1b represent the same item 11, and Figures 1c and 1d represent item 5. Two angles of each item are presented to improve the interpretation.

Figure 1

Examples of high (item 11) and low (item 05) discrimination parameters

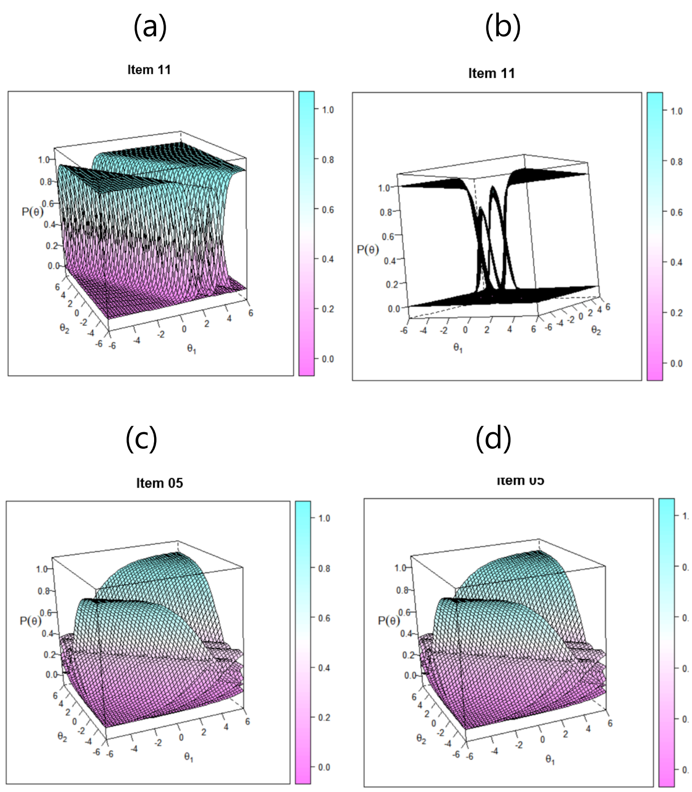


Figure 2 illustrates the distinction between highly discriminative (steep) curves and less discriminative (flatter) curves. Unlike in one-dimensional models, where characteristic curves are more straightforward to interpret, multidimensional models are more complex due to the inclusion of multiple dimensions (in this case, two). Generally, approximated curves indicate that small increases in ability result in substantial increases in the probability of selecting the correct response (high discrimination). Conversely, widely spaced curves suggest that significant increases in ability are required to achieve small

 DOI: https://doi.org/10.14210/alcance.v31n2(mai/ago).52-66

increases in the probability of a proper response (low discrimination).

Item 11 exhibited the highest value of the MDISC discrimination parameter (4.54) among all items in the model, signifying that it is highly discriminative. The steep and distinct curves for Item 11, as shown in Figures 2a and 2b, make it effective at differentiating individuals with varying levels of latent traits. These curves show that the response categories in the middle range (3 and 4) have the highest probabilities of a positive response, contributing significant information in this region of the graph. The abrupt changes in probabilities further emphasize the high discrimination power of this item.

On the other hand, Item 5 demonstrated the lowest discriminatory power (MDISC = 0.95), indicating that it is less effective at distinguishing respondents. This is evident in the flatter curves shown in Figures 2c and 2d, which represent lower response probabilities and less information in the central region of the graph.

Figure 2 also provides an alternative angle to view the characteristic curves of Items 5 and 11, offering insight into the dimensions most closely associated with each item. The orientation of the curves reveals that Item 11 is predominantly aligned with Dimension 1 (privacy), as its curves are oriented toward the θ1 axis. Meanwhile, Item 5, which had a high secondary loading as previously discussed, conveys information from both dimensions, with its curves spanning both the θ1 and θ2 axes.

Figure 2

Characteristic curve of items 11 and 05 demonstrating in which dimension they are most characteristic.

(a) (b)

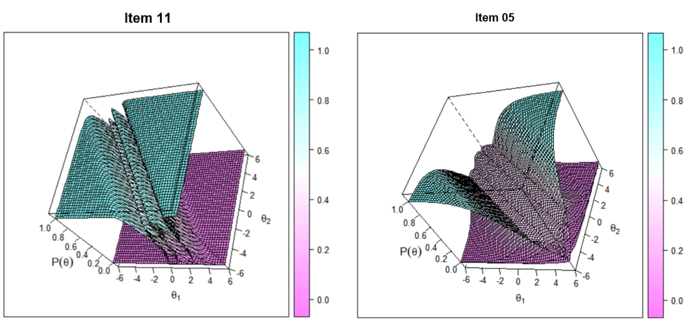
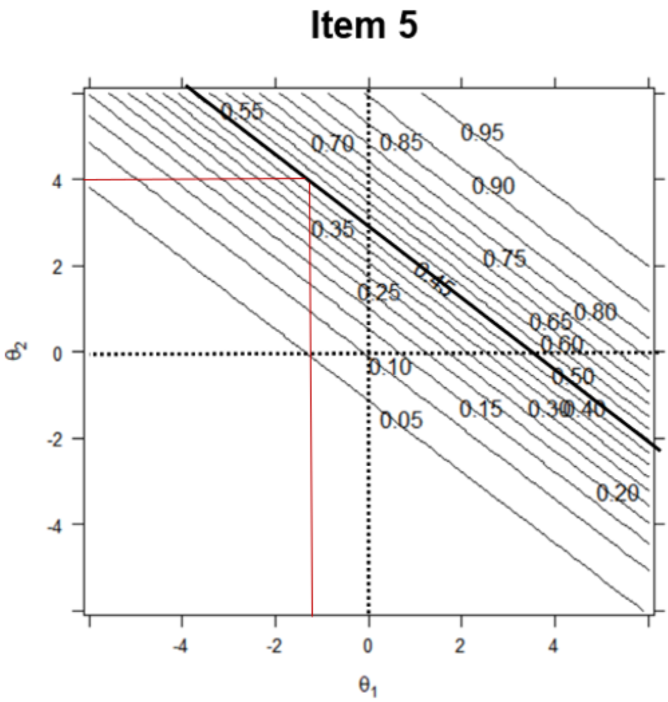


Figure 3 shows the equiprobability graph of the item's last response category ("I completely agree"). The darker line represents the 0.5 probability of an individual agreeing with the item.

Figure 3

Equiprobability lines for the last response category (I completely agree) of item 5



When analyzing the probability of a response in the final category ("I totally agree"), it becomes evident that respondents positioned above the bold line are more likely to respond positively to this item. Specifically, they are more inclined to fully agree with the statement: "When using m-banking, the privacy of my personal information is guaranteed by the operating system (Android/iOS) on my cell phone." This observation aligns with the MDIFF parameter value for the item's last category (2.24), indicating that individuals must possess a very high ability (i.e., a strong perception of privacy and reliability) to agree with this statement. Notably, this item exhibits the lowest MDISC parameter in the entire model. Consequently, its equiprobability curves or contour lines are more spaced apart. The flatter curves of the item's CCI, as shown in Figures 1c and 1d, reveal limited discrimination in distinguishing respondents. Moreover, Figure 3 demonstrates that the item primarily loads onto the reliability dimension (θ2), though it also shares a significant load with the privacy dimension (θ1). This dual loading suggests that the item captures elements of both privacy and reliability in its theoretical foundation, a topic elaborated on later in the analysis.

Figure 3's equiprobability lines illustrate the compensatory nature of the model. The red line highlights this property by showing that a low skill level on the θ1 axis (privacy dimension) can be offset by a high skill level on the θ2 axis (reliability dimension). For this specific item, lower perception of privacy and limited knowledge of information security issues are compensated by a greater perception of reliability in m-banking. Thus, a higher reliability perception balances out a weaker privacy perception. Any point on the bold equiprobability line (P = 0.5) corresponds to the same response probability, illustrating this compensatory relationship.

This multidimensional, compensatory approach enriches the literature by addressing a gap in measuring privacy and reliability in m-banking through a more robust statistical technique. Other studies, such as those by Almaiah et al. (2023) and Apaua and Lallie (2022), employ structural equation modeling to assess perceptions of reliability and privacy. Similarly, Cavus et al. (2022) use neural networks to analyze relationships between perception and attractiveness. However, these methodologies primarily focus on exploring the relationships between constructs rather than delving into the constructs themselves. The approach presented in this study advances the field by developing a standardized metric for effectively measuring the privacy and reliability constructs and their multidimensional interactions. This contributes to the work of Souza and Tezza (2021), who applied IRT to measure perceived risk and performance in Brazilian m-banking consumers, though their study adopted a unidimensional model.

CONCLUSIONS

The rapidly evolving competitive landscape and an ever-growing customer base are driving financial institutions to enhance their service channels to improve perceptions of service quality. This study aimed to contribute to the field by developing an instrument to measure the perceived quality construct, focusing on data privacy and reliability in m-banking applications. The objective was to understand how users perceive quality when interacting with these applications. A review of the concept of quality, complemented by a literature review on m-banking, revealed that "privacy and data security" and "reliability" are critical attributes in this context. Based on these attributes, a set of 26 items was developed to assess the dimensions of privacy and reliability in m-banking. To analyze the data and refine the scale, item response theory (IRT) was employed, as it facilitates a deeper understanding of the characteristics under investigation. Before estimating the IRT model, exploratory and confirmatory factor analyses were conducted to identify the multidimensionality of the construct.

This research offers significant contributions to both academia and the market. Academically, it introduces a model not yet extensively replicated in the literature, providing new insights into the perceived quality of m-banking services. The study also advances the systematization of research in this area, offering a robust theoretical framework. From a practical perspective, the findings highlight key elements of user perception regarding reliability and privacy/security in m-banking applications. For instance, concerns about sharing personal information with third parties and the security provided by mobile operating systems raise questions for users about protecting their personal data.

To ensure transparency, the study acknowledges certain limitations while offering directions for future research. The analysis was confined to privacy and reliability, which are quality dimensions but not the entirety of the multidimensional concept of quality. Future research should expand the scope to include all quality dimensions in mobile device usage, enabling a more comprehensive understanding of overall quality perception. Additionally, exploring m-banking usage in the financial routines of legal entities represents a promising avenue for further study.

REFERENCES

Almaiah, M. A., Al-Otaibi, S., Shishakly, R., Hassan, L., Lutfi, A., Alrawad, M., & Alghanam, O. A. (2023). Investigating the role of perceived risk, perceived security and perceived trust on smart m-banking application using SEM. Sustainability, 15(13), 9908.

Andrade, D., Tavares, H. & Talle, R. (2000). Teoria de Resposta ao Item: conceitos e aplicações. ABE — Associação Brasileira de Estatística.

Apaua, R., & Lallie, H. S. (2022). Measuring user perceived security of mobile banking applications. arXiv preprint arXiv:2201.03052

Albashrawi, M., & Motiwalla, L. (2019). Privacy and personalization in continued usage intention of mobile banking: An integrative perspective. Information Systems Frontiers, 21, 1031-1043.

Araujo, E. A. C. D., Andrade, D. F. D., & Bortolotti, S. L. V. (2009). Teoria da resposta ao item. Revista da Escola de Enfermagem da USP, 43, 1000-1008.

Arcand, M.; Promtep, S.; Brum, I.; & Rajaobelina, L. (2017). Mobile banking service quality and customer relationships. International Journal of Bank Marketing, 35(7), 1068-1089.

Ayala, R. (2009). The theory and practice of Item Response Theory. The guilford press. NY..

Cantú, C; Ulloa, B. (2020). The dawn of fintech in Latin America: landscape, prospects and challenges. BIS papers, 112. Disponível em: https://www.bis.org/publ/bppdf/bispap112.pdf. Acesso em: 22 jan. 2024.

Cavus, N., Mohammed, Y. B., Gital, A. Y. U., Bulama, M., Tukur, A. M., Mohammed, D., ... & Hassan, A. (2022). Emotional artificial neural networks and gaussian process-regression-based hybrid machine-learning model for prediction of security and privacy effects on M-banking attractiveness. Sustainability, 14(10), 5826.

Çelik, K., & Özköse, H. (2023). Investigation of Factors Affecting Mobile Banking Intention to Use: Extended Technology Acceptance Model. İşletme Araştırmaları Dergisi, 15(2), 1517-1530.

Chen, S. (2013). To use or not to use: Understanding the factors affecting continuance intention of mobile banking. International Journal of Mobile Communications. 10(5), 490-507.

Costa, F. (2011) Mensuração e desenvolvimento de escalas: aplicações em administração. Rio de Janeiro: Ciência Moderna.

Cronbach, L. J. (1957). Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 297-334.

de Souza, M. W., & Tezza, R. (2021). Measuring perceived risk of privacy and performance of m-banking in Brazil: An investigation applying item response theory. Latin American Business Review, 22(2), 131-161.

Federação Brasileira de Bancos - FEBRABAN- (2020) Pesquisa Febraban de Tecnologia Bancária https://cmsportal.febraban.org.br/Arquivos/documentos/PDF/pesquisa-febraban-relatorio.pdf Acesso em: 14 ago. 2021.

Federação Brasileira de Bancos (FEBRABAN). (2023). Pesquisa Febraban de Tecnologia Bancária. São Paulo, 2023. Disponível em: https://cmsarquivos.febraban.org.br/Arquivos/documentos/PDF/Pesquisa%20Febraban%20de%20Tecnologia%20Banc%C3%A1ria%202023%20-%20Volume%202.pdf Acesso em: 10 maio 2024.

Federação Brasileira de Bancos (FEBRABAN). (2024). Febraban Tech 2024: bancos à frente na segurança. Disponível em https://febrabantech.febraban.org.br/temas/seguranca/febraban-tech-2024-bancos-a-frente-na-seguranca Acesso em: 10 maio 2024.

Hanafizadeh, P., Behboudi, M., Koshksaray, A. A., & Tabar, M. J. S. (2014). Mobile-banking adoption by Iranian bank clients. Telematics and informatics, 31(1), 62-78.

Hambleton, K. (2000). Emergence of Item Response Modeling in Instrument Development and Data Analysis. Medical Care, 38(9), 60-65.

Huang, E., Lin, S., & Fan, Y. M-S-QUAL: Mobile service quality measurement. (2015). Electronic Commerce Research and Applications, 14, 126-142.

Jebarajakirthy, C., & Shankar, A. (2021). Impact of online convenience on mobile banking adoption intention: A moderated mediation approach. Journal of Retailing and Consumer Services, 58, 102323.

Jun, M., & Palacios, S. (2016). Examining the key dimensions of mobile banking service quality: An exploratory study. International Journal of Bank Marketing, 34(3), 307-326.

Kar, A. K. (2021). What affects usage satisfaction in mobile payments? Modelling user generated content to develop the "digital service usage satisfaction model". Information Systems Frontiers, 23(5), 1341-1361.

Koenig-Lewis, N., Palmer, A., & Moll, A. (2010) "Predicting young consumers' take up of mobile banking services", International Journal of Bank Marketing, 28, 410-432.

Legass, H. A., & Durmuş, M. E. (2024). Factors determining the adoption of mobile banking in Ethiopian Islamic banks: extension of technology acceptance model (TAM). Journal of Islamic Accounting and Business Research.

Liao, Z., & Cheung, T. (2008) Measuring consumer satisfaction in Internet banking: a core framework, Communications of the ACM. v.51, n.4, p. 236 -241.

Lin, H. (2013). Determining the relative importance of mobile banking quality factors. Computer Standards & Interfaces, 35, 195-204.

Malhotra, N. K. (2012). Modelagem de equações estruturais e análise de caminho. In: Pesquisa de marketing: uma orientação aplicada, 6(22), 549-575. Porto Alegre: Bookman.

Nisha, N. (2016) Exploring the dimensions of mobile banking service quality: Implications for the banking sector. International Journal of Business Analytics, 3(3), 1-24.

Nojosa, R. (2002). Teoria da Resposta ao Item (TRI): modelos multidimensionais. Estudos em Avaliação Educacional, 25, 123-166.

Parasuraman, A., Zeithaml, V., & Malhotra, A. (2005). E-S-Qual: a multiple-item scale for assessing electronic service quality, Journal of Service Research, 7(3), 213-233.

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reckase M. D. (1985) The difficulty of test items that measure more than one ability. Applied Psychological Measurement, 9, 401-412.

Reckase, M. D. (1997). A linear logistic multidimensional model for dichotomous items response data. In: W. J. van der Linden and Hambleton (Ed.), Handbook of Modern Item Response Theory (p. 271-286). New York: Springer.

Reckase, M. D. (2009). Mutidimensional Item Response Theory. New York: Springer.

Samejima F. (1974). Normal ogive model on the continuous response level in the multidimensional space. Psychometrika, 39, 111-121.

Sagib, G., & Zapan, B. (2014) Bangladeshi mobile banking service quality and customer satisfaction and loyalty. Management & Marketing, 9(3), 331-346.

Shankar, A., Datta, B., & Jebarajakirthy, C. (2019). Are the Generic Scales Enough to Measure Service Quality of Mobile banking? A Comparative Analysis of Generic Service Quality Measurement Scales to Mobile banking Context. Services Marketing Quarterly, 40(3), 224-244.

Shankar, A., Datta, B., Jebarajakirthy, C., & Mukherjee, S. (2020). Exploring Mobile banking Service Quality: A Qualitative Approach. Services Marketing Quarterly, 41(2).

Stefani, C., Azevedo, M., & Duduchi, M. (2019). Qualidade em mobile banking segundo professores e estudantes de tecnologia brasileiros. Brazilian Journal of development, 5(7), 9932-9957.

Szymanski, D., & Hise, R. (2000). E-satisfaction: Initial examination, Journal of Retailing, 76(3), 309-322.

Tezza, R., Bornia, A. C., Spenassato, D., & Trierweiller, A. C. (2016). Empirical comparison of the multidimensional models of item response theory in e-commerce. Pesquisa Operacional, 36, 503-532.

Trabelsi-Zoghlami, A., Berraies, S., & Ben Yahia, K.

 DOI: https://doi.org/10.14210/alcance.v31n2(mai/ago).52-66

(2018). Service quality in a mobile-banking-applications context: do users' age and gender matter? Total Quality Management and Business Excellence, 1-30.

Wolfinbarger, M., & Gilly, M. (2003). ETailQ: Dimensionalizing, measuring and predicting retail quality", Journal of Retailing, 79(3), 183-198.

Zaffalon, F., Prisco, A., Souza, R., Bez, J., Tonin, N., Penna, R., & Botelho, S. (2020) Comparative Study between Models that Estimate Student's Skill in Virtual Programming Environments. Revista Brasileira de Informática na Educação – RBIE, 28, 776-795.