Drawing upon shopping apps from Brazil and Spain, this thesis develops a scale to measure the online conditions that induce the flow state among users during browsing sessions. Although extensively researched, the cognitive state of flow hinges upon specific, yet under-examined conditions: balance between demands and users' requisite skills, goal clarity, and the nature of feedback received. This study aims to explore this relatively overlooked aspect of the literature, offering insights that can significantly advise the design of online shopping environments.

Advisor: Prof. Dr. Rafael Tezza (UDESC)
Co-advisor: Prof. Dr. Antoni Meseguer-Artola (UOC)

BRAZIL & SPAIN, 2024

YEAR
2024

EDUARDO HENRIQUE DE BORBA | Measurement of Online Flow Conditions in Shopping Applications: A Study Using Item Response Theory

UNIVERSIDADE DO ESTADO DE SANTA CATARINA – UDESC
PROGRAMA DE DOUTORADO ACADÊMICO EM ADMINISTRAÇÃO

UNIVERSITAT OBERTA DE CATALUNYA – UOC
PROGRAMA DE DOCTORAT EN DRET, POLÍTICA I ECONOMIA

CO-SUPERVISED Ph.D THESIS

# MEASUREMENT OF ONLINE FLOW CONDITIONS IN SHOPPING APPLICATIONS: A STUDY USING ITEM RESPONSE THEORY

EDUARDO HENRIQUE DE BORBA

BRAZIL & SPAIN, 2024

**EDUARDO HENRIQUE DE BORBA**

# MEASUREMENT OF ONLINE FLOW CONDITIONS IN SHOPPING APPLICATIONS: A STUDY USING ITEM RESPONSE THEORY

The thesis presented as a requirement for obtaining the title of Ph.D. in a co-supervision agreement between the Santa Catarina State University (UDESC) and the Universitat Oberta de Catalunya (UOC).
Advisor: Prof. Dr. Rafael Tezza
Co-advisor: Prof. Dr. Antoni Meseguer-Artola

**BRAZIL AND SPAIN**

**2024**

**Ficha catalográfica elaborada pelo programa de geração automática da Biblioteca Universitária Udesc, com os dados fornecidos pelo(a) autor(a)**

Borba, Eduardo Henrique de
Measurement Of Online Flow Conditions In Shopping Applications : A Study Using Item Response Theory / Eduardo Henrique de Borba. -- 2024.
165 p.

Orientador: Rafael Tezza
Coorientador: Antoni Meseguer-Artola
Tese (doutorado) -- Universidade do Estado de Santa Catarina, Centro de Ciências da Administração e Socioeconômicas - ESAG, Programa de Pós-Graduação em Administração, Florianópolis, 2024.

1. flow theory. 2. online flow conditions. 3. mobile commerce. 4. shopping applications. 5. item response theory. I. Tezza, Rafael. II. Meseguer-Artola, Antoni. III. Universidade do Estado de Santa Catarina, Centro de Ciências da Administração e Socioeconômicas - ESAG, Programa de Pós-Graduação em Administração. IV. Titulo.

**EDUARDO HENRIQUE DE BORBA**

# MEASUREMENT OF ONLINE FLOW CONDITIONS IN SHOPPING APPLICATIONS: A STUDY USING ITEM RESPONSE THEORY

The thesis presented as a requirement for obtaining the title of Ph.D. in a co-supervision agreement between the Santa Catarina State University (UDESC) and the Universitat Oberta de Catalunya (UOC).

**Examination Board:**

Advisor: ______________________________

Prof. Dr. Rafael Tezza
(UDESC)

Co-advisor: ______________________________

Prof. Dr. Antoni Meseguer-Artola
(UOC)

Members: ______________________________

Prof. Dra. Inma Rodríguez-Ardura
(UOC)

______________________________

Prof. Dr. Denilson Sell
(UDESC)

______________________________

Prof. Dr. Júlio da Silva Dias
(UDESC)

Florianopolis, September, 2024.

# Assinaturas do documento

Código para verificação: **3WH985GG**

Este documento foi assinado digitalmente pelos seguintes signatários nas datas indicadas:

**RAFAEL TEZZA** (CPF: 031.XXX.849-XX) em 27/09/2024 às 13:06:11
Emitido por: "SGP-e", emitido em 30/03/2018 - 12:36:24 e válido até 30/03/2118 - 12:36:24.
(Assinatura do sistema)

**DENILSON SELL** (CPF: 921.XXX.529-XX) em 27/09/2024 às 14:32:21
Emitido por: "SGP-e", emitido em 30/03/2018 - 12:36:13 e válido até 30/03/2118 - 12:36:13.
(Assinatura do sistema)

Para verificar a autenticidade desta cópia, acesse o link https://portal.sgpe.sea.sc.gov.br/portal-externo/conferencia-documento/VURFU0NfMTIwMjJfMDAwMTc1MTlfMTc1NTJfMjAyMl8zV0g5ODVHRw== ou o site https://portal.sgpe.sea.sc.gov.br/portal-externo e informe o processo **UDESC 00017519/2022** e o código **3WH985GG** ou aponte a câmera para o QR Code presente nesta página para realizar a conferência.

MESEGUER ARTOLA, ANTONI (FIRMA)

Firmado digitalmente por MESEGUER ARTOLA, ANTONI (FIRMA)
Nombre de reconocimiento (DN): c=ES, serialNumber=40932162M, sn=MESEGUER, givenName=ANTONI, cn=MESEGUER ARTOLA, ANTONI (FIRMA)
Fecha: 2024.10.08 10:36:26 +02'00'

gov.br Documento assinado digitalmente
EDUARDO HENRIQUE DE BORBA
Data: 14/10/2024 14:32:51-0300
Verifique em https://validar.iti.gov.br

gov.br Documento assinado digitalmente
JULIO DA SILVA DIAS
Data: 17/10/2024 11:25:38-0300
Verifique em https://validar.iti.gov.br

RODRIGUEZ ARDURA INMACULADA - 46588211D

Firmado digitalmente por RODRIGUEZ ARDURA INMACULADA - 46588211D
Fecha: 2024.10.16 21:12:08 +02'00'

I dedicate this thesis to those whose support and encouragement were essential to my academic journey.

## ACKNOWLEDGMENT

I am deeply grateful to my parents and grandparents, who instilled in me the importance of knowledge from an early age. I also appreciate their understanding during my absences at crucial moments and their unwavering encouragement and support, which have brought me to where I am today.

My deepest gratitude goes to God, who has been my unwavering foundation, providing me with the resilience and faith to overcome challenges and achieve my goals.

To my advisor, Prof. Rafael Tezza, I owe sincere thanks for his excellent guidance over the past four years. His readiness to assist with my doubts and academic aspirations and his support in numerous projects further motivated me to delve into quantitative methods. I also extend my heartfelt thanks to my co-advisor, Prof. Antoni Meseguer-Artola, for his steadfast support from our first email exchange. His consistently relevant advice and critical insights have been crucial to the development of this study. I am especially grateful for his willingness to open the doors of the Universitat Oberta de Catalunya (UOC) for my academic exchange.

I am also thankful to my doctoral friends who accompanied me on this journey, sharing experiences and frustrations over these four years. Additionally, I cherish the friendships I made during my unforgettable year of doctoral research in Barcelona, which enriched my time abroad.

To my long-time friends and life companion, your steadfast support and encouragement gave me strength during moments of doubt.

I extend my gratitude to the professors who accepted the invitation to be part of the examination board – Prof. Inma Rodríguez-Ardura, Prof. Denilson Sell, and Prof. Júlio da Silva Dias – for their time and valuable contributions to this thesis, both present and forthcoming. My thanks also go to the professors of the Postgraduate Department at the Administrative and Socioeconomic Sciences College (ESAG) from Santa Catarina State University (UDESC) and the Doctoral Program in Law, Politics, and Economy at UOC, who were always ready to help. I am also grateful to the members of UDESC's LabGES research group and UOC's DigiBiz group for their invaluable tips and advice.

I am deeply appreciative of the participants in the two eye-tracking user tests conducted in Brazil and Spain, whose involvement was essential to the development of this thesis.

My gratitude extends to the Coordination of Superior Level Staff Improvement (CAPES) for the financial support received over these four years through the doctoral scholarship, as well as for the financial aid provided for my doctoral research in Spain. I also thank the National Council for Scientific and Technological Development (CNPq) and the Research and Innovation Support Foundation of Santa Catarina State (FAPESC) for their contributions to the LabGES research group, enabling the acquisition of equipment that facilitated my research.

I express my gratitude to the Santa Catarina State University (UDESC), which provided quality education and supported me since my bachelor's degree, and to the Open University of Catalunya (UOC), which enabled me to complete my doctorate under a co-supervision agreement. Finally, my thanks go to everyone who, in one way or another, contributed to my personal and academic development.

"If you cannot measure it, you cannot improve it." (Lord Kelvin)

## ABSTRACT

This thesis introduces a scale designed to measure the effectiveness of shopping applications in inducing the online flow state among users. Online flow, a cognitive state experienced during internet browsing, manifests under specific conditions where users have clear goals, possess adequate skills, and receive immediate feedback. Utilizing Item Response Theory (IRT), this study developed a robust online flow conditions measurement scale grounded in prior literature on online flow, findings from two eye-tracking user tests conducted in Brazil and Spain, and qualitative experts analysis. Central to this research is Csikszentmihalyi's Theory of Flow (1975), which examines intrinsic motivation and optimal task performance. With the proliferation of the internet, scholars have reinvigorated this concept to enhance their comprehension of online environments through the lens of flow theory. This renewed scholarly interest has served as the foundation for building this thesis. The proposed scale was applied to a sample of 200 shopping applications (100 available in Brazil and 100 in Spain) to evaluate their efficacy across distinct social and economic contexts. Data analysis, employing IRT, addressed issues of dimensionality and model selection, enriching the empirical foundation of Flow Theory in online situations. Managerially, this research identifies critical elements that app developers can control to enhance flow in shopping applications, crucial for shaping effective business marketing strategies. Academically, it contributes to an under-explored area within information technology and communication, particularly in the realm of online flow conditions in virtual environments.



**Keywords:** flow theory; online flow conditions; mobile commerce; shopping applications; item response theory.

**RESUMEN**

Esta tesis presenta una escala diseñada para medir la efectividad de las aplicaciones de compras en inducir el estado de flujo en línea entre los usuarios. El flujo en línea, un estado cognitivo experimentado durante la navegación por internet se manifiesta bajo condiciones específicas donde los usuarios tienen metas claras, poseen habilidades adecuadas y reciben retroalimentación inequívoca. Utilizando la Teoría de Respuesta al Ítem (IRT), este estudio desarrolló una sólida escala de medición de condiciones de flujo en línea fundamentada en la literatura previa sobre flujo en línea, hallazgos de dos testes con usuarios basados en eye-tracking realizados en Brasil y España, y análisis cualitativos con expertos. En el centro de esta investigación se encuentra la Teoría del Flujo de Csikszentmihalyi (1975), que examina la motivación intrínseca y el rendimiento óptimo en las tareas. Con la proliferación de internet, los académicos han revitalizado este concepto para mejorar su comprensión de los entornos en línea a través de la teoría del flujo. Esta tesis se fundamenta en este renovado interés académico por el tema. La escala propuesta se aplicó a una muestra de 200 aplicaciones de compras (100 disponible en Brasil y 100 en España) para evaluar su eficacia en distintos contextos sociales y económicos. El análisis de datos, utilizando IRT, abordó cuestiones de dimensionalidad y selección de modelos, enriqueciendo el fundamento empírico de la Teoría del Flujo en situaciones en línea. Desde la perspectiva de la gestión empresarial, esta investigación identifica elementos críticos que los desarrolladores de aplicaciones pueden controlar para mejorar el flujo en aplicaciones de compras, aspecto crucial para la formulación de estrategias efectivas de marketing empresarial. Académicamente, contribuye a un área poco explorada dentro del ámbito de la tecnología de la información y la comunicación, particularmente en el contexto de las condiciones de flujo en línea en entornos virtuales.

**Palabras clave**: teoría del flujo; condiciones de flujo en línea; comercio móvil; aplicaciones de compras; teoría de respuesta al ítem.

# RESUMO

Esta tese apresenta uma escala projetada para medir a eficácia dos aplicativos de compras em induzir o estado de fluxo online entre os usuários. O fluxo online, um estado cognitivo experimentado durante a navegação na internet, manifesta-se sob condições específicas onde os usuários têm objetivos claros, possuem habilidades adequadas e recebem feedback inequívoco. Utilizando a Teoria da Resposta ao Item (IRT), este estudo desenvolveu uma escala robusta de medição das condições de fluxo online fundamentada na literatura anterior sobre fluxo online, resultados de dois testes com usuários baseado em eye-tracking realizados no Brasil e na Espanha, e análise qualitativa de especialistas. No cerne desta pesquisa está a Teoria do Fluxo de Csikszentmihalyi (1975), que examina a motivação intrínseca e o desempenho ótimo nas tarefas. Com a proliferação da internet, os acadêmicos revitalizaram este conceito para aprimorar sua compreensão dos ambientes online através da teoria do fluxo. Esta tese se fundamenta neste renovado interesse acadêmico pelo tema. A escala proposta foi aplicada a uma amostra de 200 aplicativos de compras (100 disponíveis no Brasil e 100 na Espanha) para avaliar sua eficácia em diferentes contextos sociais e econômicos. A análise de dados, utilizando IRT, abordou questões de dimensionalidade e seleção de modelos, enriquecendo o fundamento empírico da Teoria do Fluxo em situações online. Gerencialmente, esta pesquisa identifica elementos críticos que os desenvolvedores de aplicativos podem controlar para melhorar o fluxo em aplicativos de compras, crucial para moldar estratégias eficazes de marketing empresarial. Academicamente, contribui para uma área pouco explorada dentro da tecnologia da informação e comunicação, especialmente no âmbito das condições de fluxo online em ambientes virtuais.



**Palavras-chave:** teoria do fluxo; condições de fluxo online; comércio móvel; aplicativos de compras; teoria da resposta ao item.

# RESUM

Aquesta tesi presenta una escala dissenyada per mesurar l'eficàcia de les aplicacions de compra a l'hora d'induir l'estat de flux en línia entre els usuaris. El flux en línia, un estat cognitiu experimentat durant la navegació per internet, es manifesta sota condicions específiques on els usuaris tenen objectius clars, disposen d'habilitats adequades i reben un feedback inequívoc. Utilitzant la Teoria de la Resposta a l'Ítem (IRT), aquest estudi ha desenvolupat una escala robusta de mesura de les condicions de flux en línia, fonamentada en la literatura prèvia sobre el flux en línia, els resultats de dues proves d'usuari basades en el seguiment ocular realitzats al Brasil i a Espanya, i una anàlisi qualitativa d'experts. Al centre d'aquesta investigació hi ha la Teoria del Flux de Csikszentmihalyi (1975), que examina la motivació intrínseca i el rendiment òptim en les tasques. Amb la proliferació d'internet, els acadèmics han revitalitzat aquest concepte per millorar la seva comprensió dels entorns en línia a través de la teoria del flux. Aquesta tesi es fonamenta en aquest renovat interès acadèmic pel tema. L'escala proposada es va aplicar a una mostra de 200 aplicacions de compra (100 disponibles al Brasil i 100 a Espanya) per avaluar-ne l'eficàcia en diferents contextos socials i econòmics. L'anàlisi de dades, utilitzant l'IRT, va abordar qüestions de dimensionalitat i selecció de models, enriquint el fonament empíric de la Teoria del Flux en situacions en línia. A nivell de gestió, aquesta recerca identifica elements crítics que els desenvolupadors d'aplicacions poden controlar per millorar el flux en aplicacions de compra, essencials per dissenyar estratègies de màrqueting empresarials efectives. A nivell acadèmic, contribueix a una àrea poc explorada dins de la tecnologia de la informació i la comunicació, especialment en l'àmbit de les condicions de flux en línia en entorns virtuals.

**Paraules clau**: teoria del flux; condicions de flux en línia; comerç mòbil; aplicacions de compra; teoria de la resposta a l'ítem.

## FIGURES LIST

## CHARTS LIST

## TABLES LIST

## ABBREVIATIONS AND ACRONYMS LIST

| | |
|---|---|
| 1PL | One-Parameter Logistic Model |
| 2PL | Two-Parameter Logistic Model |
| AF | Average Fixation |
| AIC | Akaike Information Criterion |
| ANOVA | Analysis of Variance Hypothesis Test |
| AOI | Areas of Interest |
| ASV | Average Squared Variance |
| AVE | Average Variance Extracted |
| CAT | Computerized Adaptive Tests |
| CFI | Confirmatory Factor Analysis |
| CR | Composite Reliability |
| DT | Dwell Time |
| ET | Entry Time |
| ICT | Information and Communication Technology |
| IRT | Item Response Theory |
| KPI | Key Performance Indicators |
| MDIF | Multidimensional Difficulty Parameter |
| MDIS | Multidimensional Discrimination Parameter |
| MIRT | Multidimensional Item Response Theory |
| MSV | Maximum Shared Squared Variance |
| RMSEA | Root Mean Square Error Of Approximation |
| SABIC | Sample Adjusted Bayesian Information Criterion |
| SRMR | Standardized Root Mean Square Residuals |

## SYMBOLS LIST

| | |
|---|---|
| $\chi^2$ | Chi-square test |
| α | Cronbach's alpha |
| $>$ | Greater than |
| $\geq$ | Greater than or equal to |
| θ | Latent trait theta |
| $<$ | Less than |
| $\leq$ | Less than or equal to |
| ω | McDonald's omega |

## SUMMARY

# 1 INTRODUCTION

This chapter presents a thorough contextualization of the thesis's focus. It initiates with an overview of online commerce in Brazil and Spain, followed by a delineation of mobile commerce elements. Moreover, it articulates the research problem, outlines general and specific objectives, and furnishes justifications for the study. Lastly, it emphasizes the originality, relevance, potential impact, and research contributions within the academic sphere.

## 1.1 CONTEXTUALIZATION

The extensive adoption of mobile devices, such as smartphones and tablets, has significantly invigorated the online economy, catalyzing the expansion of mobile electronic commerce, commonly referred to as m-commerce. This form of online commerce emerged from the convergence of rapidly developing economies and the exponential proliferation of mobile Internet services (MCLEAN et al., 2020). M-commerce is a platform facilitating the transaction of information, services, and goods via mobile devices, enabling direct interactions between merchants and consumers.

According to the annual "Digital 2024: Global Overview Report" report by DataReportal, approximately 69.4% of the global population uses mobile devices. Regardless of the country being evaluated, it is noteworthy that using smartphones for shopping has become widespread.

This thesis focuses on two countries with different levels of economic development but with robust and lucrative digital markets: Brazil and Spain. Additionally, the choice is also based on an exchange program conducted in 2023 at the Universitat Oberta de Catalunya (UOC), which aimed to evaluate Spanish m-commerce, providing a comparative perspective beyond what is observed in Brazil.

Brazil and Spain stand out as significant adopters of m-commerce, with 53.6% of Brazilian internet users and 54.1% of Spanish internet users engaging in m-commerce activities (DATAREPORTAL, 2024). Notably, Spain has seen substantial growth in online turnover, surpassing 85 million euros by 2022, as reported by the National Observatory of Technology and Society (ONTSI, 2023).

Similarly, Brazilian electronic commerce revenues reached 254.4 billion reais in 2024, marking a 0.7% increase from the previous year, according to the 49th Webshoppers report (NIQ EBIT, 2024).

Moreover, the introduction of 4G Internet in 2010 played an essential role in developing a sophisticated digital infrastructure, significantly enhancing data transmission speeds in urban areas and enabling the seamless functionality of digital tools (DUHAN; SINGH, 2019). Subsequently, the advent of 5G technology in 2018 further transformed connectivity, catering to diverse digital needs by expanding internet access capabilities. This technological advancement has bolstered the framework for m-commerce, facilitating enhanced data collection and transmission capabilities essential for modern digital transactions.

The ubiquity of m-commerce is characterized by its convenience, allowing users to engage in mobile commerce activities from virtually any location (DUHAN; SINGH, 2019). The ease of downloading mobile applications, or apps, has significantly streamlined daily activities for contemporary consumers (MCLEAN et al., 2020). Shopping applications, functioning akin to virtual storefronts, require meticulous design to attract and guide users seamlessly, facilitating a flow experience that enhances engagement and encourages purchases.

King (2003) emphasizes the importance of creating online environments that provide an enjoyable user experience, which can engage users and influence their behavior. This idea is based on Csikszentmihalyi’s Flow Theory (1997). Creating environments characterized by continuous navigation, rapid response times, and minimal distractions is contingent upon acknowledging the complexity of facilitating flow in virtual spaces. Csikszentmihalyi (1997) defines flow as a holistic state of total involvement, applicable to various contexts, including virtual m-commerce environments where user engagement is pivotal to transaction completion.

The concept of online flow, perceived as the cognitive state during internet browsing, is essential for attracting and retaining consumers (NOVAK; HOFFMAN; YUNG, 2000). However, constructing effective measures for online flow in m-commerce, particularly within mobile apps, remains challenging because of its multidimensional nature (CSIKSZENTMIHALYI, 1990). While optimizing online flow presents a promising strategy benefiting both users and vendors, there remains a dearth of concrete guidance on designing flow-oriented online environments (MAHNKE; BENLIAN; HESS, 2015).

Therefore, the development of m-commerce and the improvement of the online flow experience represent critical avenues for fostering consumer engagement and driving electronic commerce growth. Future research efforts should focus on effectively measuring and optimizing online flow in diverse m-commerce contexts to refine methodologies, thereby advancing our understanding and application of flow theory in digital environments.

## 1.2 RESEARCH PROBLEM

Shopping mobile applications are becoming increasingly complex tools as the number of functions offered to consumers constantly increases (MCLEAN et al., 2020). Due to these new enhancements, app managers and developers are struggling to handle the influx of modeling and recommendations on designing and maintaining an application that awakens stream state in online consumers.

Research in several areas of technology was conducted to measure online flow (BILGIHAN et al. 2013; GAO; WAECHTER; BAI, 2015; BILGIHAN, 2016; LOPEZ; VIRTO; SAN-MARTIN, 2018; SINA; WU, 2019; AMEEN et al., 2021). Focused on the users' perception, the studies developed several items that sought to capture the spectrum of the online flow. Researchers evaluated these items in various contexts such as e-commerce, m-commerce, social commerce websites, online auctions, etc. through the analysis of different target populations: students, generation Y people, and experienced users, among others. The findings converge on the concept that online flow is a multidimensional construct (OZKARAA; OZMENA; KIM, 2017). However, few studies evaluate the characteristics of online environments capable of awakening flow in their users.

The systematic search conducted in this thesis, evidenced in Appendix A, found no research focusing on online design characteristics that lead to a flow state in users, highlighting this gap. Therefore, a deepening of this theme is necessary. From an organizational point of view, e-commerce environments leverage profitability and merge as a simple channel with consumers (SAFIEDDINE, 2016).

Currently, mobile devices are seen as an extension of the human body. People use these electronic devices daily and, regardless of value judgment, find them essential in various situations (MCLEAN et al., 2020). King (2003), based on Csikszentmihalyi's Flow Theory, highlights elements of virtual interfaces that induce online flow, known as the online flow conditions. Understanding these elements optimizes the formulation of m-commerce apps and increases the chances of achieving one of the main goals of online retailers: selling your product/service.

In this context, formulating a standardized scale that considers, as broadly as possible, the elements of apps that directly influence the online flow state is of excellent value to app developers, service providers, and users. The Item Response Theory emerges as an option for the development of a robust methodological systematic associated with a consolidated mathematical structure, capable of creating and analyzing measurement scales (EMBRETSON; REISE, 2013). In this type of scale, the focus is not only on the reduction of factors, as in factor analysis but also on the possibility of comparing items and their respondents, to further deepen knowledge about the latent variable analyzed.

According to this problem, this study proposes the following main question: How to measure online flow conditions in shopping mobile applications?

## 1.3 DEFINITION OF OBJECTIVES

### 1.3.1 General Objective

The m-commerce is increasingly fundamental to the global economy, providing significant opportunities for economic growth and innovation across countries. As mobile device usage and internet accessibility continue to expand, commercial transactions and digital interactions between consumers and businesses have become more prevalent. This sector not only drives economic activity by facilitating electronic commerce but also promotes digital inclusion and connectivity, which are crucial for socio-economic development (BORBA; TEZZA, 2023).

It is crucial, therefore, to develop virtual shopping environments that offer satisfying experiences and foster a state of flow among users. Flow is a psychological state where individuals are deeply engaged and focused during their online interactions. Environments that effectively cultivate this flow state not only enhance the efficiency of commercial interactions but also improve user satisfaction and build stronger brand loyalty (PEIFER; ENGESER, 2021).

Recognizing a gap in research on factors influencing online flow conditions within virtual shopping environments, as highlighted in the systematic search detailed in Appendix A, there is a pressing need for targeted studies in this area. Such research aims to explore and better understand the elements that contribute to creating virtual environments that meet consumer expectations and promote a positive flow state. This understanding is essential for both consumers seeking enjoyable online experiences and businesses aiming to thrive in the dynamic digital marketplace.

Due to that, the general objective of this research is to define a scale for measuring online flow conditions and to apply this measurement to a sample of mobile shopping applications from Brazil and Spain.

### 1.3.2 Specific Objectives

To achieve the overarching objective, four specific objectives were developed as a sequential framework. Initially, the study aims to clarify the concept of online flow conditions (first specific objective):

1) Elucidate the dimensions of Online Flow Conditions through a comprehensive literature review;

Subsequently, supported by eye-tracking user tests conducted in Brazil and Spain, the focus shifts to constructing measurement items (second specific objective);

2) Build items based on the literature to measure the degree of Online Flow Conditions in mobile shopping applications, supplemented by the implementation of eye-tracking users testing among users;

After that, the scale will go through validation and confirmation processes, so that it can be guaranteed that the scale is adequate to measure what is proposed, in addition, the scale will be analyzed through IRT modeling (third specific objective):

3) Analyze the scale dimensionality and evaluate the suitability of relevant item response theory models;

Finally, the last specific objective involves comparing shopping applications available in Brazil and Spain to evaluate the level of flow conditions exhibited by each (fourth specific objective):

4) Compare the proficiency levels in online flow conditions of the shopping applications available in Brazil and Spain.

## 1.4 THEORETICAL AND PRACTICAL JUSTIFICATION

The perception of online consumers regarding the intrinsic characteristics of Information and Communication Technologies (ICTs) plays a pivotal role in their adoption of these technologies and their propensity to engage in online purchasing (GUO; KLEIN, 2009; OZKARAA et al., 2017; CHEN; HSU; LU, 2018; SALLEH et al., 2019; RAHMAN et al., 2020). In this context, the concept of online flow is crucial as it enhances positive consumer perceptions, promoting an inherently enjoyable user experience that encourages exploration and facilitates purchasing behavior (KING, 2003). This state is characterized by a sense of timelessness, where users lose self-consciousness, derive pleasure, and are motivated to experiment with new offerings.

The pervasive integration of mobile devices into the daily routines of approximately 5.61 billion individuals worldwide, as reported in the "Digital 2024: Global Digital Overview" (DATAREPORTAL, 2024), has profoundly transformed application marketing by creating novel opportunities for user engagement. The success of m-commerce applications hinges significantly on their ability to cultivate a flow state among users. Therefore, developing apps that deliver an ideal and engaging experience has become a strategic imperative for organizations entering the online marketplace (KING, 2003; MAHNKE; BENLIAN; HESS, 2015).

Csikszentmihalyi's Flow Theory provides a foundational understanding of what motivates individuals to excel in tasks (CSIKSZENTMIHAYI, 1975). In today's predominantly online environment, concentration, control, and telepresence are crucial factors that drive user engagement (NOVAK; HOFFMAN; YUNG, 2000). Hence, studying online flow is pertinent as it enables online service and product providers to optimize consumer experiences, thereby supporting their economic objectives through enhanced competitiveness and market expansion.

The exploration of online flow has been a recurring theme since the proliferation of the World Wide Web in 1990. Pioneers such as Novak and Hoffman (1997) extended Csikszentmihalyi's theories to analyze consumer navigation in online environments. They underscored the importance of design elements in creating compelling user experiences across various commercial contexts on the web.

Despite a rich body of literature on electronic commerce user behavior and flow analyses (Raji et al., 2024), existing research often relies on traditional statistical approaches that may lack robustness. From a strategic standpoint, there has been limited integration of these findings with the practical realities faced by organizations, especially those focused on m-commerce.

Thus, there is a theoretical and practical imperative to consolidate this research using rigorous statistical methods. Approaches such as Item Response Theory (IRT), endorsed by Tezza et al. (2018), Peng et al. (2019), and Li et al. (2020), offer a mathematically sound framework for analyzing online environments. IRT facilitates the measurement of latent traits (difficult-to-measure characteristics) through observable items, enabling the development of a coherent and interpretable scale (HAMBLETON, 2000).

From a theoretical perspective, this study highlights the conceptualization and measurement of flow within mobile shopping apps, emphasizing the evaluation of its dimensions and the appropriateness of statistical modeling. It proposes a systematic approach to formulating items and applying theoretical models. Practically, it advocates for the creation of a standardized scale that directly interprets the relationship between online flow components and the performance of shopping apps. Such an approach aims to enhance app developers' understanding of which elements are most impactful and how changes to these factors can optimize flow experiences in mobile commerce contexts.

## 1.5 ORIGINALITY, RELEVANCE AND POTENTIAL

Electronic commerce research has evolved into a comprehensive field of study, demanding precise and coherent measurement approaches. Therefore, exploring methodologies and alternative research consolidation methods is imperative. However, systematic modeling often overlooks elements beyond users' direct and subjective preferences. This study aims to address this gap by proposing an objective framework grounded in research from the past decade on the online flow experience, thus pioneering objective modeling of design elements for flow, a previously unexplored endeavor in the literature.

Barta, Flavián, and Gurrea (2022) emphasize that designing environments focused on the flow experience is crucial for fostering loyal and satisfied customers in the virtual realm. Consequently, the design characteristics of online environments play a pivotal role in cultivating the flow experience during users' cyber journeys.

Despite a systematic search, shown in Appendix A, across databases such as Ebsco, Scopus, and Web of Science, a scarcity of research on design characteristics of online environments that induce flow among users was observed. Out of forty-nine articles retrieved, the focus predominantly lies on users' perceptions of the flow experience while using technology, neglecting intrinsic elements of virtual environment design. From Koufaris (2002), the earliest article in the sample, to Xie and Yuan (2021), the most recent, the literature primarily addresses user perceptions rather than virtual design elements, with King (2003) standing out as an exception. This theoretical gap underscores the urgent need for application-oriented studies in this area, which this study seeks to fulfill by proposing a standardized scale based on item response theory (IRT) to measure flow in mobile applications.

The choice of IRT highlights its applicability in the context of m-commerce, specifically mobile apps, highlighting the scientific relevance of this study. According to Borba and Tezza (2022), despite methodological advancements, the utilization of IRT models in IT-related fields remains underexplored. They argue for its robustness and versatility, criticizing existing scales for lacking rigor in their developmental phases.

This research anticipates significant technological relevance, proposing the development of a novel scale for measuring flow conditions in mobile shopping applications. Furthermore, it aims to advance knowledge in m-commerce and disseminate its techniques and applications within the online economy. The study also holds potential for broader impact, potentially expanding the use of IRT scales to measure flow in online environments, thereby stimulating new research avenues.

A comparison is also made between two economies with different levels of development but that progress together regarding the growth of their respective digital markets. This study aims to evaluate the levels of online flow conditions among the shopping apps available in Brazil and Spain that comprise the sample and to determine which are better equipped to induce flow in their users.

Additionally, this study contributes to the internationalization of scientific endeavors by leveraging globally recognized theories such as Flow Theory and drawing insights from international reports like the NIQ EBIT's Webshoppers and the DataReportal's Digital 2024: Global Overview Report. Focusing on m-commerce apps, a contemporary and underexplored topic, piques the interest of researchers in the technological domain, thereby enhancing UOC’s and UDESC’s visibility in international scientific discourse.

# 2 THEORETICAL FOUNDATION

This chapter covers the basic concepts of M-commerce, Flow Theory, online flow experience, measure scales, eye-tracking user testing, and Item Response Theory (IRT). In addition to the flow conceptualization, important considerations are raised about the definition of the online flow experience. The concepts of measurement scales and the specific characteristics of their formulation are addressed. Next, a conceptualization is made about the Item Response Theory along with a survey of questions relevant to its application, in this subsection are discussed some of its main models.

## 2.1 M-COMMERCE

Before discussing mobile commerce (m-commerce), it is necessary to distinguish it from its precursor, electronic commerce (e-commerce). M-commerce has evolved through two distinct phases: initially classified as an extension of e-commerce and subsequently emerging as an independent business area in addition to serving as an alternative mechanism for e-commerce itself (BORBA; TEZZA, 2023; 2021a).

Tezza et al. (2018) described e-commerce as a process that utilizes the Internet to facilitate business activities such as buying, selling, exchanging, and providing information services. They emphasize that the core of e-commerce lies in offering a platform with specialized, diverse, practical, and accessible information to enable efficient global operations at low cost. In contrast, according to Chaffey et al. (2020), the definition of e-commerce has expanded beyond electronically mediated financial transactions. It now encompasses various forms of information and experiential transactions, such as social networking and microblogging, which serve purposes beyond direct sales invitations.

E-commerce is a type of purchase and sale of products and services conducted through a computer network. Users can place purchase orders for the same products simultaneously in various places, while m-commerce brings the concept of mobility to the purchase process, allowing the user to make business transactions regardless of their geographic location. In this sense, the main difference between m-commerce and e-commerce is in the context of transaction and access, in which in m-commerce mobile devices, smartphones and tablets, are used exclusively to communicate at anytime and anywhere (DUHAN; SINGH, 2019).

According to Borba and Tezza (2023), one of the distinctive advantages of mobile devices in the realm of mobile commerce is their mobility. Unlike e-commerce transactions that are traditionally tied to stationary devices, mobile commerce services enable transactions and interactions on the go, leveraging the inherent mobility of smartphones and tablets. This mobility significantly influences how consumers present, process, and interact with retail experiences.

Wang et al. (2023) state that mobile devices offer several unprecedented functionalities such as user location detection, context sensing[1], and push delivery[2]. These capabilities have paved the way for new service categories like location-based services and context-aware applications. Due to the unique interactions facilitated by mobile devices, consumer motivations, expectations, and behaviors in mobile commerce differ significantly from those observed in traditional online settings via computers or notebooks.

Professionals and academics are increasingly studying the distinct emotional shopping experiences between traditional e-commerce and m-commerce platforms to optimize customer engagement strategies for each channel (BORBA; TEZZA, 2021b). According to Duhan and Singh (2019), m-commerce benefits significantly from mobile devices' inherent characteristics. Mobile devices provide ubiquity, allowing consumers to access information and conduct transactions anywhere and anytime via Internet connectivity. This accessibility enhances convenience and empowers consumers with instant access to product information, reviews, and comparisons, facilitating informed decision-making processes directly from their devices.

Furthermore, m-commerce excels in personalization by leveraging extensive data availability on the Internet to deliver tailored content that aligns with individual preferences and interests. The ability to deliver relevant and personalized experiences is crucial as consumers increasingly prioritize

[1] Technologies that infer the characteristics of one or more people, locations, objects, situations, or activities and use this information to adapt, synchronize, and dynamically frame situations and processes.
[2] Transfer of material posting from one storage location to another. But it can also happen between two different countries, where you need documents for border crossing.

relevance in their interactions with digital platforms. Additionally, the portability of mobile devices enables transactions to be seamlessly conducted while users are on the move, enhancing the flexibility of shopping experiences beyond the constraints of physical locations. These capabilities, combined with the potential of wireless infrastructures for widespread data dissemination, underline the transformative impact of m-commerce on consumer behavior and engagement strategies in comparison to traditional e-commerce models (DUHAN; SINGH, 2019; BORBA; TEZZA, 2021a).

Recent studies highlight the significant differences between m-commerce and e-commerce, particularly concerning the operating interface due to the compact dimensions of mobile devices. Lucas, Lunardi, and Dolci (2023) emphasize that the anatomical fit of mobile devices to the palm makes them inherently more practical than computers and notebooks, leading users to prefer functionality that aligns with their mobility needs.

Additionally, Thangavel and Chandra (2023) underscore accessibility as a critical factor in m-commerce, where the compact size of mobile devices facilitates easy carrying and storage. This accessibility extends beyond physical dimensions to encompass anytime, anywhere access enabled by wireless mobile communication networks, thereby distinguishing m-commerce by its enhanced mobility and convenience compared to traditional e-commerce platforms.

According to Zariman, Humaidi, and Abd Rashid (2023), the relationship between m-commerce and mobile applications is integral to the adaptation of digital platforms to mobile computing devices. While traditional websites may require adaptation to function optimally on smartphones or tablets, mobile applications are specifically developed to harness the unique capabilities and features of these devices. The authors define a mobile application as a dedicated software program designed to execute specific functions efficiently on mobile devices, highlighting its tailored approach to enhance user experience and functionality in the m-commerce ecosystem.

The information provided in the application must be comprehensive, relevant, and easily understandable. According to Portigal (2023), an information system should present its data in a manner that users perceive as clear and concise. In the context of a fashion items app, this means that marketers should avoid overwhelming users with excessive information and instead focus on delivering pertinent product details through the app. Failure to provide complete and relevant information can lead to user frustration and potentially result in the uninstallation of the app.

According to Aggarwal, Sharma, and Saxena (2024), investing in mobile applications for online commerce offers significant advantages, particularly in leveraging device hardware resources such as geolocation data, photo and video capabilities, telephony functions, and applying the huge functionalities of artificial intelligence. These functionalities enhance the app’s capability to deliver personalized and interactive experiences to users. Additionally, mobile applications benefit from utilizing system notifications and automatic updates, which contribute to a seamless and improved user experience. The authors also acknowledge the challenges associated with mobile applications, such as the need for specific marketing campaigns to encourage app downloads, which can increase initial costs. This requirement emphasizes the significance of employing effective incentives and engagement strategies to optimize the return on investment in mobile applications and enhance user satisfaction.

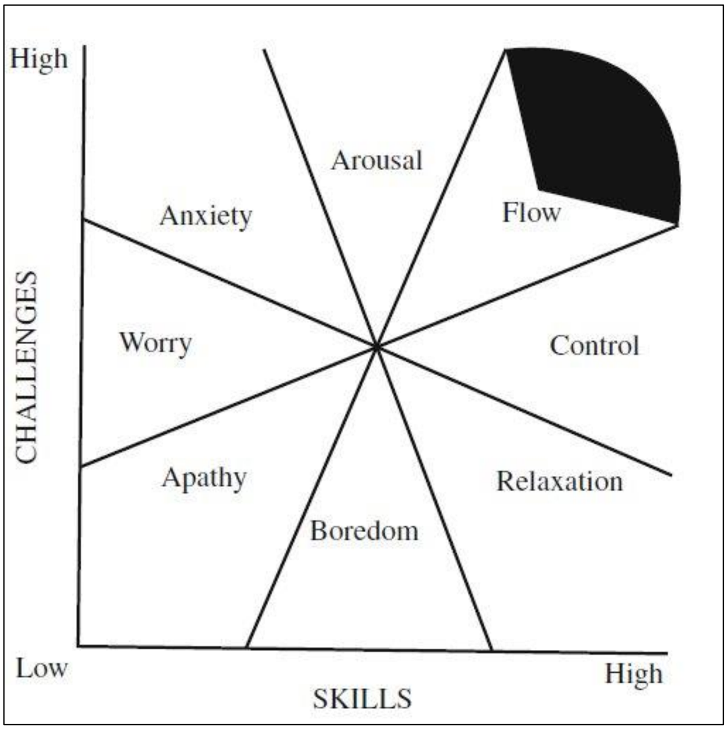
## 2.2 FLOW THEORY

Qualitative research introduced the concept of flow, initially described by Mihaly Csikszentmihalyi in 1975. Csikszentmihalyi’s investigation centered on understanding intrinsic motivation and satisfaction, revealing that individuals achieve peak performance when engaged purely for enjoyment. Interviews with climbers, chess players, composers, dancers, and basketball players highlighted their profound immersion in activities, resulting in heightened concentration and pleasure, devoid of external rewards like money or reputation.

According to Csikszentmihalyi (1975), flow represents a holistic state where individuals are completely absorbed in their current activities, regardless of the task’s nature. This immersive experience explains how individuals enter a concentrated state where irrelevant thoughts and perceptions are filtered out (CSIKSZENTMIHALYI, 1990).

Achieving flow hinges on the balance between an individual’s skills and the challenges presented by the task environment (CSIKSZENTMIHALYI, 1988), as depicted in Figure 1. When skill levels match challenge levels individuals experience flow; imbalances can lead to boredom or anxiety.

Figure 1 – Relationship between challenges and skills for the Flow process




Source: Csikszentmihalyi (1997).

Csikszentmihalyi (1975) emphasizes that flow is not confined to specific activities but is a universal phenomenon. Scholars like Zimanyi and Schüler (2021) and Bolen and Ozen (2020) corroborate that flow occurs across various actions and cultural contexts, ranging from minor to profound intensities.

Research on flow by Pelet et al. (2017) confirms its occurrence in diverse activities such as gaming, reading, or watching movies, underscoring its universality despite differences in task types and social backgrounds (Bolen & Ozen, 2020).

Flow’s components, identified through decades of research, remain stable and consistent, asserting its validity across qualitative and quantitative studies in psychology (MAHNKE et al., 2015). Chart 1 outlines these components, highlighting factors such as Demand-Skill Balance, Clear Goals, and Immediate Feedback, crucial for experiencing flow (CSIKSZENTMIHALYI, 1990).

Chart 1 – Flow Experience Components

| |
|---|
| 1. Demand-Skill Balance |
| 2. Clear Goals |
| 3. Immediate Feedback |
| 4. Concentration on the Task at Hand |
| 5. Control Paradox |
| 6. Fusion of Action and Consciousness |
| 7. Loss of Self-awareness |
| 8. Time Transformation |

Source: Prepared by the author, based on Csikszentmihalyi (1990).

Csikszentmihalyi (1990) categorizes the flow experience into two fundamental dimensions: absorption and fluency. Absorption is characterized by deep concentration, where individuals become fully immersed in an task, losing track of time and their self-awareness. On the other hand, fluency refers to the smooth and automatic nature of actions guided by internal logic, where individuals perform tasks effortlessly and seamlessly.

In the context of Information and Communication Technologies (ICT), flow manifests prominently during various activities, such as online shopping. Users often describe experiencing distinct states of flow while navigating through mobile applications and making purchase decisions (OZKARAA; OZMEN; KIM, 2017; WU; CHIU; CHEN, 2020). This phenomenon highlights how ICT environments can facilitate conditions that promote both absorption and fluency, enhancing user engagement and satisfaction in online interactions.

## 2.3 ONLINE FLOW

The concept of flow has been extensively studied and applied across various disciplines, yielding context-specific definitions (HOFFMAN; NOVAK, 2009). In the realm of online environments, the flow experience is characterized as a state where users engage with technology in a playful and exploratory manner (RODRIGUEZ-ARDURA; MESEGUER-ARTOLA, 2016). This state is marked by continuous interaction facilitated by technological devices, resulting in intrinsic pleasure, loss of self-awareness, and reinforcement of engagement (SHIH; JIN, 2011). Therefore, online flow is understood as a cognitive state prevalent during internet browsing (NOVAK, HOFFMAN; YUNG, 2000; RODRIGUEZ-ARDURA; MESEGUER-ARTOLA, 2020).

Esteban-Millat et al. (2014), Bilgihan et al. (2013), and Rodriguez-Ardura and Meseguer-Artola (2017) characterize online flow as a continuous sequence of responsive interactions facilitated by technological interactivity. Ameen et al. (2021) suggest that individuals experiencing flow during online browsing exhibit heightened states of task, alertness, focus, happiness, satisfaction, and creativity, regardless of the specific task. Consistent with Csikszentmihalyi's findings, these states are accompanied by distortions in user perception of time and self-awareness.

Bolen and Ozen (2020) emphasize that the study of online flow in Information Systems (IS) adopts a multidimensional approach. Over time, researchers have identified several key dimensions shaping user perceptions of online flow. Notably, pleasure, concentration, control, time distortion, and curiosity emerge as recurrent dimensions in systematic studies (AMEEN et al., 2021; HERRANDO et al., 2019; OZKARA et al., 2017; KIM; HAN, 2014). These dimensions are typically assessed quantitatively through surveys, probing users' experiences and sensations during interactions with technology.

Beyond user perceptions, the online flow experience can be evaluated through elements intrinsic to Information and Communication Technologies (ICTs), often intertwined with web design. These factors collectively influence the overall user experience in virtual environments, influencing factors identified as flow enhancers and impacting behaviors such as online purchasing decisions and research outcomes (SALLEH et al., 2019).

Barthelmäs and Keller (2021), building on Csikszentmihalyi's Flow Theory, assert that online flow represents a profoundly positive state of consciousness attained when perceived abilities align with the challenges undertaken. They highlight the importance of clear goals, matched skills, and immediate feedback in fostering engagement during online activities.

King (2003), referencing Novak, Hoffman, and Yung (2000), notes that prerequisites for online flow mirror those in offline contexts. Telepresence, a unique attribute of online environments, enhances user engagement by fostering a sense of immersion and participation in activities.

Achieving flow in online environments necessitates adherence to Csikszentmihalyi’s eight components of the flow experience, which are recognized as fundamental features driving users toward an optimal online flow experience. These components include:

**1) Demand-Skill Balance**

Csikszentmihalyi’s Flow theory describes this concept as the optimal match between task difficulty and user capabilities, crucial for engendering flow. In the online context, tasks should not be challenging, as this can discourage users (BARTHELMÄS; KELLER, 2021). Due to that, this research uses the term demand, not challenge. King (2003) emphasizes the importance of demands aligning with users’ prior knowledge and confidence levels to maintain engagement. Clear guidance on task progression, such as step-by-step instructions, aids users in navigating online applications effectively.

Guo and Poole (2009) caution that overly complex tasks hinder goal achievement and clarity. Simplifying task complexity enhances user experience by facilitating goal attainment and minimizing cognitive load. Hence, in online environments, ensuring tasks are both challenging and achievable optimizes user engagement.

**2) Clear Goals**

Clear goals are pivotal in sustaining user engagement in online interactions (RETTIE, 2001). Users must comprehend the purpose and anticipated outcomes of their actions to remain motivated (KING, 2003). For Barthelmäs and Keller (2021), clarity in task structure and instructions facilitates user understanding and commitment. Rettie (2001) stresses the need for transparent feedback after task completion to reinforce user motivation. Guo and Poole (2009) argue that well-organized online layouts help users concentrate on important tasks, making it easier for them to achieve their goals efficiently.

**3) Immediate Feedback**

Effective feedback mechanisms are integral to user-system interactions in online environments (RETTIE, 2001; KING, 2003). Immediate and precise feedback aids in error correction and goal achievement (BARTHELMÄS; KELLER, 2021). Barta, Flavián, and Gurrea (2022) highlight feedback as a bi-directional process crucial for enhancing user competence and confidence.

Form validations and confirmation messages exemplify clear feedback practices that guide users toward successful interactions (GUO; POOLE, 2009). In complex online systems, clear feedback mitigates user frustration and enhances flow experiences by maintaining user control and clarity.

**4) Concentration on the Task at Hand**

Concentration is essential for achieving flow states during online tasks (FINNERAN; ZHANG, 2003). King (2003) advocates for intuitive and streamlined interface designs that direct user attention and facilitate task immersion. Simple, responsive layouts and clear call-to-action prompts enhance user engagement and minimize distractions (RETTIE, 2001).

User-centric design principles emphasize the prioritization of functionality over aesthetics to optimize user experience (KING, 2003). By focusing on task relevance and responsiveness, online environments promote sustained user attention and immersion, key factors in fostering flow.

**5) Control Paradox**

Empowering users with control over their interactions enhance engagement and satisfaction in online environments (KING, 2003). Customization options, such as editable profiles and content preferences, reinforce user autonomy and reduce anxiety (VAN DER VELDEN; EMAM, 2013). Shibly, Aisbett, and Pires (2014) suggest that perceived choice in decision-making cultivates a sense of control critical for enhancing user experience.

User-friendly interfaces that facilitate error correction and undo actions provide users with a sense of security and ease (BARTA; GURREA; FLAVIÁN, 2022). This paradoxical control dynamic in online systems enhances user trust and commitment, essential for fostering flow experiences.

**6) Fusion of Action and Consciousness**

Flow experiences in online environments are characterized by effortless engagement that immerses users in their tasks (KING, 2003). Herrando, Martínez, and Hoyos (2018) note that emotionally engaging experiences in virtual contexts enhance user memory and satisfaction. Well-designed online systems minimize frustration and promote seamless interaction, improving user experience and satisfaction (HERRANDO et al., 2018).

Flow awareness in users signifies deep engagement and satisfaction with online activities (HERRANDO; MARTÍNEZ, 2018). The integration of user actions with system responsiveness fosters immersive experiences that drive user retention and positive outcomes in online environments.

**7) Loss of Self-awareness**

Effective online systems minimize user self-consciousness by promoting task-focused engagement (BARTA, GURREA; FLAVIÁN, 2022). Engaging in online experiences enables users to transcend self-awareness and concentrate fully on tasks, enhancing flow states (KING, 2003). By creating immersive and engaging environments, online platforms maximize user satisfaction and task efficiency.

**8) Time Transformation**

Flow experiences alter users' perception of time, making tasks seem shorter or longer based on engagement levels (RETTIE, 2001). Pelet, Ettis, and Cowart (2017) observe that engaging virtual tools can distort users' sense of time, emphasizing the immersive nature of online interactions. The duration and frequency of user visits indicate the success of online platforms in sustaining user interest and engagement (PELET et al., 2017).

The examination of the eight components of online flow highlights their crucial role in enhancing user engagement and facilitating seamless digital interactions. By understanding and optimizing these components organizations can effectively cultivate environments that promote user satisfaction and loyalty. In the next section, it will be demonstrated that online flow can be comprehended through three distinct domains, with this thesis focusing on the first domain, the online flow conditions.

### 2.3.1 Online Flow Conditions

According to Novak, Hoffman, and Yung (2000), who were among the first to apply Flow theory concepts to the online context, the elements related to the online flow state can be categorized into three primary areas: (1) Online Flow Conditions, (2) Online Flow Characteristics, and (3) Online Flow Consequences (see Chart 2).

Chart 2 – Areas of Online Flow

| **1. Online Flow Conditions** |
|---|
| Demand-Skill Balance<br>Clear Goals<br>Immediate Feedback |
| **2. Online Flow Characteristics** |
| Concentration on the Task at Hand<br>Control Paradox |
| **3. Online Flow Consequences** |
| Fusion of Action and Consciousness<br>Loss of Self-awareness<br>Time Transformation |

Source: Prepared by the author, based on Novak, Hoffman, and Yung (2000).

According to the literature (NOVAK; HOFFMAN; YUNG, 2000; PEIFER; ENGESER, 2021), these three areas can be understood as follows:

1. **Online Flow Conditions:** These refer to the pre-existing elements or conditions within the digital environment that facilitate the emergence of a flow state in users. This includes aspects such as clarity of goals, availability of immediate feedback, balance between challenges and skills.

2. **Online Flow Characteristics:** These are the intrinsic qualities of the flow state experienced by users during online interactions. This may involve deep immersion in the task, intense focus, and a sense of ecstasy or complete absorption in the task at hand. Flow characteristics describe how users perceive and experience this mental state while online.

3. **Online Flow Consequences:** This pertains to the positive outcomes or effects associated with the experience of online flow. This can include increased user satisfaction, greater efficiency in task performance, enhanced learning and personal development, as well as a higher likelihood of repeating online activities in the future.

These areas provide a comprehensive framework for understanding and studying how flow manifests and impacts users in digital contexts. Each area contributes to a deeper understanding of the underlying mechanisms of the online flow state and its practical implications in user experience and digital interface design (NOVAK; HOFFMAN; YUNG, 2000).

King (2003) points out the nuanced constraints of online flow, positing that users achieve flow when their interactions within digital environments remain seamless, characterized by prompt responses, clear feedback, and minimal distractions. The alignment of user skills with task challenges is pivotal for flow occurrence, prompting designers to prioritize elements facilitating these conditions.

According to Novak, Hoffman, and Yung (2000), users typically visit online platforms with predefined goals, such as gathering product information, guiding their task-oriented interactions, and focusing attention on pertinent content.

Barthelmäs and Keller (2021) contend that key conditions to flow hinge on perceived alignment between user skills and task demands, emphasizing the necessity of clear task structures and immediate, unequivocal feedback to accurately gauge performance. They propose that these elements can be consolidated into the perception of skill-task adequacy, underlining the critical role of clear goals and feedback in fostering flow.

Thus, web designers can effectively shape flow experiences by concentrating on elements classified under Area 1 (online flow conditions): (1) Demand-Skill Balance, (2) Clear Goals, and (3) Immediate Feedback. These elements, as identified in the literature (NOVAK; HOFFMAN; YUNG, 2000; BARTHELMÄS; KELLER, 2021; PEIFER; ENGESER, 2021) as prerequisites for flow, delineate the structural framework of online environments guiding user navigation. Meanwhile, characteristics and consequences categorized under online flow characteristics and online flow consequences represent inherent attributes and outcomes perceived by individuals experiencing flow states. Notably, ensuring the presence of these foundational conditions heightens the likelihood of realizing factors associated with the other two areas.

Consequently, focusing on online flow conditions is pivotal in comprehending strategies to foster flow states within digital environments. This approach aligns with the overarching objective of this thesis: to establish a scale for assessing online flow conditions and apply it to representative mobile shopping applications in Brazil and Spain.

In the subsequent section, a segment will commence focusing on the delineation of this proposed scale.

## 2.4 SCALE DEVELOPMENT

According to Gil (2002), scales are essential tools designed to objectively measure latent traits, opinions, or attitudes through a structured arrangement of items. These items are categorized into graded responses or alternatives, emphasizing the scale’s ability to discriminate between different measures.

DeVellis (2017) suggests that new scales should be developed when existing literature lacks adequate measures for the target latent trait. Therefore, a thorough literature review is crucial before initiating scale development to ensure alignment with relevant theories and to identify potential limitations.

Understanding complex concepts is facilitated through scale development, as noted by Babbie (2005), who highlights three benefits: first, exploring variable concepts enhances understanding; second, multiple perspectives capture variations more precisely; and third, scales streamline data, allowing for comparability via numerical scores that indicate ordinal positions.

Embretson and Reise (2013) highlight the role of scales in representing objects or people numerically, thereby illustrating relationships among represented factors. This representation grants scales distinguishing power, order, and rationale for studying objects.

Wilson (2005) recommends creating a construct map to identify factors contributing to the analyzed latent trait. This map aids in pinpointing aspects or behaviors that reflect different levels of the trait, guiding item development to ensure comprehensive coverage.

The interpretation of scale scores is critical, as emphasized by Cronbach (1996), who outlines three procedures: Reference to the Standard compares scores to a normative group, indicating relative positions; Reference to Content interprets scores based on representative task samples; and Reference to the Criterion correlates scores with external criteria to predict performance levels.

Moreover, Embretson and Reise (2013) stress the importance of comparing scores and utilizing appropriate statistics to summarize them effectively. They reiterate that numeric representations in scales facilitate empirical relationships between objects or people, enhancing the scale's utility in various contexts.

In the pursuit of a comprehensive understanding of latent traits, DeVellis (2017) encourages exploring alternative methodologies, such as user testing. In this study, eye-tracking user tests were employed to gain insights into participants' perceptions of the latent trait. Eye-tracking technology captures gaze patterns during activities, providing valuable data for assessing relevant aspects of participant perception.

### 2.4.1 Types of Scales

Scales are typically classified into four types based on their construction rigor and the nature of the variables they measure. The most common types are nominal, ordinal, interval, and ratio (also known as proportional) scales (EMBRETSON; REISE, 2013; PASQUALI, 1997).

The nominal scale uses numbers or symbols to represent categories of objects or people. These elements are simply labels and do not imply any order or ratio. This is the most basic level of measurement, focusing on grouping and classifying elements into distinct sets without defining any order relationship (EMBRETSON; REISE, 2013).

The ordinal scale not only identifies people, objects, or categories but also orders them along an underlying dimension. For example, it can be used to rank preferences or importance among items. Ordinal data retain all properties of nominal data and add a numerical order. This scale allows transformations that preserve data distinction and order, with natural ordering in categories being its main feature. Hence, it is also known as a scale by positions (EMBRETSON; REISE, 2013).

The interval scale quantitatively records phenomena, measuring them in terms of specific intensity relative to an arbitrary zero point. Interval scales classify and order data while also defining equal intervals between points on the scale. This allows for linear transformations where scores can be multiplied by a constant or have a constant added. Statistical parameters suitable for interval scales include the mean, standard deviation, and correlations (EMBRETSON; REISE, 2013).

The ratio scale is the most comprehensive and sophisticated. It includes a fixed and absolute zero point, representing a true minimum. On this scale, measurements are defined by the distance from the zero point, expressed in predefined units. Ratio scales support the calculation of meaningful ratios between scores. The only permissible transformation is multiplication by a constant that maintains the ratio between scores (EMBRETSON; REISE, 2013).

Data obtained from any scale type can be analyzed using descriptive statistics. The distinction between these four types of scales has significant implications for the complexity of statistical analysis. For inferential statistics, which assess the systematic nature of relationships, nominal and ordinal scale data can be analyzed using both parametric and non-parametric tests. Data from interval and ratio scales allow for these analyses and support parametric procedures as well (EMBRETSON; REISE, 2013; PASQUALI, 1997).

### 2.4.2 Item Crafting

The items of a scale are designed to manifest the latent trait under analysis, thereby operationalizing the theoretical definition of the latent trait being measured. Each item must individually reflect the latent trait of interest (BERTONCINI; SERAFIM; BORBA, 2023; DEVELLIS, 2017).

According to Wilson (2005), the initial step in item generation involves defining the format, such as multiple-choice, Likert, dichotomous responses, or others. This choice allows researchers to tailor items to their intended response format. The item set should be generated based on the levels of the latent trait identified in the construct map to ensure all levels are represented on the scale.

DeVellis (2017) emphasizes two primary considerations at the outset: 1) Avoiding redundancy in items – while grammatical redundancy should be eliminated, items should evaluate the latent trait from various perspectives to enhance reliability; and 2) Managing item specificity – overly specific items can introduce redundancy and compromise scale dimensionality. DeVellis' (2017) and Wilson's (2005) perspectives can be integrated to refine item definitions effectively.

Following these initial considerations, DeVellis (2017) outlines four steps for item creation: 1) Initially focus less on item quality criteria; 2) Develop a larger set of items than needed for the final scale; 3) Review items for quality after their creation; and 4) Consider including items with reversed meanings to detect response biases, although this should be balanced against potential participant confusion in lengthy questionnaires.

Once the initial set of items is defined, determining the number of response categories becomes crucial, particularly for closed-response items like Likert scales, which are widely used for measuring opinions or attitudes in a graduated manner (DEVELLIS, 2017), or checklists assessing the presence or absence of specific characteristics or situations (TEZZA; BORNIA; ANDRADE, 2011).

Checklists are effective tools for evaluating a wide range of aspects. In virtual contexts, they can identify interface elements through predefined items. Assessment via checklists offers advantages such as reduced evaluation costs, ease of identifying interface elements, systematic evaluation ensuring consistent results across assessors, and applicability by non-specialists due to embedded checklist knowledge (KHAJOUEI; GOHARI; MIRZAEE, 2018; FORD et al., 2022). Checklists are widely adopted in assessing online interfaces and systematically categorizing web design elements in virtual environments. Thus, considering these factors, this thesis advances a checklist-based scale template, which will be detailed in subsequent sections.

## 2.5 EYE-TRACKING TECHNIQUE

Developing measurement scales should not rely solely on literature but also on empirical methods like eye-tracking, which emerges as an innovative way to support item development (HOLMQVIST et al., 2011). An eye-tracking user testing can identify which elements attract more attention, validate item clarity, measure response time and engagement, map areas of interest, analyze attention flow, compare different groups, and detect visual biases. By tracking participants' eye movements while interacting with the scale, researchers can adjust its design to enhance clarity, fluidity, and measurement precision. This ensures effective capture of user responses while minimizing potential biases (DUCHOWSKI, 2017).

The origins of eye-tracking can be traced to Charles Bell, who in 1823 assigned movement control to the brain, classified eye movements, and described the effect of eye movement according to visual guidelines. Historically, eye-tracking was expensive and labor-intensive, requiring the researcher to directly observe and catalog each participant’s behavior. This served as a barrier for many years and limited the rate of research. However, improvements in eye-tracking technology have made this technique more accessible and practical for both the participant and the researcher. Video-based eye trackers can determine the direction of gaze very accurately by measuring the position of an infrared light’s reflection on the cornea relative to the pupil (DUCHOWSKI, 2017; CARTER; LUKE, 2020).

Eye-tracking is a process in which eye positions and movements are captured in response to a visual stimulus (NIELSEN; PERNICE, 2010). The eye-tracker – the research device used – is capable of following the trail where a person is looking and capturing eye movements. Its application can be seen in the context of assessing human attentional behavior. Eye tracking can help measure attentional behavior. In addition, eye-tracking can help observe, analyze, and understand human performance's cognitive, attentional, and visual aspects. Notably, eye-tracking has been used to evaluate the usability of newly developed equipment and systems (DUCHOWSKI, 2017).

Thus, Barreto (2012) points out that the concept of eye-tracking is directly linked to the measurement and recording of an individual’s eye movements in the face of stimuli in a real or controlled environment, thus determining the areas where they fix their attention, for how long and in what order they follow their visual analysis. The author points out that the relevance of studying eye movements is based on the idea that the elements a person visualizes indicate the current thinking prevalent in their cognitive processes. Therefore, the recording of eye movements provides a dynamic scheme for the points of attention in a given field of vision. The measurement of other elements, such as those moments when the eyes are relatively fixed, capturing information, can also reveal the relevance of certain objects viewed.

The main information provided by eye-tracking are (1) gaze fixation, the instants in which the eyes are relatively inactive, processing information, being able to measure the quantity, time, and dispersion of fixations; (2) saccades, rapid eye movements captured between fixations, providing an idea of visual hierarchy in an environment; and (3) blinks, making it possible to measure their quantity and duration, and may be indicative of the mental load required in online environments (MCINTIRE et al., 2014).

Where we look, and for how long, is influenced by cognitive processes beyond attention, such as perception, memory, language, and decision-making. While the link between the eye and the mind is not absolute (STEINDORF; RUMMEL, 2020), it is generally true that the eyes reflect the mental processing of whatever we are looking at any given time. This makes eye-tracking universally applicable to most research that explores mental processes. Due to its high temporal sensitivity, eye tracking can provide moment-by-moment insight into developing cognition, rather than simply revealing the outcome.

In addition, eye movements are largely outside of conscious control, meaning that individuals can choose what to look at and when the finer details of that movement are largely reflective. This means that eye-tracking can access unconscious information in processing (CLARKE et al., 2017).

Studies involving eye-tracking can be seen in several areas. Merino et al. (2018) sought to identify the focus of visual attention in people with motor disabilities using eye-tracking equipment. Cavalcante et al. (2020) proposed to use the analysis of eye tracking of students during computerized educational activities as a strategy for inclusive educational assessment of students with Autism Spectrum Disorder. And more recently, Gallina et al. (2022) developed a study, with eye tracking, to identify the consumer's reaction to observation and choice when exposed to a retail setting. Thus, the application of eye-tracking is broad and capable of covering several situations.

### 2.5.1 Eye-Tracking Overview in the Electronic Commerce Area

Mobile devices have significantly driven the growth of e-commerce (GREWAL et al., 2018) while eye-tracking technology has emerged as a valuable tool for objectively studying the visual behavior of online consumers (KING et al., 2019). Designing eye-tracking user testing to reflect natural user interactions with minimal interruptions is crucial for analyzing user experiences in mobile apps (HUDDLESTON et al., 2015).

Past studies have utilized eye-tracking to investigate behavioral differences across screens of varying sizes (TONKIN et al., 2011), track the purchasing processes on fashion (GUO et al., 2015) and grocery websites (Benn et al., 2015), and assess visual appeal in e-commerce settings (PAPPAS et al., 2018). Research findings indicate that visual attention significantly influences consumer evaluations of products, services (LADEIRA et al., 2019), and online marketing advertisements (KING et al., 2019). Additionally, eye-tracking has been instrumental in capturing user attention towards interface elements and images on e-commerce platforms (HAESNER et al., 2018; VIDYAPU et al., 2019), examining factors such as the impact of product image backgrounds on consumer attention (WANG et al., 2019), and the processing of online reviews (Maslowska et al., 2020). However, studies specifically focusing on search behavior in e-commerce, especially on mobile platforms, remain limited (ZILLICH; KESSLER, 2019).

Recent studies have employed eye-tracking techniques to measure the visual appeal of online environments, providing insights into users' gaze patterns when navigating websites or apps (DJAMASBI; SIEGEL; TULLIS, 2010). These studies contribute to a deeper understanding of user behavior within virtual spaces, translating user perceptions of web design elements into measurable reactions. By analyzing users' gaze behavior, aspects such as app design features, visual attention, and user attitudes can be effectively explored (DJAMASBI; SIEGEL; TULLIS, 2010; HWANG; LEE, 2017).

In the realm of mobile applications, the effectiveness of online interfaces hinges on various factors including image size, search functionalities, and textual content. Gaze fixation duration serves as a metric to identify user focal points within any online interface. The journey users undertake from entering a virtual environment to accomplishing their goals is laden with interface elements shaped by web design principles (SECKLER; OPWIS; TUCH, 2015). While these elements are pivotal in app design, limited research has explored their correlation with gaze behavior, which could empirically inform app design improvements or infer users' perceptions of online flow based on their visual interactions.

## 2.6 ITEM RESPONSE THEORY

Item Response Theory (IRT) is a statistical framework widely used in developing and analyzing assessment instruments such as tests, scales, and questionnaires. It focuses on the relationship between respondents' answers and item characteristics, enabling a more precise evaluation of latent abilities or traits (EMBRETSON; REISE, 2013). According to Foster, Min, and Zickar (2017), IRT encompasses a variety of models and techniques capable of analyzing item properties, conducting computerized adaptive testing, and assessing differential item functioning.

As highlighted by Embretson and Reise (2013), IRT's advantages include the independence of item statistics from the sample used for estimation, respondent scores being independent of scale difficulty, and the ability to assess reliability without the need for parallel tests. However, using IRT models requires strict assumptions and larger sample sizes for accurate parameter estimation, which can be considered practical limitations.

Choosing the appropriate IRT model depends on factors such as the nature of the data, model fit criteria, and dimensionality of the data (TAVARES, ANDRADE, PEREIRA, 2004; FOSTER, MIN, ZICKAR, 2017). Model selection hinges on considerations like item type, number of latent variables, and population characteristics under investigation.

The application of IRT begins with assessing the dimensionality of the dataset, which determines the number of factors required to explain variability and test hypotheses (HAIR et al., 2019). Dimensionality assessment involves methods such as examining the correlation matrix, factor analysis, nonlinear factor models (EMBRETSON; REISE, 2013), and using IRT indices for model adequacy (PASQUALI, 2003). These methods help confirm the underlying structure of the dataset and ensure the chosen IRT model fits appropriately.

In this thesis, the proposed scale is based on a three-dimensional structure derived from literature, user testing validation, and expert input. Confirming the dimensionality involves rigorous statistical procedures to validate this structure, aligning with theoretical expectations and ensuring the model’s robustness across different populations.

IRT comprises various mathematical models tailored to different research needs. Hence, this study briefly discusses both unidimensional and multidimensional IRT models, emphasizing the latter’s suitability for the proposed scale structure.

### 2.6.1 Unidimensional Item Response Theory

According to Reckase (2009), the basic representation of the IRT model is described by the following probability equation:

$$P(U = u|\theta) = f(\theta, \eta, u) \tag{1}$$

Where $\theta$ represents the parameter describing the characteristics of the respondent, $\eta$ represents the vector of parameters describing the item characteristics, $U$ represents the variable response or score on a specific test, and $u$ is a response possibility in dichotomous models: 1 (positive response) or 0 (negative response). Additionally, $f$ is the function that describes the relationship between the parameters and the probability of a response $P\ (U = u)$.

The most basic and commonly used unidimensional models are the Rasch Model (RASCH, 1960) and its extension, the one-parameter logistic model, 1PL (LORD, 1952; LORD & NOVICK, 1968). These models focus only on one main parameter, the item difficulty. This indicates how easy or difficult it is for the evaluated apps to possess the characteristics expressed by each item.

The unidimensional two-parameter logistic model (2PL), introduced by Birnbaum (1968), was developed to accommodate variation in item discrimination. This model incorporates a discrimination parameter in addition to the difficulty parameter, allowing for the assessment of each item's ability to differentiate between varying levels of the underlying latent trait and to identify the primary dimension it measures. The unidimensional 2PL model permits items to differ in both difficulty and discrimination levels, and it does not require the assumption of a single trait being assessed. This model is represented as follows:

$$P(U_{ij} = 1 \mid \theta_j, a_i, b_i) = \frac{1}{1+e^{-a_i(\theta_j - b_i)}} \qquad (2)$$

Where:

$P\ (U_{ij} = 1\ |\theta_j, a_i, b_i)$ = probability of app $j$ to have the characteristic from item $i$ ;

$U_{ij}$ = response from app $j$ to item $i$ ;

$\theta_j$ = latent trait for app $j$ ;

$a_i$ = discrimination parameter of item $i$ ;

$b_i$ = difficulty parameter of item $i$.

Ideal for dichotomous items, the unidimensional 2PL model estimates a single discrimination parameter for each item and not a general value of discrimination. Thus, the model provides estimates of the quality of each item. A scale adjusted through this model will have items with distinct levels of quality, and different contributions regarding the general provision of information (DE AYALA, 2009).

In the unidimensional 2PL model, items that do not reach a minimum discrimination value should be eliminated; and probability curves of different items may cross due to the inclination of the curves (REACKASE, 2009). In addition, this modeling is capable of estimating difficulty parameters for each item that composes the latent trait scale (ANDRADE; TAVARES, TAVARES, VALLE, 2000).

It is important to highlight that this model will initially be applied to each dimension derived from the measurement scale of online flow conditions, treating these dimensions as distinct unidimensional scales. Subsequently, the development of a multidimensional model relevant to the data in this study will be introduced, as it is hypothesized that the multidimensional approach may better suit the data.

### 2.6.2 Multidimensional Item Response Theory

Multidimensional Item Response Theory (MIRT) models are crucial for assessing the complexity inherent in measurement contexts, recognizing that each item on a scale may correspond to multiple latent dimensions (Wilson & Hoskens, 2005). This approach evaluates individuals not just on a single trait but across several interrelated dimensions, which can either be independent or correlated with each other. In the realm of information and communication technology (ICT) research, this capability is particularly advantageous as it allows for comprehensive modeling of interactions between various latent traits and observed variables (RECKASE, 2009; EMBRETSON; REISE, 2013).

In ICT assessments, diverse dimensions such as user engagement, system usability, and technological proficiency play pivotal roles, contributing uniquely to the overall evaluation framework. Traditional unidimensional methods often oversimplify these multidimensional constructs, whereas MIRT offers a more nuanced perspective by simultaneously considering multiple dimensions and their interdependencies. This approach enhances measurement precision by accounting for how each dimension influences item responses, thereby providing a more accurate portrayal of individuals' abilities and attitudes within ICT contexts (EMBRETSON; REISE, 2013; SAMEJIMA, 1969).

MIRT's ability to handle complex constructs and its integration with advanced statistical tools like R software empowers researchers to conduct sophisticated analyses that reveal intricate patterns and relationships in ICT research. By facilitating the development of robust measurement instruments tailored to diverse ICT applications, from user interface design to system performance evaluation, MIRT serves as a powerful methodological framework in advancing our understanding and assessment of ICT systems and applications (RECKASE, 2009).

Reckase (2009) describes the multidimensional two-parameter logistic model as follows:

$$P(U_{ij} = 1 \mid \theta_{jk}, a_{ik}, d_i) = \frac{1}{1+e^{-\left(\sum_{k=1}^{n} a_{ik}\theta_{jk}+d_i\right)}} \qquad (3)$$

Where:

$P\,(U_{ij} = 1 \mid \theta_{jk}, a_{ik}, d_i)$ = probability of app $j$ to have the characteristic of item $i$ ;

$k$ = number of latent trait dimensions, varying from 1 to $n$ dimensions;

$a_{ik}$ = discrimination parameter for item $i$ on dimension $k$ ;

$\theta_{jk}$ = latent trait for app $j$ on dimension $k$ ;

$d_i$ = scalar difficulty parameter of item $i$.

Since the discrimination parameters $a_{ik}$ form a vector set, it can be represented by a single encompassing vector known as the Multidimensional Discrimination Parameter (MDIS) of item $i$. The equation governing this parameter is expressed as follows (RECKASE, 2009), where $k$ is the number of latent trait dimensions varying from 1 to $n$ dimensions:

$$MDIS_i = \sqrt{\sum_{k=1}^{n} a_{ik}^2} \qquad (4)$$

Utilizing the MDIS parameter, it is possible to compute the multidimensional difficulty parameter (MDIF) of item $i$, reflecting the vector's spatial positioning (both direction and magnitude). The equation governing this parameter is expressed as follows (RECKASE, 2009):

$$MDIF_i = \frac{d_i}{MDIS_i} \tag{5}$$

A positive value of $MDIF_i$ indicates a challenging characteristic for an app to have, while a negative $MDIF_i$ denotes easier characteristics. The magnitude of $MDIF_i$ corresponds to the distance from the origin to the point of maximum incline along the vector $MDIS_i$, representing the difficulty level of the item. It is imperative to note that this interpretation of $MDIF_i$ pertains solely to the direction specified by the vector $MDIS_i$ (DE AYALA, 2009).

Another widely known multidimensional IRT model that fits the characteristics of the scale defined in this thesis is the Bifactor model (GIBBONS; HEDEKER, 1992). This model is an extension of traditional IRT models that seeks to explain individuals' responses to a set of test items by considering the presence of two underlying latent factors. These factors can represent different abilities or traits that influence an individual's performance on the test items.

The Bifactor model allows for the assessment of the presence of two dimensions of the latent trait. It is particularly useful when it is believed that the scale being evaluated is influenced by more than one dimension (REISE; MOORE; HAVILAND, 2010). This is the case for the developed scale, which measures the conditions of online flow in shopping apps through three dimensions: Demand-Skill Balance, Clear Goals, and Immediate Feedback.

It is noteworthy that the use of the Bifactor model can assist in constructing test items that better discriminate between different levels of ability in both dimensions. This can result in more balanced tests that are more representative of the abilities intended to be measured (REISE; MOORE; HAVILAND, 2010).

Gibbons and Hedeker (1992) describe the Bifactor model according to the following equation:

$$P(U_{ij} = 1 \mid \theta_{Gj}, \theta_{Sjk}, a_{Gi}, a_{Sik}, d_i) = \frac{1}{1+e^{-\left(a_{Gi}\theta_{Gj}+\sum_{k=1}^{n} a_{Sik}\theta_{Sjk}-d_i\right)}} \quad (6)$$

Where:

$P(U_{ij} = 1 \mid \theta_{Gj}, \theta_{Sjk}, a_{Gi}, a_{Sik}, d_i) =$ probability that app $j$ has the characteristic of item $i$ ;

$k$ = number of latent trait dimensions, varying from 1 to $n$ dimensions;

$\theta_{Gj}$ = general latent trait for app $j$ ;

$\theta_{Sjk}$ = specific latent trait for app $j$ on dimension $k$ ;

$a_{Gi}$ = discrimination parameter of item $i$ on the general factor;

$a_{Sik}$ = discrimination parameter of item $i$ on dimension $k$;

$d_i$ = scalar difficulty parameter of item $i$.

According to Gibbons et al. (2007), the Bifactor model is relevant whenever items share a common characteristic. The presence of item subgroups typically introduces meaningful associations in testing that cannot be captured with the complete assignment of loadings to the general factor alone. Additionally, according to the authors, this separation of factors improves estimation accuracy.

Gibbons et al. (2007) further assert that the constraints of the Bifactor model led to greater simplification of probability equations by requiring assessment of only two integral dimensions, thereby enabling analyses of models with a large number of dimensions, permitting conditional dependence among identified item subgroups, and often providing a more parsimonious solution than full-information factor analysis due to its lack of constraints.

## 2.7 CHAPTER SUMMARY

In this chapter, a paramount focus was placed on the detailed establishment of the theoretical-empirical underpinnings of the thesis across a series of distinct subsections. The examination of M-commerce (Section 2.1) intricately delved into its operational definition and intricate interplay with mobile applications, thus illuminating its pivotal role within the dynamic expanse of the digital market sphere. Following this academic elucidation, a comprehensive discourse unfurled about Csikszentmihalyi's Flow Theory (Section 2.2), thus laying a fundamental groundwork for comprehending the nuanced concept of flow within the realm of digital interfaces. Subsequently, the discourse seamlessly progressed to scrutinize online flow (Section 2.3), wherein the focal point was directed towards discerning the involved web design parameters that support a continuous and captivating online user experience. The critical development basis of a precisely constructed scale (Section 2.4) aimed at quantifying these online flow conditions was underscored, with the concomitant inclusion of insightful eye-tracking user tests (Section 2.5) serving as instrumental tools in aiding the refinement of said scale. Furthermore, the cautious application of item response theory (Section 2.6) in the evaluative process of the scale accentuated the fundamental import of multidimensional models in dissecting and comprehending user experiences within the intricate web-based background. This chapter intricately melded theoretical frameworks with empirical queries, thereby engendering a robust and academic foundation underlying the thesis.

# 3 EPISTEMOLOGICAL AND METHODOLOGICAL PROCEDURES

This chapter presents the epistemological research assumptions, study paradigm, methodological procedures, and conceptual analysis model. Also presents both user tests carried out in Brazil and Spain, and the scale definition, highlighting the items justification, expert analysis, pre-test, population, sample, and data collection techniques used.

## 3.1 EPISTEMOLOGY, PARADIGM, AND METHODOLOGY

Epistemological definitions deal with how the researcher treats reality in each research context (MAY, 2011; JONKER; PENNIK, 2010). In the study here, the objectivity of the facts is admitted, that is, the social facts exist free from the perception of individuals and the researcher’s actions. And this fact ends up characterizing an epistemology focused on the positivist line. Regarding social sciences, positivist epistemology seeks to clarify individuals’ behaviors in the most objective way possible (RIBEIRO JUNIOR, 1996).

Thus, the researcher will position himself as an external observer, who analyzes the phenomenon and describes it objectively (MAY, 2011; RIBEIRO JUNIOR, 1996). In the positivist view, human behavior is elucidated in terms of cause and effect.

Positivism states that one can only believe that a theory is appropriate if it is demonstrated by valid scientific methods (RIBEIRO JUNIOR, 1996). Thus, research data are factors derived from the social environment and the reactions of individuals to this environment (MAY, 2011).

Paradigms in science approach systems of thought that present visions of reality (MORGAN, 1980). According to May (2011), when focused on social sciences, these systems are broad and open, so they allow researchers to make comparisons between existing paradigms, to evaluate their respective strengths and weaknesses. In the field of social and organizational theories, the role of paradigms as a view of social reality was analyzed by Burrell and Morgan (1979), who suggested four paradigms: Functionalist; Interpretive; Radical Humanist; and Radical Structuralist.

Because of this classification, this research is part of the functionalist paradigm, which, as exposed by Morgan (1980), believes that society is real and concrete, and has a systemic character that can produce a state of ordered and regulated elements. The ontological assumptions in the functionalist paradigm indicate a belief in the possibility of achieving an objective and naked social science of values. The scientist acting as an external factor approaches the phenomenon only through rigor and scientific method techniques (MORGAN, 1980).

In short, the functionalist paradigm is regulatory and pragmatic, an understanding of society generating useful empirical knowledge. Morgan (1980) claims that this paradigm assumes that human behavior is always connected to the real world, full of concrete and tangible social relationships.

Methodological procedures are described in terms of addressing the research problem and method of investigation (QUIVY; CAMPENHOUDT, 1998; MAY, 2011). As far as the approach is concerned, it refers to how the problem is addressed – quantitatively; qualitatively; or through mixed method, which combines quantitative and qualitative approaches – (MAY, 2011). The method consists of the methodological procedure’s formalization, describing the paths adopted by the researcher to analyze the phenomenon or domain of interest (QUIVY; CAMPENHOUDT, 1998).

Regarding the approach, one can classify research as quantitative, qualitative, or mixed methods (CRESWELL, 2007). Given this classification, this research adopts a mixed method, as it is considered to be the most appropriate way to approach the problem and achieve the study objectives.

To help the development of the specific objective that deals with the definition of a scale to measure the degree of online flow conditions in shopping apps, and in addition to the theoretical-empirical foundation, an eye-tacking user tests was carried out, in Brazil and Spain – further on, its steps will be explained. After the creation of these items, they were presented to experts in the area, and web designers were able to give an opinion on the developed scale. And later then will be held soon pre-test.

After the scale was created, it was applied to a sample of shopping apps available in the Brazilian and Spanish app stores. The answers will be obtained through the manual collection. The evaluator accesses the app that will be analyzed and responds, after quick navigation through the online environment, whether it has the characteristic under evaluation or not.

## 3.2 RESEARCH STEPS OPERATIONALIZATION

Studies involving the formulation of scales emphasize the need to highlight what will be measured. The latent trait definition helps the researcher to know the limits of the object of study and avoids the measurement of something that is not desired. In addition to the definition, the researcher needs to highlight the stages and processes practiced during the evaluation of the phenomenon; thus, how the latent trait, theoretically defined, was operationalized (DEVELLIS, 2017).

This study will analyze the latent trace of Online Flow Conditions, in mobile shopping applications, therefore, the focus is given to the level of m-commerce. Online flow refers to the state that arises during online browsing, being characterized by a continuous sequence of responses facilitated by interactivity with the mobile device used. The scale developed to measure this latent trait is a theoretical measure, since it analyzes the objective of online flow characteristics in the apps' interface; variables cannot be measured directly (DEVELLIS, 2017).

King (2003) evaluates elements of web design that enable flow in online environments and based on Csikszentmihalyi, highlights eight components, evidenced in item 2.2 of this study. Novak, Hoffman, and Yung (2000), pioneers of online flow analysis, separate these components into areas to better understand how to analyze them.

The online streaming experience takes place under a narrow set of situations. King (2003) states that users can experience the flow only if their travels through the online environment seem uninterrupted, with instant feedback and understandable goals. In addition, the author states that users who experience online flow feel that their knowledge satisfies the availability challenges required. In this context, the app developer should focus on the components of the online stream that precede or cause the flow. Since the design of cyberspace can only be influenced by elements that precede and cause flow.

Users visit online environments with pre-existing goals, and as these goals evolve, as the user completes tasks, their attention becomes drawn to an original task (KING, 2003; NOVAK; HOFFMAN; YUNG, 2000). Therefore, the main elements that designers can control come from the flow conditions area. The other components are intrinsic characteristics or consequences of the flow, factors that web design does not directly project. However, if the online environment has the conditional factors reported, these online flow components will be achieved.

The study here in question aims to the online flow conditions – constituted by the components of (1) Demand-Skill Balance, (2) Clear Goals, and (3) Immediate Feedback – in shopping mobile apps, and it emerges as a focus for the development of the items that will make up the measurement scale.

Briefly, after defining the research question and objectives, the study methodology proposed here followed the order: (1) conducting a bibliography search in databases; (2) carrying out two eye-tracking user tests to obtain data that assist in the elaboration of the measurement scale; (3) definition of the scale items, with the help of previous literature and the results of the eye-tracking user tests; (4) presentation of the scale to five specialists and scale correction according to their notes; (5) applying a short-term pre-test; (6) application of the scale to the shopping app sample through manual data collection, and, finally (7) the scales statistic evaluation, focusing in the IRT suitability check. Figure 2 highlights those steps through a visual model.

Figure 2 – Research steps operationalization

| | | |
|---|---|---|
| Research Question | How to measure online flow conditions in shopping mobile applications? | |
| General Objective | Define a scale for measuring online flow conditions and to apply this measurement to a sample of mobile shopping applications from Brazil and Spain | |
| Specific Objectives | 1) Elucidate the dimensions of Online Flow Conditions through a comprehensive literature review;<br>2) Build items based on the literature to measure the degree of Online Flow Conditions in mobile shopping applications, supplemented by the implementation of eye-tracking user tests among users;<br>3) Analyze the scale dimensionality and evaluate the suitability of relevant item response theory models;<br>4) Compare the proficiency levels in online flow conditions of the shopping applications available in Brazil and Spain. | |
| Systematic Search | 1) SCOPUS.<br>2) EBSCO.<br>3) Web of Science (including SCIELO). | |
| Eye-tracking user testing | 18 shopping app users | 9 participants in Florianopolis, Brazil<br>9 participants in Barcelona, Spain |
| Scale Definition – 47 items | | |
| Scale Experts Analysis | 1) Professor in a Design department and PhD in Mechanical Engineering.<br>2) Professor in an Administration department and PhD in Administration.<br>3) Professor in a Design department and PhD in Design.<br>4) Professor in a Graphic Design department and PhD in Design.<br>5) Professor in a Department of Languages and Computer Systems and PhD in Computer Science. | |
| Scale Pre-test | | |
| Scale Application (Data Collect) | 200 Shopping Apps | 100 shopping apps available in Brazil<br>100 shopping apps available in Spain |
| Scale Statistical Evaluations | | |

Source: Prepared by the author, 2024.

## 3.3 EYE-TRACKING USER TESTING

To support the development of items comprising the online flow measurement scale, two eye-tracking user tests were conducted: one in Brazil and another in Spain. This methodological approach was employed to enhance the robustness of scale construction. Through these user tests, it was possible to achieve several purposes: (1) identifying pertinent elements by determining which components of the online interface attract the most user attention, (2) assessing cognitive load to pinpoint interface areas that impose higher cognitive demands on users, thereby highlighting potential difficulties, (3) validating scale items by observing where users direct their gaze when responding to flow-related questions, ensuring these items adequately capture desired aspects, and (4) comparing user groups to discern specific enhancements tailored to different segments (DUCHOWSKI, 2017).

Therefore, these eye-tracking user tests provide valuable insights into user interaction with online interfaces, elucidate factors influencing flow perception, and validate scale items that effectively measure these aspects. Integrating these findings will contribute to the development of a more accurate and reliable measurement scale for assessing online flow conditions.

### 3.3.1 Brazilian Eye-Tracking User Testing

The first user test was carried out in Brazil, in August 2023, and consisted of a simulation of the use of a shopping app. It was developed at the Design Management Center and Design and Usability Laboratory (NGD/LDU), at the Federal University of Santa Catarina (UFSC), in Florianopolis. Concerning population, it consists of users of online shopping apps, from Brazil. And, regarding the participants, nine volunteers were selected. The participants were users with previous experience in the use of online shopping environments, of both sexes. They were divided into three groups, to observe the divergences that exist in the different age groups:

- Group 1 with people aged from 18 to 30 years, young people;
- Group 2 with people aged from 31 to 49 years, middle-aged;
- Group 3 with people aged from 50 years or more, older.

This division into groups was made to make a comparison between the age groups regarding adaptability to the elements of online flow conditions.

The object of research was the shopping app Shopee. Which, according to the State of Mobile report (2022), produced by Data.ai in partnership with Similarweb, was the most popular shopping app in Brazil in 2021. Throughout the user test, a single smartphone (iPhone SE 2020) was used, with an iOS operating system, always connected to the same wi-fi network. The activities were carried out in an environment with a white background, and free of distractions, to maintain the isonomy of the user test.

The equipment used for data collection (eye-tracker) was the Senso Motoric Instruments (SMI) Eye Tracking Glasses 2W model, which is accompanied by the BeGaze software, capable of storing and analyzing the collected data. This tool captures eye movements through a pair of glasses that have a set of cameras, capable of mapping the gaze of its users. The infrared from the glasses creates reflections on the surface of the participant's eye, while the information system captures the data to provide the location of the reflection in the pupil, in addition to generating analyses on areas of greater interest according to the movement of the eyes during the execution of the user test (NIELSEN; PERNICE, 2010).

It is noteworthy that before starting data collection, a pre-test was carried out to better understand the tool that would be used and to evaluate the progress and duration of each of the activities that would be proposed. In this crucial stage, some details of the user test were corrected, such as the reduction in the number of activities, the need to use a single Smartphone for all participants, and the need to number and name each video that was generated by the software.

Participants were invited to contribute to the research through the presence of the moderator researcher and an assistant professor with experience in the use of the equipment. Data collection was as follows:

(1) each participant was positioned in a chair facing the smartphone, at an appropriate distance to perform the activities, in which the shopping application that would be used in the simulation was open;

(2) the eye-tracker glasses were positioned on the participant's head, so that it was comfortable, and then the equipment was calibrated so that the gaze capture was as dependable as possible;

(3) soon after, the moderator explained how the user test would work, emphasizing that the steps would be recorded, both in video and audio, and that each participant would perform three activities, which are summarized in Chart 3;

Chart 3 – Tasks performed by the Brazilian participants

| |
|---|
| Starting from the "Me" tab, go to the "Home" tab, find the search bar, and click on it, then search for a Computer. |
| Starting from the "Me" tab, access the "Official" tab, search for the "Food and drinks" category, and click on it, then in the "Official Stores" section access the first store. |
| Starting from the "Me" tab, go to the "Home" tab, find the "Exclusive coupon for first purchase" section and click on the "Free Shipping Coupon", then access the "Conditions" only for the "Free Shipping" coupon. |

Source: Prepared by the authors.

(4) When the participant considers that he or she has performed a task, he or she is instructed to say aloud that he or she has finished so that a video can be computed, and he or she can move on. Until all activities are performed, one at a time, and the user test is over.

Data collection with all nine participants lasted about five hours. The realization of pre-tests proved to be essential so that it was not necessary to repeat any of the activities and to run as realistically as possible.

### 3.3.2 Spanish Eye-Tracking User Testing

The second user test was carried out in Spain, in June 2024, and consisted of a simulation of the use of a shopping website, instead of an app. Anwyl-Irvine et al. (2019) highlight that the use of user testing with websites can support studies focused on both mobile and non-mobile platforms. It was developed at the “Social, Behavioural, and Cognition – SBC” laboratory of the Universitat Oberta de Catalunya (UOC) in Barcelona. As in the Brazilian user test, the Spanish users of online shopping apps are considered the population. The participants group was composed of nine volunteers who participated, who were also equally divided into three groups according to age groups (18 to 30 years old; 31 to 49 years old; and 50 years old or older).

The object of research was Amazon’s shopping website, which, according to the Statista Search Department (2023), is, to date, the most used e-commerce by Spaniards, covering 81% of the population. During the user test, a single computer (Lenovo) with Windows Vista operating system was used, always connected to the same internet via Wi-Fi; and the activities were carried out in a room free of distractions, to preserve the isonomy of the user test.

The eye-tracker used for data collection was the Tobii T60 model, accompanied by the Tobii Studio software, which stores and analyzes the collected data. Like the equipment used in the Brazilian user test, this one is also capable of capturing eye movements through cameras. However, unlike the equipment used in Brazil, this is not a tool in the shape of glasses; It is a machine that looks like a conventional computer, which, in the context of online shopping environments, is restricted only to simulating the use of websites.

As in the Brazilian user test, in this one a pre-test was also carried out before starting data collection, to better understand the tool that would be used and to evaluate the progress and duration of each of the activities that would be proposed. At this stage, some details of the user test were corrected, such as the addition of activities to better evaluate the participants, and the need to use a mouse and keyboard attached to an eye-tracking machine.

The participants were invited to contribute to the research through the presence of the moderator researcher, who had undergone previous training on how to use the eye tracker. Data collection was as follows:

(1) each participant was positioned in a chair facing the Tobii T60 eye-tracker, there was an adequate distance to perform the activities;
(2) Next, the equipment was calibrated, with the participant looking directly at the screen;

(3) soon after, the moderator explained how each guest would participate in the research, emphasizing that the steps would be recorded, both in video and audio, and that each participant would perform five activities during the user test, which are summarized in Chart 4;

Chart 4 – Tasks performed by the Spanish participants

| |
|---|
| Browse to the location where you can log in and access it. Then, look for the field where you enter your mobile number or email and click on it. |
| Go to the website's search bar. |
| Go to the "Best Sellers" section of the website. Then, choose one of the products offered and click on it. |
| Go to the website's shopping cart. |
| Go to the "Flash Deals" section. Then, go to the "Furniture" category, choose one of the furniture offered, and click on it. |

Source: Prepared by the authors.

(4) The eye-tracker already had the activities pre-programmed so that the participant could perform them autonomously, without pauses, following the commands that appeared on the screen. The user test ended when the participant performed the five tasks.

Data collection with all nine participants lasted around five and a half hours. Here the performance of pre-tests also proved to be equally essential for the user test to take place properly and without the need for repetitions.

In the subsequent section, the results are displayed and discussed concerning the research question of the study. Which revolves around research on the presence of online flow conditions in online shopping environments, considering web design elements, and flow constraints, previously highlighted.

## 3.4 SCALE CONSTRUCTION AND DATA COLLECTION

In this section, the process of scale definition for the study will be outlined, focusing on several critical stages essential for ensuring the validity and reliability of the measurement instrument. These stages include Items Justification, Expert Analysis, Pre-Test Evaluation, description of Population and Sample characteristics, and the methodology employed for Data Collection. Each stage plays a crucial role in the development and refinement of the measurement scale, aiming to capture accurate and meaningful data relevant to the research objectives. By systematically addressing these components, the study aims to establish a robust foundation for assessing the phenomena under investigation with clarity and precision.

### 3.4.1 Items Justification

As previously noted, Novak, Hoffman, and Yung (2000), pioneers in the study of online flow, categorize this flow state into three primary areas: conditions, characteristics, and consequences. This thesis focuses exclusively on flow conditions – elements in web design that stimulate the emergence of flow in users. The proposed scale for measuring Online Flow Conditions builds upon three critical components of flow: Demand-Skill Balance, Clear Goals, and Immediate Feedback incorporating insights from existing literature and expert opinions.

Flow is not easily self-diagnosed but rather defined by the equilibrium between user demands and abilities. Rittie (2001) suggests that fostering flow in virtual environments aligns actions with user capabilities, facilitating decision-making during navigation. King (2003) emphasizes the importance of users feeling at ease upon entering these environments. Introducing onboarding screens can enhance this experience by providing a clear, step-by-step introduction to key commands (HIGGINS, 2020). Thus, the following item was defined:

Item 1) The application has onboarding, an initial step-by-step that indicates how its use works.

Regarding app developers, Guo and Poole (2009) emphasize the significance of understanding user skill levels and goals to design interfaces that enhance the flow experience. An adaptable interface that allows users to adjust the complexity of the virtual environment according to their skills supports the continuity of online flow. Additionally, minimizing the initial registration process aligns with the idea of creating a user-friendly environment that promotes seamless navigation and encourages flow (King, 2003). The approach of allowing users to explore first and register later is pivotal in this context. Based on that, these items were defined:

Item 2) The application features an adjustable interface, enabling users to tailor it to their preferences.

Item 3) The application allows its use without the need for user registration.

To enhance the online streaming experience, motivating and inspiring users as they navigate the virtual environment is essential. Mahnke, Benlian, and Hess (2015) emphasize the critical role of accurate recommender systems, particularly for users driven by curiosity. Recommendations, whether general or personalized based on user history, significantly enhance user experience and perceived value (AYADA; HAMMAD, 2023). Kim (2010) underscores the impact of prominently displaying product recommendations on the virtual environment's homepage. Moreover, Ayada and Hammad (2023) highlight the importance of providing users with customization options for receiving product information, such as through messages or emails. Therefore, the following items were defined:

Item 4) Users can access a list that compiles the history of viewed items within the application.

Item 5) The application shows personalized product recommendations based on the user's history.

Item 6) Users can select their preferred method for receiving notifications from the application, such as via email, SMS, WhatsApp, or other available channels.

Flow-enabled online environments require clear and easy-to-perform activities, so the speed of task completion is intricately linked to the presence of flow in the environment (KING, 2003; NOVAK, HOFFMAN AND YUNG, 2000; FINNERAN AND ZHANG, 2005). Oliveira et al. (2023) highlight that the shorter the active time in the system, the greater the users' perception that there is a balance between demands and their skills. And Ayada and Hammad (2023) corroborate this, by highlighting that slow loading times frustrate users while browsing. Thus, the absence of a loading icon while browsing, due to the speed of the information system, can be an indication of a virtual environment enabled to flow. Therefore:

Item 7) The application does not display a loading icon during use.

Flexible, adjustable, and easy-to-use interfaces direct the user to the flow state. In such a way, information control mechanisms, such as zooming, have been used to help the user find, access, and focus on the intended information (PUNCHOOJIT AND HONGWARITTORRN, 2017). Hamza and Salivia (2015) point out that since childhood there is already an ease in handling Zoom on the screen of mobile devices. Then, the item is defined:

Item 8) It is possible to use zoom to enlarge product images.

The way you view the content displayed in a virtual environment influences the online flow. Thus, according to Mahnke, Benlian, and Hess (2015), providing infinitely scrolling pages when displaying product lists helps this feeling. To develop an application with an appropriate interface design that attracts new users, Jiang et al. (2019) concluded, through nonlinear modeling, that the horizontal layout criterion must be met. In addition, the authors state that it is essential to make the design suitable for the width of the mobile device, which can be done by limiting vertical scrolling. Findings corroborated by Chan et al. (2002), which help in the definition of the items:

Item 9) The application has a horizontal layout and vertical scrolling.

Item 10) The application has infinitely scrollable product sections.

Adaptive virtual interfaces, which allow the user to adjust, strengthen the sense of online flow (KING, 2003). In addition, Ayada and Hammad (2023) highlight that personalization is an important aspect of app design, as it is capable of improving the relationship with users from different cultures and situations. The authors also concluded, through Survey, that customization is also an important factor. Able to improve the online experience and increase the perception of the app's value to the user. Based on this:

Item 11) The application enables modification of the user's profile, including inserting a personal photo, creating a username, and other related features.

Item 12) The application facilitates customization of its components within the account settings section.

In online environments, the sensation of flow arises from a balance between what is demanded virtually, and the skills possessed by the user (TRIBERTI; DI NATALE; GAGGIOLI, 2021). These skills can be based on previous knowledge. Currently, the two major operating systems in place for mobile devices are iOS and Android (KOCH; KERSCHBAUM, 2014). In the context of acquired skills, the routine use of one, rather than the other, familiarizes the user with the most commonly used system to the detriment of the other. Thus, the existence of an application with versions for both operating systems is essential so that users can flow, regardless of the system used. Therefore:

Item 13) There is a version of the application for iOS and Android.

Users have specific goals when browsing the internet (MAHNKE; BENLIAN and HESS, 2015). And the virtual environment in which they are inserted must provide tools that help them achieve these goals in the best way. Choi, Kirshner, and Wu (2016) point out that to alleviate the cognitive burden associated with choice overload, apps encourage users to save to a list of products they may be interested in buying. At any time, the user can review the list of saved items to make actual purchases. In this context, the possibility of saving products in a list helps the user in their journey and corroborates the generation of flow. Supriadi (2019) points out that a history list is essential for users to keep track of their records in the application used. That said:

Item 14) The application allows users to save items to user-created lists.

Item 15) The application enables users to add items to a favorites list.

Item 16) The application allows users to share lists of items with other users.

An online environment enabled for online flow is reflected through simple and clear navigation (KING, 2003). For this, there are web design elements that support this perception, such as signposts, words that point out the meaning of longer entries and act as a visual index that helps the user access important sessions more quickly (KIM; HUA; NAFISEH, 2011); breadcrumbs trails, a type of schema that reveals the user's location in a virtual environment to create a hierarchical structure (SHITKOVA et al., 2015); "you are here" landmarks (KING, 2003); and even a navigation bar, which help users find the best path so that they can easily form a mental model of the virtual environment they are navigating. Therefore, the following items were defined:

Item 17) The application has breadcrumb trails that allow the user to identify which part of the application they are in, referring to a general structure.

Item 18) The application provides a section where users can answer questions, such as a FAQ (Frequently Asked Questions) section or direct contact with customer service.

Item 19) The application features “you are here” style location landmarks, symbols, or highlights so that the user can easily form a mental model of the app.

Item 20) The application has a navigation bar to access its main information, for example, user profile, main page, etc.

To improve consumers’ online flow experiences, Mahnke, Benlian, and Hess (2015) point out that it is sensible to pursue a strategy based on increasing the user’s ability to maintain their motivation towards a pre-existing goal. To maintain motivation, continuous progress toward the existing goal must be ensured, according to the authors, through the provision of information in quantity and quality, as well as search engine optimization and filtering. According to this, the following items were defined:

Item 21) The application enables users to apply various filters to the lists of available products.

The quality of the information available in the virtual environment is crucial to support the potential reception of relevant information. Mahnke, Benlian, and Hess (2015) point out that quantity information is also needed to increase absorption and exploratory behavior. In terms of shopping apps, the amount of information is provided through a variety of product ranges and advertisements. Therefore:

Item 22) The application’s homepage features announcements, such as special birthday promotions, Black Friday promotions, etc.

Item 23) After completing an action, the application displays motivational information, such as promotions and most purchased products, to encourage continued user engagement.

The expeditiousness of navigation holds paramount significance in the realm of online streaming, as emphasized by works such as that of King (2003) and more recent contributions by Oliveira et al. (2023). In this context, the provision of buttons facilitating rapid access to specific content assumes pivotal importance. Notably, Kim (2010) accentuates the ubiquity of shopping cart buttons within online retail environments, emphasizing their imperative accessibility. Furthermore, scholars such as Saffer (2014) and Neil (2014) advocate for the inclusion of features enabling users to ascertain the number of products housed within the shopping cart without necessitating direct entry into it. Therefore, the following items were defined:

Item 24) The application includes a visible button for accessing the shopping cart.

Item 25) Users can obtain the number of items in the shopping cart without needing to access it directly.

Mahnke, Benlian, and Hess (2015) highlight regardless of the type of user (those directed to specific goals or those who are just navigating) in both there is evidence of flow. Moreover, the two profiles are closely related since it is possible to alternate existing goals and formulate new ones. These are directly linked to important information that the virtual environment must offer, such as product characteristics, order fulfillment, payment, and delivery methods (WU; CHIU and CHEN, 2020). Ayada and Hammad (2023) highlight that delivery updates, interactions with customer service, and receiving notifications during the shopping process enhance users' experience while achieving their goals. Kim (2010) also points out that the home page of the virtual environment should have an easily accessible section that indicates the status of the products purchased. Therefore, the following items were created:

Item 26) Users have the option to qualify for free shipping within the application.

Item 27) The application has a customer support section.

Item 28) The application allows more than two payment methods.

An online environment designed for flow has consistency. Which means providing an organization of information consistently over time. This enhances the expertise effects that lead to better orientation in a virtual environment. Mahnke, Benlian, and Hess (2015) state that consistency implies both a consistent layout (arrangement of elements) and a consistent design (colors, fonts, etc.). Therefore:

Item 29) The application uses the same font, color scheme, and layout on different pages.

Online browsing should be fluid, fast, and uninterrupted (KING, 2003). Therefore, the online environment must respond correctly and quickly to incoming commands. Supriadi (2019) points out that responsive applications can provide a button to go back to the previous page or state, enabling flow-oriented navigation. According to this, the following item was defined:

Item 30) The application provides a visible button to go back to the previous page or state.

Feedback is a crucial two-way information flow process in online environments, as highlighted by Barta, Flavián, and Gurrea (2022). It involves the exchange of information between the virtual environment and its users. When users perceive control over their interactions with technology, the feedback process is enhanced, fostering perceptions of competence, self-confidence, and increased self-determination. This contributes to the generation of the flow state, which is pivotal for the success of virtual environments.

Effective feedback, as emphasized by Ayada and Hammad (2023), significantly enhances user experience. King (2003) further illustrates that in flow-enabled environments, feedback can take various forms: navigation widgets, performance variables (such as server load or download progress bars), floating or visited links, and confirmation messages. So, based on that the items were defined as:

Item 31) The application provides feedback in the form of navigation widgets, such as menus, toolbars, etc.

Item 32) The application provides feedback in the form of performance variables, such as loading server icons, memory usage amount, etc.

Item 33) The application has a progress bar, for example, in the cases of displaying the loading of a page, how much is left to get free shipping, etc.

Item 34) The application presents a confirmation message before executing specific actions, such as deleting items from the favorites list or shopping cart.

Item 35) The application has visual differentiation between visited and unvisited sessions.

Users should receive clear and concise information when interacting with an interface, through immediate feedback. Ayada and Hammada (2023) state that this may involve providing visual information and auditory cues, such as messages, symbols, or sounds, to indicate when an action has been completed or when an error has occurred. The presence of feedback can help users understand how to use the interface, as well as assist in the creation of a flow-enabled virtual environment (KING, 2003). That said, the items were defined as:

Item 36) The application delivers feedback audibly.

Item 37) The application provides feedback in the form of symbols that express information, such as emojis, numbers, etc.

Item 38) The application provides feedback in the form of a banner notification.

Feedback transmission is part of a flow-enabled environment (FLAVIÁN and GURREA, 2022), and micro-interactions are the simplest, briefest, and easiest ways to convey feedback to users (SAFFER, 2014). Through subtle hints, this type of interaction can help users improve their experience. Among the examples of micro-interactions in online shopping environments, Saffer (2014) and Neil (2014) point out the ability to highlight incorrect entries in the information system, the renaming of buttons (e.g., "add to cart" and "add another to cart" if this product already exists; or "buy now" and "buy another", if you already have a purchase history). That said, the following items were defined:

Item 39) When entering incorrect input data, the application notifies the user of the error.

Item 40) Some buttons react to user actions by changing nomenclature, such as "add to cart" and "add another to cart", "buy now" and "buy another now", etc.

Mahnke, Benlian, and Hess (2015) assert that online shopping environments must maintain consistency, which involves consistently providing organized information over time. A consistent layout includes various elements such as a search bar, navigation bar, and widgets, among others. In this context, the options of product search stand out as an essential component for any application aiming to provide a positive, efficient, and personalized user experience. It facilitates navigation, provides feedback, and generates valuable insights into user behavior. Accordingly, the following items were developed:

Item 41) The application offers more than one method for product search, such as a search bar, sorting and filtering options, and voice commands, among others.

Item 42) The application has an automatic entry feature in the search bar, that is, as something is typed, possible product options are already provided.

Furthermore, different forms of login can demonstrate that the application prioritizes agility and speed, thereby facilitating the online flow (Kim, 2010). This approach can also serve as a means of providing feedback to the user by offering an alternative way to access the platform. Consequently, the item was defined as follows:

Item 43) The application provides alternative methods of logging in, such as authentication via social media accounts like Facebook, Google, and others.

When users log into a virtual shopping environment, a practical method for providing feedback about the available products is through generic product recommendations (Ayada and Hammad, 2023). This approach enables users to begin browsing and enter a state of online flow. Consequently, an item was defined as follows:

Item 44) The application offers generic product recommendations on the homepage.

According to Mahnke, Benlian, and Hess (2015), user-generated content fosters social proximity, which is crucial for satisfying individual interests and supporting the state of online flow. The authors also highlight that the affective quality of feedback information contributes to the flow and can be enhanced by offering media richness, such as providing images, audio, or videos of the products. They note that prior evidence demonstrates the impact of media richness on emotional responses and flow in online commerce contexts. Therefore, the following items were created:

Item 45) The application allows users to read and submit product reviews.

Item 46) The application allows users to submit images, audio, or videos of the products in reviews.

Another effective micro-interaction, highlighted by Saffer (2014), involves sending personalized messages to the user (e.g., welcome, farewell, inquiries) as a means to strengthen the relationship between the shopping app and the user. Consequently, the following item was defined:

Item 47) The application sends messages of a personal nature to the user, such as welcome, good morning, good evening, etc.

Therefore, these 47 items were defined to compose the Online Flow Conditions measurement scale.

### 3.4.2 Expert Analysis

Through a bibliographic review, a total of 47 items were identified for inclusion in the scale. These items were subjected to evaluation by a panel of five experts using a structured assessment form. The expert panel comprised one PhD specializing in Business Administration from a public university in Brazil, three experts in Web Design, all from various public universities in Brazil, and one expert in Computer Science from a public/private university in Spain. This diverse panel was assembled to rigorously assess the content validity of the scale, a crucial aspect emphasized in scale development by DeVellis (2017).

Following DeVellis's guidelines, each item underwent evaluation by experts based on two primary criteria: (1) Clarity, which assessed the clarity of item wording and the absence of ambiguity, and (2) Relevance, which gauged how well each item aligned with the intended latent trait. Experts rated each item on a scale from 1 (low) to 3 (high). Furthermore, the evaluation form provided experts with the opportunity to offer recommendations and suggestions for refining the instrument. The item evaluation form and references for each item are detailed in Appendix C.

Overall, the items received positive evaluations, with all items achieving moderate to high relevance based on median scores. Boone and Boone (2012) emphasize that the median is a more appropriate measure of central tendency for Likert scale data because it respects the ordinal nature of these scales and avoids the assumption of equal intervals between points, which can lead to inaccurate interpretations. The mean can be misleading, especially with non-normally distributed data, as it assumes equal spacing between response options, which is incorrect for ordinal data.

Experts provided constructive feedback on item wording, the inclusion of examples for clarity enhancement, and suggestions for overall improvement. Among the various strategies observed, certain items have been modified through the inclusion of illustrative examples to enhance comprehension, represented by items 31 – The application provides feedback in the form of navigation widgets, such as menus, toolbars, etc., and 43 – The application provides alternative methods of logging in, such as authentication via social media accounts like Facebook, Google, and others. Additionally, adjustments such as word substitutions or insertions have been implemented to aid clarity, as demonstrated in item 4 – Users can access a list that compiles the history of viewed items within the application, in which it was inserted the word “viewed” to clarify the king of product. Furthermore, there are instances where clarification of less familiar terms within items has been undertaken, as verified by item 17 – The application has breadcrumb trails that allow the user to identify which part of the application they are in, referring to a general structure, in which it was explained the term “breadcrumb”.

A summary of the experts’ evaluations is provided in Appendix D. This section offers a concise overview of the assessments conducted by the experts, listing their outcomes on the subject matter under review.

### 3.4.3 Pre-Test

A pre-test evaluation was conducted to standardize item analysis. Ten randomly selected shopping apps were assessed using the scale. This process aimed to establish consistent evaluation criteria based on uniform subjective premises, irrespective of the app under analysis. The results indicated the necessity to adjust the wording of 5 items to ensure consistency across all assessments.

Adjustments were made in item 3 – The application allows its use without the need for user registration, the term “sign up” was replaced with “registration” for better understating, and item 10 – The application has infinitely scrollable product sections, the term “infinite scroll” was replaced by “infinitely scrollable product sections” for the same reason. Additionally, on item 6 – Users can select their preferred method for receiving notifications from the application, such as via email, SMS, WhatsApp, or other available channels, and item 40 –Some buttons react to user actions by changing nomenclature, such as “add to cart” and “add another to cart”, “buy now” and “buy another now”, etc., it was inserted more examples to clarify the context. Furthermore, in item 18 – The application provides a section where users can answer questions, such as a FAQ (Frequently Asked Questions) section or direct contact with customer service, the meaning of the acronym was given. No substantial changes were made, just small adjustments that were considered necessary to clarify the item.

### 3.4.4 Population and Sample

Regarding the measurement scale, its items were applied to a sample of 200 mobile shopping applications (100 from Brazil and 100 from Spain) available in the mobile app store. The population consisted of all shopping apps accessible in each country’s respective app store. Appendix B provides charts displaying the specific shopping apps included in the research sample.

According to Embretson and Reise (2013), one prerequisite for sample selection is heterogeneity in the latent trait of interest. Thus, shopping apps were chosen due to their diverse range of activities, aligning with the general objective of this thesis. The sample of shopping apps was selected conveniently using a non-probabilistic approach, based on apps listed in the Top Charts of the shopping category in the Brazilian and Spanish Apple stores.

There is no consensus on the optimal sample size for using IRT (DOWNING, 2003; WONGTADA; RICE, 2008). However, studies applying IRT in the technology field have utilized samples ranging from 208 to 361 elements (MITSUTAKE et al., 2020; TEZZA; BORNIA; ANDRADE, 2011; MOREIRA JR. et al., 2013; ACHTENHAGEN; WINTHER, 2014).

### 3.4.5 Data Collection

Following the development and pre-testing of the measurement scale, the next phase involves its application to a sample of shopping apps. The data collection approach aligns with methodologies previously employed by Al-Khalifa (2010), Stepchenkova et al. (2010), Tezza et al. (2011), Moreira Jr. et al. (2013), and Tezza et al. (2016), where the scale items serve to assess the information system rather than capturing user perceptions. Hence, this manual data collection adheres strictly to a checklist format. The scale comprises objective items designed to ascertain the presence or absence of specific characteristics associated with online flow conditions within each evaluated shopping app. These items feature dichotomous responses (0 or 1).

Data collection took place from February 2024 to April 2024. All 200 applications were accessed using the same mobile device (iPhone SE), and a consistent Wi-Fi connection was maintained throughout to ensure uniformity in data collection conditions. Data were recorded and compiled in a Microsoft Excel spreadsheet for subsequent conversion and analysis using data analysis software.

## 3.5 SCALE ASSESSMENT AND STATISTICAL METHODOLOGY

To achieve the third specific objective – analyze the scale dimensionality and evaluate the suitability of relevant item response theory models – the next step involves statistical analysis of the data collected using the measurement scale. This phase will include dimensionality analyses through confirmatory factor analysis to determine the number of dimensions and assess each item’s suitability within those dimensions. Additionally, the adequacy of various IRT models will be evaluated to identify the model that best fits the observed data patterns. The statistical software R will be utilized for both IRT modeling and factor analysis, facilitating a comprehensive examination of the scale’s underlying structure and model fit.

Briefly, the statistical analysis followed these three steps:

Chart 5 – Statistic methodology steps

| | |
|---|---|
| 1. | Development of an assessment framework to verify the predetermined number of dimensions in the scale through an assessment of item quality. |
| 2. | Validation and confirmation of the dimensions identified in the initial phase through statistical analysis and examination of the content relationship of each item within these dimensions. |
| 3. | Suitability check of the IRT models to the obtained data. |

Source: Elaborated by the author, 2024.

For statistical significance purposes, the analyses were performed considering a significance level of 5% (p-value < 0.05) in all statistical tests, and the evaluations of the IRT modeling. Below, each of these three stages is described in greater depth.

### 3.5.1 Assessment Framework

Authors such as Fleck and Bourdel (1998) and Podsakoff, Ahearne, and MacKenzie (1997) advocate for an empirical criterion in retaining dimensions with strong theoretical foundations. This criterion involves evaluating item quality based on factor loadings to identify and eliminate items that do not effectively measure constructs within the instrument, thus avoiding potential spurious dimensions.

Before commencing the empirical investigation of items and dimensions, it is important to note that although the items were derived from literature and validated by experts, they had not undergone quantitative evaluation for quality after data collection. Nevertheless, the scale was expected to consist of three dimensions.

To confirm the scale's dimensionality, factor analysis was conducted using the "psych" package in R software. This analysis determined the number of dimensions present and assessed item quality based on their factor loadings. Rotation methods such as varimax (for orthogonal factors) or oblique methods like oblimin (for correlated factors) were employed to enhance understanding of the relationships between dimensions and items.

Factor loadings in factor analysis serve as critical indicators of the strength of associations between items and latent factors in the dataset. While there is no universally agreed threshold, Hair et al. (2019) and Tabachnick and Fidell (2013) suggest that loadings below 0.3 indicate weak associations, possibly suggesting poor item alignment with factors or susceptibility to cross-loading. Loadings between 0.3 and 0.5 suggest moderate associations, indicating a reasonable yet improvable relationship. Loadings above 0.5 indicate strong associations, where items significantly contribute to explaining variance in the construct. These interpretations are crucial for assessing the reliability and validity of measurement instruments and guiding refinements to factor structures in research studies.

Additionally, communalities above 0.2 are essential in factor analysis to ensure each item adequately contributes to explaining identified factors. Low communalities suggest inadequate representation of items by the included factors, potentially compromising result validity by insufficiently capturing construct variance (HAIR et al., 2019; TABACHNICK; FIDELL, 2013). Therefore, items with factor loadings below 0.3 or communalities below 0.2 are considered of low quality and were excluded from the analysis.

After eliminating low-quality items, the remaining dimensions are re-evaluated. The potential correlation between dimensions is also examined to determine the appropriate rotation type for distributing item loadings across dimensions. Based on this assessment, each item’s loadings across dimensions are reviewed to determine its placement within the dimension where it exhibits the highest loading.

### 3.5.2 Validation and Confirmation

Following the analyses initiated in the preceding section, the next step involves concurrently revisiting the initial procedures for item evaluation and checking the number of dimensions within the item set. This step aims to validate and confirm these dimensions using R software.

Validation refers to demonstrating that a measurement scale effectively measures its intended construct. This includes assessing content validity (whether scale items represent the theoretical construct), construct validity (whether the scale measures the intended construct), and other forms of validity like convergent and discriminant validity (DEVELLIS, 2017). The process is continuous and multifaceted, involving statistical analyses, the already done expert reviews, and comparisons with established measures.

Confirmation, on the other hand, entails verifying the replicability and consistency of previous analysis results (DEVELLIS, 2017). This typically involves additional analyses to confirm factor structure, internal consistency, and other psychometric properties of the scale, ensuring reliability for subsequent studies.

The initial step involves applying Parallel Analysis (CHERNYSHENKO et al., 2001) and Kaiser's criterion (KAISER, 1958) to evaluate the scree plot of eigenvalues. In Parallel Analysis, a comparison is made between a simulated data set and the actual data set, with eigenvalues exceeding those of the simulated set being considered for extraction. Kaiser's criterion, on the other hand, identifies eigenvalues greater than 1 as significant for factor extraction. Additionally, Confirmatory Factor Analysis (CFA) using fit indices from the "lavaan" and "semPlot" packages in R software was conducted to evaluate the scale's internal reliability.

To confirm internal consistency, Cronbach's Alpha (CRONBACH, 1951) was used, supplemented by McDonald's Omega (MCDONALD, 1986) to address the limitations of Cronbach's Alpha (SIJTSMA, 2009; BRUNNER; NAGY; WILHELM, 2012). Both indices targeted values above 0.7, aligning with guidelines that values over 0.6 are considered acceptable (NUNNALL; BERNSTEIN, 1994).

Convergent validity will be assessed through factor loadings, composite reliability, and Average Variance Extracted (AVE) for each item. Factor loadings indicate item-concept relationships, with higher values indicating stronger convergent validity as seen in the previous 3.5.1 section. Composite reliability evaluates internal consistency, aiming for a threshold of 0.7, while AVE measures variance captured relative to measurement error, with values above 0.5 indicating satisfactory convergent validity (HAIR et al., 2019). Fornell and Larcker (1981) emphasize that when the AVE is less than 0.5, but the composite reliability exceeds 0.6, convergent validity remains affirmed. This is because composite reliability provides a broader assessment of the internal consistency of the construct, considering the overall reliability of its indicators.

Item-total correlation, typically acceptable above 0.3 (CLARK; WATSON, 1995), also contributes to convergent validity, measuring each item's contribution to the total score reliability. Field (2018) acknowledges that while 0.3 is a strong indicator of item reliability, in practical applications, values around 0.2 can still be considered adequate, providing that the overall reliability and validity of the scale are acceptable.

Discriminant validity will be confirmed through AVE, Maximum Shared Variance (MSV), and Average Shared Variance (ASV) values, ensuring each construct's distinctness from others. When the square root of AVE for each construct exceeds the correlations with every other construct in the model, it will confirm discriminant validity (HAIR et al., 2019).

Upon completing these analyses, a refined item set will be identified before progressing to explore the IRT modeling.

### 3.5.3 IRT Models Suitability

The final stage of the statistical methodology for scale evaluation involves assessing the appropriateness of existing IRT models with the resulting items. Initially, the feasibility of treating the three dimensions in the scale separately was tested by applying a unidimensional two-parameter IRT model to each set of items. This approach aims to examine the independent behavior of each dimension and explore the potential for autonomous treatment.

Subsequently, the adequacy of the multidimensional two-parameter IRT model was evaluated, considering the item set resulting from the factor analysis and the presence of three dimensions. This model is available in the R software through the "mirt" package. Additionally, a validation of the MIRT model results was conducted using a Confirmatory tool. This technique examines the covariances among the three dimensions present in the scale and emphasizes the most significant discrimination parameters in each dimension, aiming to confirm the outcomes derived from the MIRT model.

In the evaluation process, various models are compared using the Sample Adjusted Bayesian Information Criterion (SABIC) (BOZDOGAN, 1987) or the sample-adjusted version, along with the Akaike Information Criterion (AIC)

(AKAIKE, 1973). Lower values of these criteria indicate a superior fit of the model to the data, signifying a greater level of modeled information.

Furthermore, an analysis of variance (ANOVA) test, recommended and implemented in the R software by Chalmers (2012), was employed. Chalmers (2012) suggests that by conducting a generic ANOVA, specifically via the chi-square ($\chi^2$) test based on likelihood comparisons between models, the optimal number of dimensions for the best data fit can be determined. This test compares models under the null hypothesis that there is no disparity in information between the evaluated model and the model with one fewer dimension, indicating that the addition of another dimension does not notably enhance the model’s informational capacity. Conversely, the alternative hypothesis proposes a substantial increase in information for the model with additional dimensions. Four models were scrutinized: one assuming a single dimension, another assuming two dimensions, a third assuming three dimensions, and a fourth assuming four dimensions.

The statistic $M_2$, as discussed by Joe and Maydeu-Olivares (2010), resembles the chi-square statistic and is widely utilized in model fit assessment in IRT. A lower value of $M_2$ signifies a superior model fit compared to others. Additionally, a comparison of Root-Mean-Square Error of Approximation (RMSEA) values aids in validating the model selection, with values below 0.05 regarded as optimal.

An assessment of the adequacy of the Bifactor model, also available in the R software through the “mirt” package, was conducted. A comparison of model fit indices (AIC, SABIC, $M_2$, chi-square, and RMSEA) between the Bifactor model and the MIRT was performed to select the most appropriate model.

Theoretical verification is also undertaken, considering the concepts of online flow conditions embedded in each item and the estimation of parameters in each model.

## 3.6 CHAPTER SUMMARY

The chapter dedicated to epistemological and methodological procedures intricately delved into a series of interconnected subsections, each playing a pivotal role in establishing a solid research framework. Commencing with an exploration of epistemological assumptions and study paradigms (Section 3.1), the chapter laid the groundwork for subsequent methodological procedures. As for the research characterization, it predominantly adopted a quantitative approach, complemented by a qualitative exploratory foundation aimed at comprehending the field of online flow conditions in shopping applications and serving as a basis for item development.

The chapter proceeded to elucidate constitutive definitions and the operationalization of research steps (Section 3.2), offering valuable insights into the utilized methodology. The proposed study methodology involved a structured sequence: (1) conducting a literature search in databases; (2) executing two eye-tracking user tests to acquire data facilitating the creation of a measurement scale; (3) defining scale items concerning prior literature and outcomes of the eye-tracking user tests; (4) presenting the scale to five experts for feedback and subsequently refining it based on their inputs; (5) conducting a preliminary test; (6) administering the scale to a sample of 200 shopping app users through manual data collection; and finally, (7) evaluating the scale's statistical properties with a focus on assessing its suitability using Item Response Theory (IRT).

The discourse then proceeds to the eye-tracking user tests conducted in Brazil and Spain (Section 3.3), emphasizing the cultural variations and insights extracted from these studies. Furthermore, the chapter addressed the development of a scale (Section 3.4) to measure Online Flow Conditions, through three dimensions (Demand-Skill Balance; Clear Goals, and Immediate Feedback), highlighting the process of defining the 47 constitutive items within the scale. Lastly, the chapter delved into the statistical methodology used to evaluate the scale (Section 3.5), underscoring the rigorous analyses conducted to ensure the reliability and validity of the measured variables, and the following application of the IRT modeling. This chapter synthesized epistemological foundations, methodological frameworks, and empirical user tests, laying a solid groundwork for the subsequent research events.

# 4 RESULTS AND RESEARCH CONTRIBUTIONS

This chapter presents the results of the eye-tracking user testing presented in section 3.3 and also presents the scale assessment described in section 3.5, along with a discussion of these results using statistical theory and the concepts of online flow as frameworks for analysis.

## 4.1 EYE-TRACKING USER TESTING RESULTS AND DISCUSSION

Following the previous concepts exposed, this section presents the results of the two conducted user tests, accompanied by a corresponding discussion of these findings.

The investigation into web design elements relevant to online flow conditions continues to be a relatively underexplored area in research. As a result, a comprehensive literature search was conducted, yielding a compilation of elements (see Table 1) intended to facilitate the assessment of online flow conditions through eye-tracking methodologies. These elements serve as foundational criteria for examining how specific features of web design contribute to the facilitation or hindrance of flow experiences in online environments. This approach not only aims to enhance understanding of user interaction dynamics but also to provide empirical insights into optimizing web interfaces to foster optimal user experiences conducive to flow.

Table 1 – Evaluated elements of online flow conditions

| **Online flow conditions components** | **Evaluated elements** | **References** |
|---|---|---|
| Demand-Skill Balance | Speed of task completion | King (2003); Novak, Hoffman and Yung (2000); Finneran and Zhang (2005); Oliveira et al. (2023) |
| | Amount of time spent on each task | Oliveira et al. (2023) |
| | Quantity of gaze fixation points | Baharum et al. (2019) |
| | Fulfillment of the path of the gaze | |
| | Size of gaze fixings | Baharum et al. (2019); Boardman and McCormick (2022) |
| Clear Goals | Labels | Punchoojit and Hongwarittorrn (2017) |
| | Titles | |

| | | |
|---|---|---|
| | Icons | |
| | Logo | King (2003); Rettie (2001); Guo and Poole (2009) |
| | Information scent | King (2003); Blackmon (2012) |
| | Breadcrumb navigation | King (2003); Shitkova et al. (2015) |
| | Signposts | King (2003); Mothers, Van Geel and Cozijn (2006) |
| | Navigation bar | King (2003) |
| | Search bar | Mahnke, Benlian and Hess (2015) |
| Immediate Feedback | Server Load Icon | King (2003); Guo and Poole (2009); Barta, Flavián and Gurrea (2022); Saffer (2014) |
| | Sensory feedback in the form of visual changes, buttons, icons, navigation widgets, messages, and warnings. | |
| | Progress bar | King (2003); Saffer (2014) |
| | Micro-interactions | |

Source: Prepared by the author, 2024.

The elements evaluated in Table 1 assisted in validating some of the previously presented items (section 3.4.1). Participant gaze during task completion confirmed the relevance of certain characteristics of virtual shopping environments. Among the items that were refined based on the user test results are: 7 – The application does not display a loading icon during use; 17 – The application has breadcrumb trails that allow the user to identify which part of the application they are in, referring to a general structure; 20 – The application has a navigation bar to access its main information, for example, user profile, main page, etc.; 24 – The application includes a visible button for accessing the shopping cart; 25 – Users can obtain the number of items in the shopping cart without needing to access it directly; 32 – The application provides feedback in the form of performance variables, such as loading server icons, memory usage amount, etc.; 37 – The application provides feedback in the form of symbols that express information, such as emojis, numbers, etc.; 38 – The application provides feedback in the form of a banner notification; and 41 – The application offers more than one method for product search, such as a search bar, sorting and filtering options, and voice commands, among others.

Regarding the user tests, the presence of a Demand-Skill Balance was assessed primarily through the speed of task accomplishment, and through the amount of gaze fixations required to achieve web design elements correlated with online flow conditions. Thus, the age group of users who obtain lower quantities is better adapted to the online shopping environment.

Regarding the presence of clear objectives and immediate feedback in the online environments studied, the evaluation will be made around the ease of finding the respective elements, perceived through the size and quantity of the gaze fixations. Thus, the generation of minor fixations by a user is an indication of a virtual environment that is better designed for him.

King (2003), throughout his research, emphasizes that enabling flow in online environments is not an easy task for information systems developers. However, Barta, Flavián, and Gurrea (2022) state that the presence of flow in these environments is capable of increasing the perception of virtual aesthetics by users and awakening the feeling of satisfaction. So, developing flow-enabled virtual environments is about ensuring that the user experience will be enhanced.

Kristianto (2017) further points out that the streaming experience has a direct and positive effect on the attitude towards the use of social media systems by younger users. In the area of information systems, Morris and Venkatesh (2000) believe that age is a truly relevant and present factor in the context of technology adoption. In addition, the authors highlight the age group of 18 to 33 years (considered young) as the real users of technology associated with the computational area. In this context, it has been assumed that online shopping environments are designed to awaken the state of online flow more easily in younger users.

Barthelmäs and Keller (2021) point out that when the user’s objectives are clear, their skills are up to what is requested of us by the information system during its use, and the feedback received is unequivocal, users engage in the proposed task, whatever it may be. Thus, the highlighted elements served as the basis of the investigation during the user tests conducted, and their relevance was evaluated through the eyes of the participants. It is worth noting that, if the online environment has the three conditioning elements reported, the components of the other two areas of the online flow will be more easily reached by users. Thus, an adequate screening of these elements is essential for the development of online environments.

Flow-enabled online environments require clear and easy-to-perform activities, so the speed of task completion is linked to the presence of flow in the environment (King, 2003; Novak, Hoffman & Yung, 2000; Finneran and Zhang, 2005). Recently, Oliveira et al. (2023) obtained significant results in an user test focused on the use of an information system created by them, which indicates that there is an inverse relationship between the amount of time of users' actions and the conditions of the flow experience. The authors point out that (1) the shorter the active time in the system, the greater the users' perception that there is: a presence of clear objectives and a balance between the demands and their skills; and (2) the shorter the average response time after positive feedback, the greater the perception of immediate feedback.

Thus, during each user test, the time of performance of the proposed activities was computed, to perceive which age group of the participants was more apt to flow through the online environments evaluated.

In Brazil, the time of each of the three activities was computed. In the user test carried out in Spain, only the total time of the five activities of each participant was computed, due to the characteristics of the eye-tracker used (see Table 2).

Table 2 – Time of user tests in Brazil and Spain (in minutes and seconds)

| | | **Brazil** | | | | **Spain** |
|---|---|---|---|---|---|---|
| | | **Task 1** | **Task 2** | **Task 3** | **Total** | **Total** |
| **Group 1** | Participant 1 | 24' | 21' | 32' | 1'17' | 1'38' |
| | Participant 2 | 22' | 19' | 20' | 1'01' | 1'25' |
| | Participant 3 | 19' | 19' | 23' | 1'01' | 1'28' |
| | Mean (SD) | 22' (2,5) | 20' (1.2) | 25' (6.2) | 1'06' (9,2) | 1'30' (6,8) |
| **Group 2** | Participant 4 | 20' | 32' | 21' | 1'13' | 1'69' |
| | Participant 5 | 34' | 49' | 27' | 1'50' | 1'79' |
| | Participant 6 | 33' | 31' | 25' | 1'29' | 1'42' |
| | Mean (SD) | 29' (7.8) | 37' (10.1) | 24' (3.1) | 1'31' (18,6) | 1'63' (19,1) |
| **Group 3** | Participant 7 | 39' | 29' | 32' | 1'40' | 2'45' |
| | Participant 8 | 43' | 38' | 40' | 1'61' | 2'57' |
| | Participant 9 | 35' | 32' | 33' | 1'40' | 2'60' |
| | Mean (SD) | 39' (4,0) | 33' (4,6) | 35' (4,4) | 1'67' (12,1) | 2'54' (7,9) |

Note: SD - Standard deviation
Source: Prepared by the author, 2024.

In both user tests, the division into three groups showed that Group 1 (18 to 29 years old) performed the activities in a shorter time, and had a more homogeneous time distribution, due to the presence of a low standard deviation of the mean. In terms of online flow, the assumption that online environments are better prepared for the flow of younger people is confirmed, since the participants in Group 1 perform the activities more quickly and are more adapted to the resources available virtually. Speed and flow are intimately intertwined variables, a flow-enabled online environment demands tasks that are quick to complete (King, 2003; Oliveira et al., 2023). Thus, the speed with which the participants in Group 1 (considered young) performed the activities, compared to the other two groups, generates indications that the virtual environments are more qualified for the flow of this specific group.

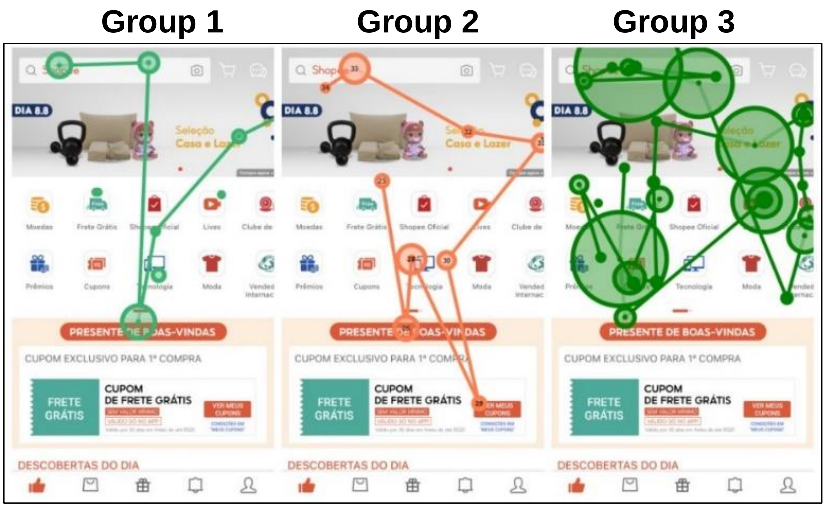
The search for web design elements that generate a reduction in the time it takes to perform tasks by Group 2 (31 to 49 years old) and Group 3 (50 years or older) is crucial, since if the virtual environment is better adapted to the skills of users over 30 years of age, consequently, the experience of Group 1 components will be enhanced.

The user test collected all of the participants' eye movements as they performed the given tasks. First, the recorded data were analyzed using gaze path graphs. These graphs trace the gaze pattern of each participant through a series of circles indicating fixation and lines indicating saccades. The size of the circles represents the duration of a fixation. Short fixations are indicated by small circles, while larger circles indicate a longer fixation. Therefore, an evaluation was made in terms of fixation points, recorded while the participants observed specific elements in the virtual environments. The smaller the fixation point required for the participant to locate a given element, the better its location on the interface will be, as a short fixation of gaze was required to locate it (Baharum et al., 2019).

Concerning the user test developed in Brazil, Figure 3 highlights the graph of the most characteristic gaze path of one participant in each group, during Task 1, in which they were asked to find the search bar of the application and click on it.

The participant in Group 3 (50 years or older) took a closer look at the app's navigation tab, fixing his gaze longer on the search bar and several other icons on the interface; The large circles reveal this fact. Rayner (1998) states that a longer fixation duration is often associated with deeper and more laborious cognitive processing. So, for Group 3, the elements of online flow conditions captivated their gaze, but it took more time to process what they saw. Rayner (1998) also points out that it should be considered that the duration of gaze fixation is an idiosyncratic measure and can only represent a behavioral or structural characteristic peculiar to an individual or group.

Figure 3 – Gaze path graphs of a participant from each age group of the Brazilian user test during Task 1 (navigation tabs)




Source: Results from the BeGaze software prepared by the author.

Observing the participant chosen from Group 1 (18 to 30 years old), it can be seen that he quickly completed the task, fixing his gaze on a smaller number of elements of flow conditions. Tracing a path of the gaze composed of smaller circles and fewer of them. Unlike Groups 2 and 3, which observed more elements in the interface, and even for a longer time. Smaller fixation points indicate a more improved location of elements in the interface (Baharum et al., 2019), and, in this sense, the gaze path characteristic of Group 1 reveals an ease in finding elements of flow conditions and performing tasks in the evaluated online environment.

The initial findings, highlighted in Figure 2, already reveal indications that online shopping environments are designed to awaken the state of online flow more easily in younger users. The Spanish user test corroborates this understanding.

As in the Brazilian user test, in Spain, participants were also asked to find elements of flow conditions within the evaluated online environment. Among the five activities proposed, Figure 4 shows the graph of the path of the most characteristic gaze of each participant, during a stage of Task 1, in which the participants were asked to look for the place where it is possible to log in and click on it.

Figure 4 – Gaze path graphs of a participant from each age group of the Spanish user test during Task 1 (navigation pages)


Source: Results from the Tobii Studio software prepared by the author.

The results highlighted the same conjuncture observed in the Brazilian user test. Short gaze fixations on elements of flow conditions, such as navigation bars, icons, and buttons, by Group 1 (18 to 30 years), which indicates a rapid cognitive process of assimilation of what was seen. In addition, in the path graph of the gaze characteristic of Group 1, few fixation points appeared in comparison with the other two groups evaluated, which Emphasizes a greater adaptability, on the part of the participants of Group 1, to the online environment in which they were inserted. So, the Spanish user test supports the idea that online shopping environments are designed to awaken the state of online flow more easily in younger users.

Concerning elements of online flow conditions, highlighted by King (2003), Rettie (2001), and Guo and Poole (2009), which drew the attention of the participants and corroborated the assumption that elements of online flow conditions captivate the users' gaze, the labels, titles, icons, and signposts full of information scent stand out.

According to Blackmon (2012), information scent is linked to the ability of certain web design elements to serve as a clue to what is next. In the case of online flow, this is essential since people have specific goals and need the help of the information system to achieve them. In both user tests, there were several fixations on these elements, regardless of age group, as can be seen in Figures 2 and 3.

Maes, Van Geel, and Cozijn (2006) point out that signposts in online environments are essential for navigation. According to the authors, signposts are, for the most part, words that point out the meaning of longer entries and act as a visual index that helps the user access what they want faster. In Figure 3 it is possible to see these elements in the header of the evaluated website. In addition, it is observed that it is on them that the participants of all groups fixed their gaze for the longest time, as indicated by the circles of the graph. This reveals an awakening of the cognitive process that seeks to understand what is seen.

The union of elements full of information scent and signposts is ideal since both quickly help the user to better understand the environment in which they are inserted. During the user tests, it became clear that this combination caught the attention of the participants. Thus, these elements, which correspond to the presence of clear objectives, proved to be able to contribute to the enablement of flow and corroborate the assumption that the elements of flow conditions captivate the eye and help users in online navigation.

Another evaluation is made in terms of fixing the participants’ gaze on certain elements existing in the online shopping environments evaluated. And now, to visualize the data extracted from the eye trackers, heat maps have been drawn up that show the distribution and concentration of the participants’ attention through different colors. Heatmaps operate on a graded system of colors: red (longer fixation), yellow, green, and blue (shorter fixation). Each map elaborated here brings together, in a single image, the joint view of the three participants of each group evaluated.

Figure 5 highlights the heat maps of the three groups of the Brazilian user test after they were asked to access the Home tab in the application (Task 2). All participants easily accessed the requested tab through the navigation bar at the bottom of the app. As highlighted by Guo and Poole (2009), navigation bars help the user move through the virtual environment, so their presence is crucial in a flow-enabled platform.

Figure 5 – Heat map for each age group of the Brazilian user test during Task 2 (navigation tabs)


Source: Results from the BeGaze software prepared by the author.

The presence of clear objectives in virtual environments is observed through easy access to navigation tabs is also related (King, 2003). These tabs are accessed through a bar that can contain symbols (as in the Brazilian user test application) or words (as in the Spanish user test website).

Figure 3 shows that Group 3 fixed their gaze longer on the elements that were in the navigation bar of the application, indicating an awakening of the cognitive process to what was being observed; and that Group 1 observed in a moderate way the other elements that existed on the screen, emphasizing that even though they had a specific objective, other elements of flow conditions caught the attention of this group. The exchange of clear information between the user and the information system, the core of a good online flow, was mediated through the existence of a navigation bar.

In flow-enabled online shopping environments, the presence of a navigation bar is essential, as breadcrumb navigation as well. According to Shitkova et al. (2015), this technique used in interfaces is a type of schema that reveals the user's location in a virtual environment, to create a hierarchical structure.

On websites, the presence of breadcrumb navigation is more frequent, and during Task 3 of the Spanish user test, this becomes clear (see Figure 6). In Task

3, participants were initially asked to access a section on the website called “Best Sellers”. This section was located in a navigation bar at the top of the shopping website and was broken down into a trail of other product categories, making it easy to access.

Figure 6 – Heat map for each age group of the Spanish user test during Task 5 (navigation pages)


Source: Results from the Tobii Studio software prepared by the author.

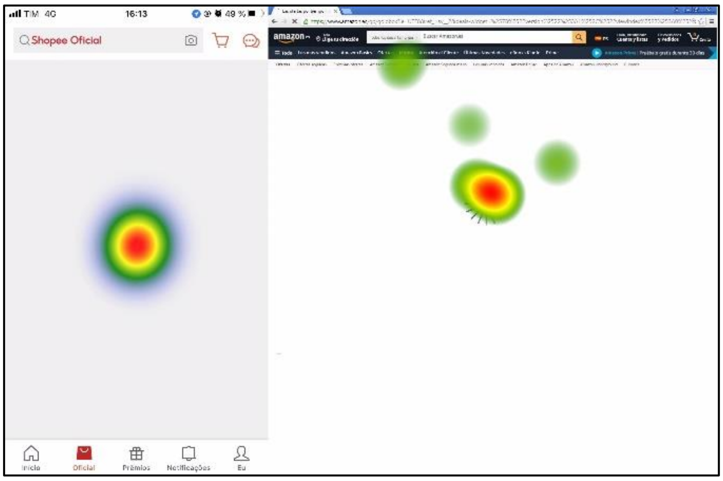
Participants in each of the three groups evaluated differed in their agility in accessing the requested section in such a way that Group 1 (18 to 30 years old) quickly located the section in the navigation bar, while Group 3 (50 years or older) took longer to access the requested section. In this shopping environment, once again what is observed is an adaptation of the interface to the online flow of younger users.

Barthelmäs and Keller (2021) point out that the search for flow consists of the existence of clear objectives, in the sense of a good understanding of the task structure, which is often based on clear instructions. However, they point out that in addition to the existence of previous objectives, the balance between demands and skills is essential. There is no point in having well-formulated goals if I do not have the right tools to achieve them.

The participants in Groups 1 (18 to 30 years old) and Group 2 (31 to 49 years old) showed sufficient skills to perform the activities, due to the fixation of looking at what was requested without major distractions. On the other hand, those in Group 3 (50 years or older) showed that the formulation of these skills was built during the use of the information system since the participants’ gaze wandered in several regions of the navigation page.

During the activities in the two eye-tracking user tests, a loading icon sometimes appeared in the evaluated application and website. This icon is seen as valuable feedback on what is happening (Ayada and Hammad, 2023), i.e., information from the next tab/page is being loaded. Regarding this element of flow conditions, Figure 7 highlights two heat maps. In the center, there is a wide fixation to look precisely at the loading icon during a stage of the requested activities.

Figure 7 – Heat map of Group 3 for the Brazilian user test (right navigation tab) and the Spanish user test (left navigation page)




Source: Results from the BeGaze software prepared by the author.

The two images were brought as an example since this fact was repeated in the three groups evaluated in both user tests. All the participants of the user tests were captivated by the loading icon, understood in the literature as a conditioning element of online flow.

Regarding the user tests carried out, the participants agreed on the fact that this element of online flow is essential. Sensory feedback, transmitted through warnings and messages, also caught the attention of the participants in both user tests and manifested themselves in an essential way for communication with the virtual environment evaluated.

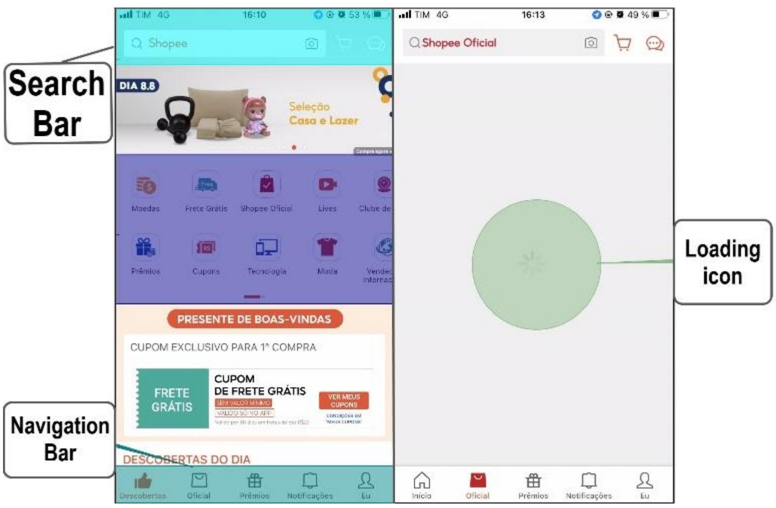
Rettie (2001) and King (2003) point out that the development of a virtual environment should be based on the objective of the system serving as a guide to the user's final goal. The system should always guide the user if they are on the right path. And, in this context, the role of sensory feedback, such as visual changes, messages, or warnings, present in both online shopping environments evaluated, proved to be crucial for participants to continue on their virtual journey, in addition to captivating their gaze.

Still, concerning investigating which age group of users is best conditioned to enter the state of online flow, the BeGaze software, used in the data analysis of the Brazilian user test, allows a deepening of the analysis through quantitative indicators. In it, areas of interest (AOI) were selected, drawing a boundary around a region of the object, and then calculating quantitative measures of eye movement, known as key performance indicators (KPI).

Figure 8 shows, in two navigation tabs of the evaluated application, three AOIs selected based on the elements of online flow conditions, two referring to the clear objectives component (search bar and navigation bar), and one referring to the unequivocal feedback (loading icon), from Table 1.

Three metrics were used: Entry Time (ET), the average duration from the beginning of the video to the first reach of an AOI through the eyes of the participants in each group; the Average Fixation (AF), the sum of the average fixation (sum of the fixation time divided by the number of fixations) in the AOI of each participant divided by the number of participants in the group; and Dwell Time (DT), the sum of all fixations and drawings within an AOI for all participants in the group divided by the number of participants.

Figure 8 – Areas of interest (navigation tabs)




Source: Results from the BeGaze software prepared by the author.

Through the indicators, the speed with which the participants looked for the first time at the highlighted elements of online flow conditions was observed, and the depth of the cognitive processing carried out was evaluated, through the duration of the fixations of gaze on these elements. To make it easier to understand, Table 3 summarizes the calculated metrics.

Table 3 – Key performance indicators of the Brazilian user test (in milliseconds)

| **AOI** | **Search Bar** | | | **Navigation Bar** | | | **Loading icon** | | |
|---|---|---|---|---|---|---|---|---|---|
| **KPI** | **ET** | **AF** | **DT** | **ET** | **AF** | **DT** | **ET** | **AF** | **DT** |
| **Group 1** | 349,7 | 75,4 | 188,4 | 0,1 | 74,8 | 174,6 | 0,1 | 254,9 | 254,9 |
| **Group 2** | 446,4 | 121,9 | 556,9 | 0,1 | 59,1 | 199,5 | 0,1 | 99,7 | 99,7 |
| **Group 3** | 2859,5 | 133,0 | 1319,9 | 0,1 | 112,8 | 908,8 | 0,1 | 277,0 | 277,0 |

Note: AOI = Area of interest; KPI = Key performance indicators; ET = Entry time; AF = Average fixation; DT = Dwell time.
Source: Results from the BeGaze software prepared by the authors.

Since a longer fixation is associated with deeper cognitive processing (Rayner 1998), through Average Fixation it is noticed that the participants in Group 3 (50 years or older) fixed their gaze for a longer time on the navigation bar, indicating that there is an imbalance between the demand to find the bar and the skill possessed by them.  compared to the other two groups. On the other hand, when participants were asked to find the search bar and click on it in Task 1, those in Group 1 (18 to 30 years old) observed it for a longer period. So even for the youngest users, the position of the search bar within an app can be somewhat controversial. It is also noteworthy that the loading icon was observed for a longer period by the participants in groups 1 and 3. Here it is understood that the element captivated the attention of the participants, as expected, and not as something that demanded a greater cognitive effort.

Still about the three AOI highlighted, the indicators of Average fixation and Dwell time were the same in the Loading Icon area, and this is because the element appeared alone and quickly in a single navigation tab, not allowing variation between fixations and saccades, however, both simultaneously reveal the prominence of this element when it is fixed by the gaze of all participants. Dwell Time reveals the sum of the durations of all fixtures and balconies that reached each AOI evaluated. And, about the search bar and the navigation bar, Group 3 (50 years or older) obtained higher values, compared to the other two groups, indicating a longer cognitive process for the understanding of these two elements. For Group 2 (31 to 49 years old), finding the navigation bar proved to be easier than finding the search bar, when comparing the respective Dwell time values. On the other hand, for Group 1 (18 to 30 years old), the assumption that online environments are enabled to awaken the flow state in younger people is confirmed, since they needed fewer fixations and balconies to reach the elements of flow conditions evaluated.

As previously highlighted, the speed of resolution of activities is indicative of the state of online streaming (King, 2003; Novak, Hoffman & Yung, 2000; Finneran and Zhang, 2005, Oliveira et al., 2023). In this context, the three groups evaluated, when requested, looked at the application’s Navigation Bar immediately, which can be perceived through the Entry Time indicator close to zero. In addition to signaling the existence of a balance between demand and skill, this element indicative of clear objectives is intrinsic to the participants' prior knowledge, easy to find, and well disseminated.

Regarding the AOI of Loading Icon, an element of unequivocal feedback, it is confirmed to be able to catch the eye of the participants immediately. For this area, Entry Time indicators close to zero were obtained for the three groups evaluated. This is due to its central and moving characteristics, but also to its wide cognitive diffusion as an indicative element of immediate feedback. Unlike what was previously noted, the participants took time to find the AOI Search Bar. The participants in Group 2 (31 to 49 years old) were the ones who found this element more quickly. The speed with which an AOI is achieved is indicative of clarity regarding its position. In this case, comparatively, participants in all groups took longer to find the search bar than the navigation bar. Searching for the correct positioning of the elements of online flow conditions is essential (King, 2003), and in these cases, the elements are well positioned.

During each of the user tests carried out, in both Brazil and Spain, the participants’ gaze showed that the elements of online flow conditions do indeed draw the user’s attention and are fundamental for them to flow through the virtual environments evaluated. In addition, factors that refer to the transmission of unequivocal feedback and the generation of clear objectives proved to be essential for the construction of an online flow channel. This is because the fixation of the participants’ gaze on certain regions of the virtual environments highlighted the indispensability of the web design elements that condition the flow.

When navigating through the online shopping environments evaluated, the online flow elements captivated the gaze of all participants, regardless of age. It is noteworthy that the findings of the two user tests, evaluated separately by groups, converged when it came to groups of the same age group.

Evidence that the online shopping environments assessed are designed to arouse the state of online flow more easily in younger users (18 to 30 years old) was found through cross-group comparison. The participants in Group 1 (18 to 30 years old) not only performed the activities more quickly, but also showed a greater ability to identify the requested elements, and therefore more likely to flow.

Agility is essential to carrying out online activities, and the age difference between the participants confirmed the idea that the virtual environments evaluated are better designed for the flow of younger users. Since the participants of Group 1, in both user tests, performed the activities more quickly.

The literature around online flow highlights that the speed of task completion is intricately linked to the awakening of the flow state in virtual environments. Users in Group 3 (50 years or older) and those in Group 2 (31 to 49 years old) spent more time performing the activities. While the participants in Group 1 performed the task, in some cases, in half the time spent by Group 3.

The KPIs from the Brazilian user test also corroborated the previous findings. The low dwell time (an indicator of the sum of the duration of all fixations and balconies in an area of the virtual environment) of Group 1, in the search bar (188.4 ms) and the navigation bar (174.6 ms), in comparison with the other two groups, corroborates the assumption that the online environments evaluated are better able to awaken flow state in younger people. Since they required fewer fixations and saccades during the cognitive process of identifying the elements of flow conditions evaluated.

In conclusion, the user tests underscored the practical application and managerial significance of the elements of online flow conditions extracted from the literature (refer to Table 1). While the exploration of flow conditions in virtual environments encompasses broader aspects beyond those specifically highlighted, the current state of understanding in this area remains somewhat nascent. Nonetheless, these identified elements have proven instrumental in guiding the construction of the measurement scale for online flow conditions. They have provided valuable insights into the essential components of virtual environments that contribute significantly to creating environments conducive to facilitating optimal flow experiences. By elucidating these crucial factors, app developers and managers are better equipped to enhance user engagement and satisfaction in online settings, ultimately aiming to optimize the overall user experience.

## 4.2 SCALE ASSESSMENT RESULTS AND DISCUSSION

This section will present the results of the scale assessment and statistical methodology (elucidated in section 3.5), focusing on its dimensions confirmation and validation, followed by the exploratory implementation of Item Response Theory (IRT) modeling.

### 4.2.1 Scale Assessment Framework

According to the theoretical framework, the 43-item set was anticipated to reflect three distinct dimensions. However, the Parallel Analysis scree plot (as presented in Appendix E) did not corroborate this expectation, revealing the presence of more than three dimensions.

However, to achieve better results in assessing the scale’s dimensionality, there is a need to eliminate low-quality items. According to Hair et al. (2019), these items can impact the overall assessment of the scale. Therefore, a factor analysis was conducted. This analysis allows for the removal of low-quality items based on factor loadings and communalities, and it aids in confirming the three expected dimensions.

The factor analysis revealed that four items should be removed from the analysis due to their null variance within the dataset. These items are listed in Chart 6.

Chart 6 – Removed invariant items

| | |
|---|---|
| 9 | The application has a horizontal layout and vertical scrolling. |
| 13 | There is a version of the application for iOS and Android. |
| 29 | The application uses the same font, color scheme, and layout on different pages. |
| 35 | The application has visual differentiation between visited and unvisited sessions. |

Source: Elaborated by the author, 2024.

All the evaluated shopping apps exhibit characteristics represented by items 9, 13, and 29. These include features like a horizontal layout with vertical scrolling, consistent use of font, color scheme, layout across pages, and availability on Android and iPhone platforms. Conversely, item 35, which pertains more to websites than apps, was absent in all apps.

Continuing the analyses, the presence of three dimensions in the scale was assumed, as justified in the theoretical foundation. The correlation between the dimensions and the quality of the items was then assessed through the factor analysis.

Table 4 – Correlation matrix to the three-dimension model

| | **D1** | **D2** | **D3** |
|---|---|---|---|
| **D1** | 1.00 | | |
| **D2** | 0.10 | 1.00 | |
| **D3** | 0.19 | 0.14 | 1.00 |

Source: Results from the R software prepared by the author.

Table 4 displays the correlation coefficients for the three extracted factors (D1, D2, and D3). The low correlation coefficients observed among these factors indicate that they are relatively independent of one another. This suggests that an orthogonal rotation technique, which assumes that the factors are uncorrelated, would be suitable for the Factor Analysis. Consequently, an orthogonal rotation technique was applied to clarify the factor structure and improve interpretability. The results of this orthogonal rotation are detailed in Table 5, which provides a clearer understanding of the distinctiveness and contribution of each factor to the overall analysis.

Table 5 – Factor loadings for the three-dimension model (first round)

| Item | Dimensions | | | Communality |
|---|---|---|---|---|
| | D1 | D2 | D3 | |
| 1 | -0.04 | -0.15 | 0.37 | 0.16 |
| 2 | -0.36 | 0.42 | 0.37 | 0.44 |
| 3 | 0.43 | 0.16 | -0.14 | 0.23 |
| 4 | 0.11 | 0.65 | 0.16 | 0.45 |
| 5 | 0.13 | 0.73 | 0.25 | 0.60 |
| 6 | -0.04 | 0.49 | -0.08 | 0.24 |
| 7 | -0.24 | 0.14 | -0.22 | 0.12 |
| 8 | 0.22 | 0.24 | 0.03 | 0.10 |
| 10 | -0.07 | 0.13 | 0.32 | 0.12 |
| 11 | -0.32 | 0.36 | 0.60 | 0.59 |
| 12 | -0.14 | 0.56 | -0.06 | 0.33 |
| 14 | 0.12 | 0.45 | 0.03 | 0.22 |
| 15 | 0.47 | 0.20 | 0.03 | 0.25 |
| 16 | 0.24 | 0.36 | -0.05 | 0.18 |
| 17 | 0.06 | 0.47 | -0.13 | 0.24 |
| 18 | 0.43 | 0.27 | 0.09 | 0.26 |
| 19 | 0.18 | 0.46 | 0.17 | 0.26 |
| 20 | 0.25 | 0.39 | 0.14 | 0.23 |
| 21 | 0.38 | 0.40 | 0.00 | 0.30 |
| 22 | 0.29 | 0.02 | 0.04 | 0.08 |
| 23 | 0.53 | -0.05 | 0.13 | 0.30 |
| 24 | 0.87 | 0.11 | 0.06 | 0.77 |
| 25 | 0.63 | -0.08 | 0.08 | 0.41 |
| 26 | 0.57 | 0.11 | 0.06 | 0.33 |
| 27 | 0.38 | 0.12 | -0.05 | 0.15 |
| 28 | 0.39 | 0.22 | 0.14 | 0.21 |
| 30 | 0.14 | 0.05 | 0.02 | 0.02 |
| 31 | 0.33 | 0.06 | 0.07 | 0.11 |
| 32 | 0.24 | -0.07 | 0.33 | 0.17 |
| 33 | 0.26 | 0.25 | 0.05 | 0.13 |
| 34 | 0.45 | -0.19 | 0.59 | 0.58 |
| 36 | -0.23 | 0.07 | 0.52 | 0.32 |
| 37 | -0.11 | -0.13 | 0.41 | 0.19 |
| 38 | 0.47 | -0.01 | 0.41 | 0.38 |
| 39 | 0.21 | -0.08 | 0.67 | 0.50 |
| 40 | 0.09 | 0.26 | 0.30 | 0.16 |
| 41 | 0.26 | 0.24 | 0.36 | 0.26 |
| 42 | 0.13 | 0.26 | 0.31 | 0.17 |
| 43 | 0.14 | 0.35 | 0.58 | 0.47 |
| 44 | 0.24 | 0.11 | 0.48 | 0.29 |
| 45 | 0.22 | 0.34 | 0.49 | 0.40 |
| 46 | 0.22 | 0.49 | 0.44 | 0.48 |
| 47 | 0.07 | 0.02 | 0.14 | 0.02 |

Source: Results from the R software prepared by the author.

Items with factor loadings below 0.3 and/or communalities below 0.2 are considered poor and should be removed from the analysis (HAIR et al., 2019). Following this criterion, the 16 bolded items in Table 8 were excluded, and a new Factor Analysis was then conducted with the remaining items. After another two rounds, an additional 6 items were removed, leaving 21 considered good items remaining (see them in Table 6).

Table 6 – Factor loadings for the three-dimension model (last round)

| Item | Dimensions | | | Communality |
|---|---|---|---|---|
| | **Demand-Skill Balance** | **Clear Goals** | **Immediate Feedback** | |
| 2 | 0.68 | -0.27 | 0.16 | 0.57 |
| 4 | 0.56 | 0.09 | 0.26 | 0.38 |
| 5 | 0.69 | 0.22 | 0.28 | 0.60 |
| 6 | 0.54 | 0.07 | -0.10 | 0.31 |
| 11 | 0.55 | -0.35 | 0.53 | 0.71 |
| 12 | 0.63 | 0.05 | -0.13 | 0.42 |
| 15 | 0.20 | 0.53 | 0.03 | 0.32 |
| 18 | 0.32 | 0.60 | 0.10 | 0.47 |
| 21 | 0.34 | 0.54 | 0.01 | 0.40 |
| 23 | -0.15 | 0.45 | 0.18 | 0.26 |
| 24 | -0.06 | 0.89 | 0.16 | 0.82 |
| 25 | -0.17 | 0.68 | 0.12 | 0.51 |
| 26 | 0.04 | 0.60 | 0.16 | 0.39 |
| 34 | -0.26 | 0.28 | 0.82 | 0.82 |
| 36 | 0.16 | -0.24 | 0.45 | 0.29 |
| 38 | -0.12 | 0.37 | 0.63 | 0.55 |
| 39 | -0.07 | 0.14 | 0.77 | 0.62 |
| 41 | 0.28 | 0.29 | 0.43 | 0.35 |
| 43 | 0.46 | 0.09 | 0.64 | 0.63 |
| 44 | 0.08 | 0.13 | 0.51 | 0.28 |
| 45 | 0.26 | 0.16 | 0.44 | 0.29 |

Source: Results from the R software prepared by the author.

The analyses conducted using R software indicated a cumulative variance of 48% across the three dimensions, which is considered satisfactory. As a result, the three-dimensional model demonstrates both qualitative and quantitative coherence. In the following section, confirmation of these dimensions will be conducted, further substantiating their conceptual robustness and empirical validity.

Regarding the 28 items identified as poor and subsequently removed during the Factor Analysis, several conclusions can be drawn. From a statistical perspective, these items are not correlated with the other items, implying that they are not theoretically aligned to measure the conditions of online flow in shopping applications. Many items were excluded from the analysis due to low variance, indicating that these characteristics are commonly found across most applications and are already well-established, as necessary. This makes statistical analysis difficult, as it relies on response patterns and their relationships. For example, items like 8 – It is possible to use zoom to enlarge product images, 10 – The application has infinitely scrollable product sections, 20 – The application has a navigation bar to access its main information, for example, user profile, main page, etc., 31 – The application provides feedback in the form of navigation widgets, such as menus, toolbars, etc., among others.

From a practical standpoint, the individual evaluation of the removed items can aid in understanding their exclusion. The reason for their removal may be linked to poor item wording or a weak relationship with the dimension they were initially designed to measure. For instance, authors such as King (2003) and Guo and Poole (2009) emphasize that onboarding – an initial step-by-step guide that indicates how to use the application – can facilitate navigation in online environments. However, this concept, represented in item 1 – The application has onboarding, an initial step-by-step that indicates how its use works, is less common today as most users are already familiar with the web design of shopping applications.

Regarding feedback provided to users during app usage, item 47 – The application sends messages of a personal nature to the user, such as welcome, good morning, good evening, etc., may be perceived by users as irrelevant, leading to their exclusion from the analysis. Other items, such as 31 – The application provides feedback in the form of navigation widgets, such as menus, toolbars, etc., and 32 – The application provides feedback in the form of performance variables, such as loading server icons, memory usage amount, etc., may already be widely understood as essential in shopping applications, rendering their inclusion in the scale unnecessary.

It is crucial to note that the development of this scale to measure flow conditions is innovative, as the existing literature on this topic is relatively limited, mainly comprising suggestions from various authors. Moreover, these suggestions are often not categorized into specific dimensions beforehand. Therefore, factor analysis plays a critical role not only in eliminating poorly performing items based on their loadings but also in identifying which items are pivotal for the scale.

In the following section, the three extracted dimensions will be validated using previously established instruments.

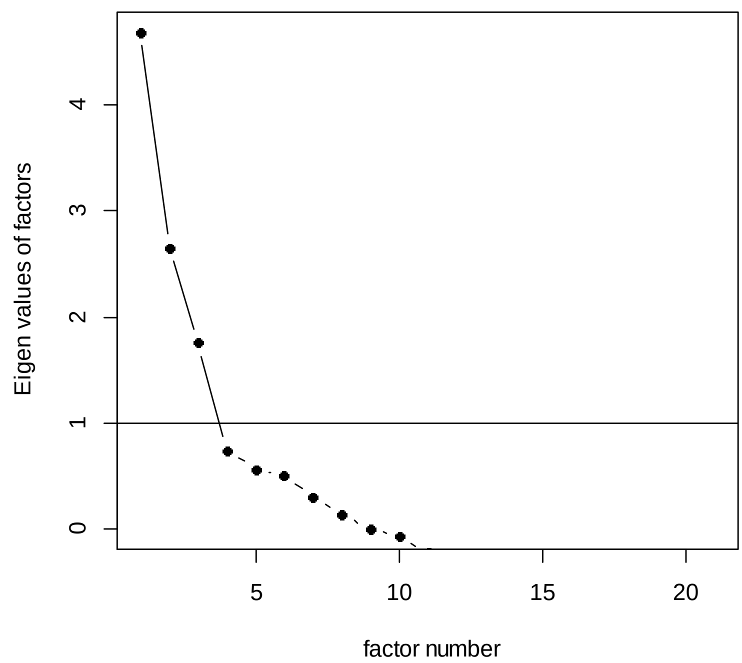
### 4.2.2 Scale Validation and Confirmation

Continuing the analysis initiated in the preceding section, the objective now shifts to the validation and confirmation of the outcomes derived from the factor analysis. This validation process was conducted concurrently with the initial analysis to ensure consistency and reliability.

The method employed to validate the dimensions of the construct was empirical, relying on both the factor loadings of individual items and their conceptual relevance. The Kaiser criterion played a pivotal role in determining the number of dimensions in the scale, as it stipulates that only eigenvalues exceeding 1 in factor analysis are significant indicators of distinct dimensions. Additionally, a confirmatory factor analysis was conducted to further substantiate the scale's internal consistency.

According to Kaiser's criterion (KAISER, 1958), which posits that eigenvalues greater than 1 indicate the number of dimensions, Figure 9 illustrates the presence of three dimensions. This finding is further supported by Parallel Analysis, confirming the identification of three distinct dimensions within the construct through factor analysis.

Figure 9 – Factor analysis scree plot for the three-dimension model




Source: Results from the R software prepared by the author.

To confirm the outcomes of the factor analysis conducted, a structural model was rigorously tested using a confirmatory factor analysis on the complete set of 21 items. The detailed structural model output generated through R software has been documented and is accessible for review in Appendix F. Additionally, comprehensive assessments of internal reliability and convergent validity were conducted for each dimension individually. These analyses provide a thorough evaluation of the reliability and consistency of the dimensions under study. The detailed findings from these evaluations are meticulously presented and summarized in Table 7, offering insights into the robustness and validity of the measurement model employed.

Table 7 – Internal reliability and convergent validity of the dimensions

| Dimension | Item | Factor loading | Item-total correlation | Communality | CR<br>AVE | α<br>ω |
|---|---|---|---|---|---|---|
| Demand-Skill | 2 | 0.68 | 0.47 | 0.57 | 0.65 | 0.68 |
| Balance | 4 | 0.56 | 0.41 | 0.38 | 0.36 | 0.85 |
| | 5 | 0.69 | 0.52 | 0.60 | | |
| | 6 | 0.54 | 0.26 | 0.31 | | |
| | 11 | 0.55 | 0.42 | 0.71 | | |
| | 12 | 0.63 | 0.35 | 0.42 | | |
| Clear Goals | 15 | 0.53 | 0.31 | 0.32 | 0.62 | 0.67 |
| | 18 | 0.60 | 0.32 | 0.47 | 0.24 | 0.78 |
| | 21 | 0.54 | 0.27 | 0.40 | | |
| | 23 | 0.45 | 0.24 | 0.26 | | |
| | 24 | 0.89 | 0.63 | 0.82 | | |
| | 25 | 0.68 | 0.46 | 0.51 | | |
| | 26 | 0.60 | 0.38 | 0.39 | | |
| Immediate | 34 | 0.82 | 0.50 | 0.82 | 0.73 | 0.67 |
| Feedback | 36 | 0.45 | 0.20 | 0.29 | 0.27 | 0.72 |
| | 38 | 0.63 | 0.38 | 0.55 | | |
| | 39 | 0.77 | 0.48 | 0.62 | | |
| | 41 | 0.43 | 0.22 | 0.35 | | |
| | 43 | 0.64 | 0.42 | 0.63 | | |
| | 44 | 0.51 | 0.34 | 0.28 | | |
| | 45 | 0.44 | 0.31 | 0.29 | | |

Note: CR – Composite reliability; AVE – Average variance extracted; α – Cronbach's alpha; ω – McDonald's omega.
Source: Results from the R software prepared by the author.

Cronbach's alpha values for each dimension were approximately 0.7, which were considered acceptable in the technological context according to Daud et al. (2018). It is noteworthy that Nunnally and Bernstein (1994) suggest that values between 0.60 and 0.80 are considered acceptable. Additionally, McDonald's omega, recognized as a qualitatively superior index by Brunner, Nagy, and Wilhelm (2012), exceeded 0.7. Consequently, the internal reliability of all self-reported constructs was deemed minimally adequate.

Convergent validity is underpinned by the entirety of factor loadings, a cornerstone affirmed by Hair et al. (2019), who stipulate a minimum acceptable threshold of 0.3 for each item's factorial loading. The item-total correlation values ranged from 0.2 to 0.89. Field (2018) acknowledges that values around 0.2 can still be considered adequate, provided the overall reliability and validity of the scale are acceptable.

Additionally, composite reliability surpasses 0.6 in each construct, and the Average Variance Extracted (AVE) exceeds 0.2 in all cases. Fornell and Larcker (1981) argue that convergent validity can still be considered adequate even if the AVE is less than 0.5, provided that the composite reliability exceeds 0.6. Specifically, they suggest that while AVE is a critical measure of how well a set of items captures a construct, composite reliability assesses the internal consistency of the items. If composite reliability is high (greater than 0.6), it indicates that the items consistently measure the same construct, which can still support convergent validity even if the AVE is below 0.5. Thus, while AVE is important, strong composite reliability can still affirm the construct's validity.

Furthermore, AVE for each construct exceeds its Maximum Shared Variance (MSV) and Average Shared Variance (ASV) values (see Table 7 and Table 8). Moreover, the square root of AVE for each construct surpasses all correlations between that construct and every other construct, confirming the robust discriminant validity of the model's constructs.

Table 8 – Discriminant validity of the dimensions and bivariate correlations

| Dimension | MSV | ASV | Demand-Skill Balance | Clear Goals | Immediate Feedback |
|---|---|---|---|---|---|
| **Demand-Skill Balance** | 0.048 | 0.026 | (0.60) | | |
| **Clear Goals** | 0.109 | 0.056 | 0.060 | (0.493) | |
| **Immediate Feedback** | 0.109 | 0.079 | 0.220 | 0.330 | (0.522) |

Note: MSV – Maximum Shared Squared Variance; ASV – Average Squared Variance;
( ) Square Root of Average Variance Extracted.
Source: Results from the R software prepared by the author

Therefore, the scale resulting from the factor analysis demonstrates internal consistency, supported by the convergence of these three indices in confirming the coherence of the obtained results.

Subsequently, the assessment of the adequacy of Item Response Theory models commenced, utilizing the scale comprised of the 21 items derived from the Factor Analysis.

### 4.2.3 IRT Models Suitability Check

In this section, the application of Item Response Theory (IRT) modeling will be explored, beginning with unidimensional models for each dimension, followed by multidimensional and bifactor applications. In addition to a comparison between the proficiency levels of the applications available in Brazil and Spain concerning each of the three dimensions of online flow conditions.

#### 4.2.3.1 Multiple Unidimensional IRT Model Suitability Check

In an attempt to apply the unidimensional IRT model, a multiple unidimensional analysis was conducted, considering the presence of three individual constructs. The aim was to assess the feasibility of treating each dimension as an independent scale.

The first model, Mod1, included items from the Demand-Skill Balance dimension; Mod2 included items from the Clear Goals dimension; and Mod3 comprised items from the Immediate Feedback dimension. The parameter estimation results for these models are presented in Tables 9, 10, and 11.

Table 9 – Parameter estimation for Mod1

| Item | a | d |
|---|---|---|
| 2 | 1.144 | -0.784 |
| 4 | 3.911 | -3.090 |
| 5 | 20.216 | -13.202 |
| 6 | 0.676 | 1.008 |
| 11 | 1.074 | -1.074 |
| 12 | 0.937 | 0.702 |

Note: a – discrimination parameter; d – scalar difficulty parameter.
Source: Results from the R software prepared by the author.

The items in Mod 1, corresponding to the Demand-Skill Balance scale, exhibited issues with the estimation of parameters for two items, indicating the inadequacy of the unidimensional model. The removal of these unstable items also destabilizes other items, underscoring that the independent use of the scale may be unfeasible.

Table 10 – Parameter estimation for Mod2

| Item | a | d |
|---|---|---|
| 15 | 0.948 | 1.593 |
| 18 | 0.995 | 2.239 |
| 21 | 0.861 | 2.314 |
| 23 | 0.755 | 1.291 |
| 24 | 26.411 | 28.644 |
| 25 | 3.409 | 2.594 |
| 26 | 1.167 | 0.729 |

Note: a – discrimination parameter; d – scalar difficulty parameter.
Source: Results from the R software prepared by the author.

Regarding Mod 2, which relates to the Clear Goals scale, two items displayed inconsistent discrimination values, while the remaining items exhibited satisfactory discrimination values.

Table 11 – Parameter estimation for Mod3

| Item | a | d |
|---|---|---|
| 34 | 3.127 | 0.822 |
| 36 | 1.009 | -2.768 |
| 38 | 2.396 | 3.408 |
| 39 | 1.946 | 0.999 |
| 41 | 1.121 | 3.000 |
| 43 | 1.257 | 0.107 |
| 44 | 1.346 | 2.411 |
| 45 | 0.850 | -0.255 |

Note: a – discrimination parameter; d – scalar difficulty parameter.
Source: Results from the R software prepared by the author.

Similar to the other two modules, Module 3, which focuses on the Immediate Feedback scale, also included an item with estimation issues. Applying unidimensional Item Response Theory (IRT) modeling to these three dimensions, treated independently as separate scales, revealed problems in estimation during this exploratory model evaluation. Removing these adverse items from the analysis could further reduce the already limited sample size. However, the critical observation here is that the unidimensional modeling approach demonstrated instability.

Considering a multidimensional Item Response Theory model aligns more closely with the concept of online flow conditions, as it allows for the examination of potential interconnectedness among items. Therefore, the analysis proceeded to estimate a Multidimensional Item Response Theory (MIRT) model in the subsequent section, which is expected to provide more statistically robust results.

#### 4.2.3.2 MIRT Model Suitability Check

Building upon the conclusions and verifications elucidated in preceding sections, it is now assumed that the scale encompasses three dimensions with multidimensionality inherent. Accordingly, a multidimensional model was adopted utilizing Item Response Theory (IRT). As expounded upon in Chapter 2, IRT confers distinct advantages, principally by treating items on an individual basis rather than solely in terms of factor aggregation. This method facilitates an in-depth analysis of each item’s quality, difficulty level, and subsequently, the estimation of scores for respondents.

Table 12 presents the parameter estimates for the 21-item set resulting from the Factor Analysis using the multidimensional item response theory (IRT) model, as well as the orthogonal rotation.

Table 12 – Multidimensional IRT results (first round)

| **Item** | **Dimensions** | | | **d** |
|---|---|---|---|---|
| | **D1** | **D2** | **D3** | |
| 2 | 0.330 | -0.917 | 1.379 | -0.998 |
| 4 | 0.249 | 0.296 | 3.451 | -2.742 |
| 5 | 3.713 | 7.115 | 38.216 | -24.414 |
| 6 | -0.267 | -0.219 | 0.814 | 1.037 |
| 11 | 1.437 | -1.190 | 1.470 | -1.613 |
| 12 | -0.320 | -0.305 | 1.081 | 0.712 |
| 15 | -0.066 | 0.843 | 0.528 | 1.655 |
| 18 | 0.078 | 1.072 | 0.964 | 2.591 |
| 21 | -0.025 | 0.782 | 0.807 | 2.510 |
| 23 | 0.333 | 0.823 | -0.019 | 1.407 |

| | | | | |
|---|---|---|---|---|
| 24 | 14.190 | 62.438 | 6.422 | 78.014 |
| 25 | 0.367 | 2.409 | -0.053 | 2.277 |
| 26 | 0.419 | 1.179 | 0.420 | 0.842 |
| 28 | 0.546 | 0.739 | 0.274 | 1.295 |
| 34 | 2.688 | 1.071 | -0.035 | 0.936 |
| 36 | 1.148 | -0.441 | 0.188 | -2.961 |
| 38 | 1.984 | 1.087 | 0.096 | 3.412 |
| 39 | 1.969 | 0.413 | 0.138 | 1.106 |
| 43 | 1.937 | 0.082 | 1.682 | 0.233 |
| 44 | 0.911 | 0.298 | 0.584 | 2.281 |

Note: d – scalar difficulty parameter; D1, D2, and D3 – dimensions extracted.
Source: Results from the R software prepared by the author.

Two items (item 5 and item 24) exhibited incompatible discrimination parameters, leading to their exclusion from the analysis. Following two subsequent rounds, three additional items were also removed, resulting in a total of 16 remaining items. Consequently, Table 14 displays the parameter estimates for this set of items derived from the Factor Analysis using the multidimensional item response theory (IRT). The alignment of items with their respective dimensions identified by the factor analysis is evident (see Table 6 and Table 13).

Table 13 – Multidimensional IRT results (last round)

| **Item** | **Dimensions** | | | **d** | **MDIS** | **MDIF** |
|---|---|---|---|---|---|---|
| | **Demand-Skill Balance** | **Clear Goals** | **Immediate Feedback** | | | |
| 2 | 1.150 | -0.682 | 0.555 | -0.858 | 1.448 | -0.593 |
| 6 | 2.173 | 0.180 | 0.269 | 1.669 | 2.197 | 0.760 |
| 12 | 1.995 | 0.130 | -0.068 | 0.980 | 2.000 | 0.490 |
| 15 | 0.425 | 0.892 | 0.237 | 1.616 | 1.016 | 1.590 |
| 18 | 0.673 | 1.506 | 0.450 | 2.742 | 1.710 | 1.604 |
| 21 | 1.034 | 1.143 | 0.433 | 2.842 | 1.601 | 1.775 |
| 23 | -0.568 | 1.014 | 0.225 | 1.456 | 1.184 | 1.230 |
| 25 | -0.403 | 1.183 | 0.212 | 1.455 | 1.268 | 1.148 |
| 26 | 0.038 | 1.537 | 0.337 | 0.809 | 1.574 | 0.514 |
| 36 | 0.445 | -0.800 | 1.840 | -3.749 | 2.055 | -1.824 |
| 38 | -0.416 | 0.634 | 1.574 | 2.828 | 1.747 | 1.619 |
| 39 | -0.607 | 0.451 | 1.708 | 0.971 | 1.868 | 0.520 |
| 41 | 0.270 | 0.527 | 1.116 | 3.120 | 1.263 | 2.470 |
| 43 | 1.040 | 0.512 | 2.624 | 0.251 | 2.869 | 0.087 |
| 44 | -0.059 | 0.562 | 1.016 | 2.282 | 1.163 | 1.963 |
| 45 | 0.241 | 0.478 | 0.987 | -0.282 | 1.123 | -0.251 |

Note: d – scalar difficulty parameter; MDIS – multidimensional discrimination parameter; MDIF – multidimensional difficulty parameter.
Source: Results from the R software prepared by the author.

In Table 13, aside from delineating the discrimination parameters of each item across dimensions, with emphasis placed on higher values, two additional parameters are presented to assist a deeper comprehension of the multidimensional model. These parameters include the multidimensional discrimination parameter (MDIS) and the multidimensional difficulty parameter (MDIF) for each item, both contributing pivotal insights into the nuanced interplay of factors characterizing the model under scrutiny.

The MDIS is calculated as the square root of the sum of the squared discrimination parameters for each item across all dimensions. The higher the MDIS, the greater the multidimensional discrimination power of the item. Table 10 also includes the value of the scalar difficulty parameter (d) for each item, which, according to Reckase (2009), is related to the item's difficulty. However, the interpretation of this parameter cannot be directly equated to the unidimensional model's difficulty parameter. In the multidimensional model, the d parameter is a scalar, thus presenting only a single value for each item. An interpretable parameter is the negative scalar difficulty parameter (-d) divided by the MDIS resulting in the multidimensional difficulty parameter (MDIF). The MDIF indicates the distance from the origin of the space to the inflection point in the direction of the discrimination vector. Therefore, the higher the MDIF parameter value, the greater the item's difficulty.

The discrimination parameter, akin to factor loading, can be arbitrary across dimensions, assuming that there might be some form of rotation; these loadings can be compared by identifying similar items within the same dimension, much like in factor analysis (RECKASE, 2009). It is anticipated that the discrimination parameter of a given item will be higher in the dimension where it exhibits greater discrimination power, or in other words, where it provides more information.

Therefore, by evaluating the individual values of the discrimination parameters across dimensions (see discrimination parameters for each item in Table 13), it is observed that the configuration aligns with the three-dimensional structure presented in the factor analysis (see Table 9). Thus, the prior factor analysis supports the findings of the IRT modeling.

A confirmation of the MIRT model results was also performed using Confirmatory IRT, and the findings are detailed in Table 14. The results corroborate and confirm the assignment of each item to its respective dimension based on its discrimination, as previously illustrated in Table 14.

Table 14 – Confirmatory IRT results

| Item | Dimensions | | | d |
|---|---|---|---|---|
| | Demand-Skill Balance | Clear Goals | Immediate Feedback | |
| 2 | 0.841 | 0 | 0 | -0.712 |
| 6 | 2.889 | 0 | 0 | 2.025 |
| 12 | 1.726 | 0 | 0 | 0.901 |
| 15 | 0 | 0.957 | 0 | 1.592 |
| 18 | 0 | 1.678 | 0 | 2.723 |
| 21 | 0 | 1.181 | 0 | 2.517 |
| 23 | 0 | 0.709 | 0 | 1.273 |
| 25 | 0 | 1.138 | 0 | 1.397 |
| 26 | 0 | 1.609 | 0 | 0.824 |
| 36 | 0 | 0 | 1.201 | -2.911 |
| 38 | 0 | 0 | 1.499 | 2.635 |
| 39 | 0 | 0 | 1.378 | 0.832 |
| 41 | 0 | 0 | 1.249 | 3.111 |
| 43 | 0 | 0 | 2.137 | 0.197 |
| 44 | 0 | 0 | 1.3 | 2.379 |
| 45 | 0 | 0 | 1.167 | -0.284 |

Note: d – scalar difficulty parameter.
Source: Results from the R software prepared by the author.

An evaluation of the amount of information captured by models with varying numbers of dimensions was conducted. Following the criteria proposed by Bartolucci et al. (2007), one-, two-, three-, and four-dimensional models were analyzed, revealing evidence to support the presence of three dimensions in the 16-item set. This trend is depicted in Table 15.

Table 15 – Model comparison indices

| Model | Log of max. likelihood | AIC | SABIC | $M_2$ | $\chi^2$ test | RMSEA |
|---|---|---|---|---|---|---|
| One-dimension | -1573.743 | 3211.486 | 3215.653 | 253.760 * | - | 0,082 |
| Two-dimension | -1535.125 | 3164.250 | 3170.370 | 162.294 * | 77.236 * | 0,064 |
| Three-dimension | -1510.017 | 3142.034 | 3149.977 | 91.171 * | 50.216 * | 0,032 |
| Four-dimension | -1502.109 | 3152.218 | 3161.854 | 76.388 | 15.815 | 0,034 |

Note: * p-value < 0,05; AIC – Akaike Information Criterion; SABIC – Sample Adjusted Bayesian Information Criterion; $M_2$ – Fit Index; $\chi^2$ – chi-square test; RMSEA – Root-Mean-Square Error of Approximation.
Source: Results from the R software prepared by the author.

The decrease in the Akaike Information Criterion (AIC) and Sample Adjusted Bayesian Information Criterion (SABIC) values up to the three-dimensional model, the attainment of a significant p-value in the chi-square ($\chi^2$) test for the three-dimensional model, the observation that the three-dimensional model exhibits the lowest Root-Mean-Square Error of Approximation (RMSEA) value, and the least significant $M_2$ statistic value collectively advocate for the consideration of the item set as possessing three dimensions. For added clarity, Table 7 enumerates the 16 items forming the resultant scale that measures de online flow conditions.

Chart 7 – Set of items that make up the scale according to MIRT

| | |
|---|---|
| **Demand-Skill Balance** | |
| 2 | The application features an adjustable interface, enabling users to alter it to their preferences. |
| 6 | Users can select their preferred method for receiving notifications from the application, such as via email, SMS, WhatsApp, or other available channels. |
| 12 | The application facilitates customization of its components within the account settings section. |
| **Clear Goals** | |
| 15 | The application enables users to add items to a favorites list. |
| 18 | The application provides a section where users can answer questions, such as a FAQ (Frequently Asked Questions) section or direct contact with customer service. |
| 21 | The application enables users to apply various filters to the lists of available products. |
| 23 | After completing an action, the application displays motivational information, such as promotions and most purchased products, to encourage continued user engagement. |
| 25 | Users can obtain the number of items in the shopping cart without needing to access it directly. |
| 26 | Users have the option to qualify for free shipping within the application. |
| **Immediate Feedback** | |
| 36 | The application delivers feedback audibly. |
| 38 | The application provides feedback in the form of a banner notification. |
| 39 | When entering incorrect input data, the application informs you that there is an error. |
| 41 | The application offers more than one method for product search, such as a search bar, sorting and filtering options, and voice commands, among others. |
| 43 | The application provides alternative methods of logging in, such as authentication via social media accounts like Facebook, Google, and others. |
| 44 | The application offers generic product recommendations on the homepage. |

| | |
|---|---|
| 45 | The application allows users to read and submit product reviews. |

Source: Elaborated by the author, 2024.

A classification of each item based on its discriminatory power or informativeness within the construct, as well as its difficulty level, derived from the estimation of multidimensional difficulty parameters, respectively, are provided in Table 16. The aim is to rank the items on the scale, identifying which items exhibit the strongest relationship with the dimensions and which are more challenging to observe in online environments.

Table 16 – Ranking of items according to the multidimensional discrimination (MDIS) and difficulty (MDIF) parameters

| **Rank** | **Item** | **MDIS** | **Item** | **MDIF** |
|---|---|---|---|---|
| 1 º | 43 | 2.869 | 41 | 2.470 |
| 2 º | 6 | 2.197 | 44 | 1.963 |
| 3 º | 36 | 2.055 | 21 | 1.775 |
| 4 º | 12 | 2.000 | 38 | 1.619 |
| 5 º | 39 | 1.868 | 18 | 1.604 |
| 6 º | 38 | 1.747 | 15 | 1.590 |
| 7 º | 18 | 1.710 | 23 | 1.230 |
| 8 º | 21 | 1.601 | 25 | 1.148 |
| 9 º | 26 | 1.574 | 6 | 0.760 |
| 10 º | 2 | 1.448 | 39 | 0.520 |
| 11 º | 25 | 1.268 | 26 | 0.514 |
| 12 º | 41 | 1.263 | 12 | 0.490 |
| 13 º | 23 | 1.184 | 43 | 0.087 |
| 14 º | 44 | 1.163 | 45 | -0.251 |
| 15 º | 45 | 1.123 | 2 | -0.593 |
| 16 º | 15 | 1.016 | 36 | -1.824 |

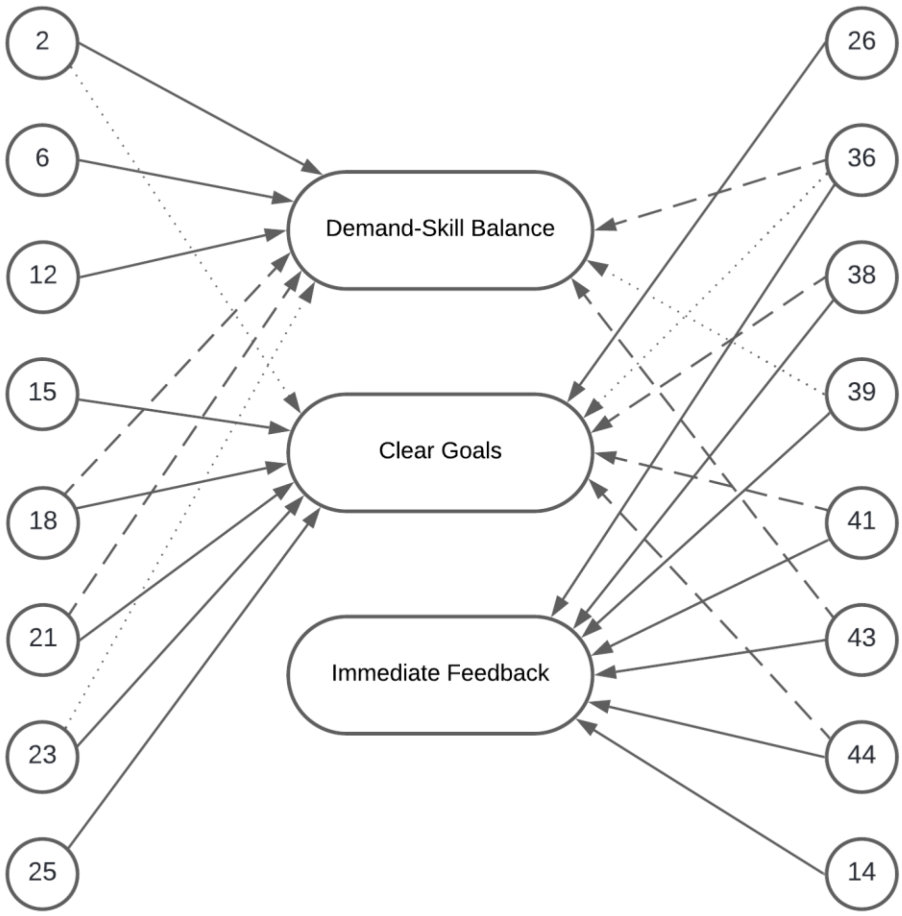
Source: Results from the R software prepared by the author.

It is noteworthy that the three most discriminative items on the scale, namely those with the highest information content within their respective dimensions, are items 43 – The application provides alternative methods of logging in, such as authentication via social media accounts like Facebook, Google, and others –, 6 – Users can select their preferred method for receiving notifications from the application, such as via email, SMS, WhatsApp, or other available channels –, and 36 – The application delivers feedback audibly.

Additionally, the three items with the highest level of difficulty, representing the most challenging online flow conditions within the application, are items: 41 – The application offers more than one method for product search, such as a search bar, sorting and filtering options, and voice commands, among others. –, 44 – The application offers generic product recommendations on the homepage –, 21 – The application enables users to apply various filters to the lists of available products. Therefore, offering various ways for users to search for products proves to be a challenging feature for a shopping app to achieve. While a search bar is expected as a basic feature, providing additional options to users highlights its importance. Another statistically challenging yet significant feature revolves around product recommendations. Both personalized and generic product offerings contribute to generating a state of flow, emphasizing the need for app developers to enhance this aspect to deliver a richer user experience. Another challenging aspect concerns offering product filters to users. A diverse range of filters is crucial for navigation, aiding in decision-making, and managing product displays. Thus, a broader array of filters can enrich the user experience and potentially increase sales by facilitating product discovery.

Items 38 – The application provides feedback in the form of a banner notification – and 18 – The application provides a section where users can answer their questions, such as a FAQ (Frequently Asked Questions) section or direct contact with customer service – are also considered challenging. Providing feedback to the user, regardless of the format, is crucial for enhancing navigation and fostering a state of flow. This interaction between the information system and the user must be executed optimally, with the utilization of banners proving to be significant. Furthermore, it is important to acknowledge that users may encounter questions during navigation, and the online platform should be equipped to address these inquiries effectively to mitigate user abandonment.

Figure 10 – Graphical relationship representation of the remaining items




Note: Solid lines represent the strongest connection; dashed lines represent the secondary positive connection; dotted lines represent the secondary negative connection.
Source: Results from the R software prepared by the author.

Figure 10 illustrates part of the complexity involved in multidimensional modeling. Solid lines represent interactions with the highest discriminatory load, while dashed lines represent secondary interactions with discriminatory load in the same direction as the primary one, and the dotted lines represent secondary inverse strong connection. Only the second connections with an intensity greater than 0.50 were highlighted. Loadings of +0.50 or higher are considered practically significant, and loadings greater than +0.70 are indicative of a well-defined structure (HAIR et al., 2019).

It is possible to observe that some items have secondary connections, such as item 43 – The application allows users to perform social login, an authentication method using social media accounts such as Facebook, Google, and others –, which belongs to the Immediate Feedback dimension but also shows a secondary connection with the Demand-Skill Balance dimension. This secondary connection makes sense, as logging in may require previous skills, especially when it involves a new authentication method. Another item in the Immediate Feedback dimension that also connects with another dimension is item 44 – The application makes generic product recommendations on the home page –, which also indicates a connection with the Clear Goals dimension. This highlights the importance of providing product recommendations that align with users' profiles.

Item 21 – The application enables users to apply filters to the product listings offered – also has a secondary connection, indicating that the ability to apply filters is not only something that users require but also something that may involve previous skill. Mahnke, Benlian, and Hess (2015) emphasize in their work that virtual shopping environments should provide filtering options, and the results of the IRT modeling reveal a multidimensionality of this guidance.

Item 41 – The application's home page has a search bar – and 38 – The application provides feedback in the form of a banner notification – also revealed secondary connections, both with the clear goals dimension, which is coherent since access to the search bar and receiving notifications should align with the user's objectives, as well as assist them in navigation.

To test the resulting scale (Chart 7) in real shopping applications and compare Brazil and Spain, the online flow conditions were estimated for the 200 apps in the research sample. These estimates are derived based on the standard normal scale N(0,1), indicating a mean of zero and a variance of one. The results can be seen in Appendix G. Table 17 presents the percentages of applications where the degree exceeds the average level for each of the three dimensions of online flow conditions, categorized by nationality.

Table 17 – Percentage of shopping applications that have an above-average level for each dimension of online flow conditions (%)

| Shopping app availability | Demand-Skill Balance | Clear Goals | Immediate Feedback |
|---|---|---|---|
| **Brazil** | 56 | 35 | 58 |
| **Spain** | 34 | 67 | 45 |

Source: Results from the R software prepared by the author.

A distinct pattern emerges from the data presented in Table 17, highlighting contrasting proficiency levels across different dimensions of online flow conditions between shopping applications in Brazil and Spain. In Spain, a significant majority of applications excel in Clear Goals (67% of apps), demonstrating a strong capability in organizing virtual layouts. Conversely, Brazilian applications demonstrate higher proficiency in the dimensions of Demand-Skill Balance and Immediate Feedback, with 56% and 58% of apps, respectively, indicating a focus on tailoring the virtual environment to individual user needs and ensuring prompt interaction.

This contrast underscores varying emphases in the design and functionality of shopping applications in each country. Brazil appears to prioritize features that enhance user engagement through personalized experiences and responsive interaction channels. In contrast, applications available in Spain emphasize clarity in task goals and organization within the virtual interface.

Regarding areas that require improvement to foster a more conducive flow-enabled virtual environment, the findings suggest distinct challenges in each context. In Brazil, the area of Clear Goals stands out with only 34% of applications surpassing the average proficiency level. In Spain, a similar challenge is noted in the dimension of Demand-Skill Balance, where 34% of applications demonstrate potential for enhancement. These insights underscore the importance of understanding cultural and contextual factors in shaping user experiences and interface design decisions in digital environments.

For a comprehensive analysis, four applications were selected from the research sample, comprising two available in Brazil (BR 1 and BR 2) and two in Spain (SP 1 and SP 2) applications. The objective is to emphasize the individual analyses applicable to each shopping application, underscoring the managerial insights enabled by IRT modeling. This approach facilitates the identification of the specific strengths and areas needing improvement for each application, with a focus on enhancing the creation of a flow-enabled virtual environment. The findings detailing these analyses are presented in Table 18.

Table 18 – Online flow conditions level estimation for 4 apps based on the MIRT model

| Shopping app | Demand-Skill Balance | Clear Goals | Immediate Feedback |
|---|---|---|---|
| **BR 1** | 0.828 | 0.356 | 0.940 |
| **BR 2** | 1.405 | -0.295 | 0.537 |
| **SP 1** | 0.939 | 0.600 | 0.282 |
| **SP 2** | 1.026 | 0.779 | 0.654 |

Source: Results from the R software prepared by the author.

It is verified that app BR 1 exhibits a greater mastery of items related to the generation of Immediate Feedback, thus needing to further develop its flow conditions, particularly in the dimension of Demand-Skill Balance, and especially in the dimension of Clear Goals. It is noteworthy that negative values reflect that the evaluated app has a proficiency below the average of the 200 evaluated apps, indicating that these areas deserve more attention. This is evident in app BR 2, which generated a degree of online flow conditions of -0.295 in the dimension of Clear Goals. The eye-tracking user test conducted highlighted that this dimension is closely related to the presence of web design elements that assist the user in achieving their objectives, such as accessible navigation bars, labels, and signposts, among other elements. Therefore, a focus on these elements could be beneficial for this app. The other two apps available in Spain (SP 1 and SP 2) exhibit a higher proficiency in the dimension of Demand-Skill Balance. Both would need to focus on elements involving Immediate Feedback, such as the presence of sensory feedback in the form of visual changes, buttons, icons, navigation widgets, messages, or warnings.

However, in a multidimensional framework, a thorough examination of proficiency estimates among respondents, such as the level of online flow conditions of a shopping application in this context, risks oversimplifying the model. Given the complexity and aggregation of information inherent in multidimensional structures compared to their unidimensional counterparts, it becomes imperative to evaluate each dimension individually. This requires a comprehensive examination of each item's association with the level of online flow conditions for each application, as represented in Table 19. Herein lies the probabilistic assessment of correctness (positive response regarding the described characteristic) for each item, based on estimates of flow condition levels assigned to the four applications.

Table 19 – Probability of positive response to the 16-item set for the 4 selected applications

| Item | Shopping apps | | | |
|---|---|---|---|---|
| | BR 1 | BR 2 | SP 1 | SP 2 |
| 2 | 0.724 | 0.554 | 0.699 | 0.717 |
| 6 | 0.784 | 0.614 | 0.794 | 0.813 |
| 12 | 0.749 | 0.635 | 0.779 | 0.788 |
| 15 | 0.686 | 0.502 | 0.697 | 0.730 |
| 18 | 0.735 | 0.467 | 0.745 | 0.781 |
| 21 | 0.750 | 0.529 | 0.773 | 0.802 |
| 23 | 0.585 | 0.330 | 0.584 | 0.608 |
| 25 | 0.621 | 0.355 | 0.624 | 0.654 |
| 26 | 0.658 | 0.409 | 0.668 | 0.700 |
| 36 | 0.792 | 0.316 | 0.697 | 0.736 |
| 38 | 0.739 | 0.285 | 0.661 | 0.718 |
| 39 | 0.771 | 0.211 | 0.673 | 0.732 |
| 41 | 0.730 | 0.440 | 0.692 | 0.740 |

Source: Results from the R software prepared by the author.

From a practical standpoint, the analysis presented in Table 19 demonstrates, based on the level of online flow conditions estimated by the three-dimensional multidimensional model, the items theoretically mastered by the app. Considering the concept of anchor items defined by Beaton and Allen (1992) and Andrade et al. (2000), where one of the criteria to characterize an item at a certain point on the scale is a probability greater than 0.65 at that point on the scale, it is observed that app SP 2 dominates most of the items, except for item 23, which has a probability of a positive response below 0.65.

It is also noted that despite app BR 2 showing a higher level of Demand-Skill Balance (as depicted in Table 14), it has a probability below 0.65 (P = 0.529) of positively responding to item 21, which, according to the model analysis, pertains to a feature associated with the application's ability to allow users to apply filters to the listings of offered products. This is particularly due to it being an item with a difficulty parameter (MDIF = 1.775) among the highest on the scale (according to Table 13). Such analysis can be extended to the remaining items on the scale, meaning that items considered more difficult tend to result in a lower probability of the apps possessing that particular characteristic.

The comprehensive examination of each app, coupled with the assessment of the probability of a positive response for each item, serves to enhance the practical understanding of multidimensional modeling. This approach enables a nuanced evaluation of individual items, highlighting specific areas for improvement. Conversely, the overarching analysis of the degree of online flow conditions offers a relative indication of the dimensions or concepts requiring refinement or further exploration within the app's framework.

#### 4.2.3.3 Bifactor Model Suitability Check

The study also applied IRT bifactor modeling, suitable for assessing latent traits with multiple dimensions, such as the three dimensions of Online Flow Conditions evaluated here. It employed a confirmatory two-factor approach, defining dimensions and grouping the 16 items from the model under Online Flow Conditions in shopping applications; model parameter estimates are detailed in Table 20.

Table 20 – Results from de Bifactor model

| Item | Online Flow Conditions | Demand-Skill Balance | Clear Goals | Immediate Feedback | d |
|---|---|---|---|---|---|
| 2 | 3.444 | -1.942 | 0 | 0 | -1.667 |
| 6 | 2.447 | 1.59 | 0 | 0 | 2.028 |
| 12 | 1.551 | 0.925 | 0 | 0 | 0.922 |
| 15 | 0.308 | 0 | 0.924 | 0 | 1.6 |
| 18 | 0.517 | 0 | 1.547 | 0 | 2.684 |
| 21 | 0.879 | 0 | 1.211 | 0 | 2.758 |
| 23 | -0.574 | 0 | 0.793 | 0 | 1.37 |
| 25 | -0.42 | 0 | 1.466 | 0 | 1.573 |
| 26 | -0.057 | 0 | 1.646 | 0 | 0.835 |
| 36 | 0.721 | 0 | 0 | 1.292 | -3.16 |
| 38 | -0.22 | 0 | 0 | 1.755 | 2.858 |
| 39 | -0.292 | 0 | 0 | 1.934 | 1.004 |
| 41 | 0.664 | 0 | 0 | 1.19 | 3.201 |
| 43 | 1.294 | 0 | 0 | 2.093 | 0.252 |
| 44 | -0.098 | 0 | 0 | 1.354 | 2.421 |
| 45 | 0.329 | 0 | 0 | 0.98 | -0.269 |

Note: d – scalar difficulty parameter.
Source: Results from the R software prepared by the author.

In the bifactor structure, the level of association of the items with the general factor, which in this case represents the Online Flow Conditions in shopping apps, can be observed. Analyzing the loadings associated with the general factor reveals that most items previously identified as characteristic of the Clear Goals and Immediate Feedback dimensions exhibited low loadings on the general factor. Furthermore, some items showed negative loadings, mathematically reflecting the negative loadings observed in both the secondary factors and the multidimensional model presented.

However, examining the secondary loadings of these items shows consistency in the intensity and direction of the parameters, indicating that the characteristics related to the Clear Goals and Immediate Feedback dimensions have a distinct orientation from those associated with the Demand-Skill Balance. This suggests they may represent orthogonal characteristics that cannot be considered as part of a general factor.

Specifically, the fact that the bifactor model assumes orthogonality between the secondary dimensions and the general factor limits the suitability of this model to constructs that have a general factor orthogonal to the other subdimensions, which is not the case for the construct under analysis. In this sense, the latent trait of online flow conditions cannot be represented by a general dimension orthogonal to the subdimensions, at least not for the construct developed in this study. Thus, the comparison of these three models suggests that online flow conditions are a non-unidimensional characteristic, decomposable into three dimensions.

The suitability of the bifactor model versus the MIRT model, both assuming three dimensions, was evaluated using SABIC and AIC information criteria, RMSEA, chi-square test, and M2 fit index, detailed in Table 21.

Table 21 – Bifactor and MIRT models comparison indices

| **Model** | **Log of max. likelihood** | **AIC** | **SABIC** | **$M_2$** | **$\chi^2$ test** | **RMSEA** |
|---|---|---|---|---|---|---|
| Bifactor | -1524.046 | 3144.092 | 3150.342 | 123.036 * | - | 0.044 |
| MIRT | -1510.017 | 3142.034 | 3149.977 | 91.171 * | 28.058 * | 0.032 |

Note: * p-value < 0,05; AIC – Akaike Information Criterion; SABIC – Sample Adjusted Bayesian Information Criterion; $M_2$ – Fit Index; $\chi^2$ – chi-square test; RMSEA – Root-Mean-Square Error of Approximation.
Source: Results from the R software prepared by the author.

Table 21 demonstrates that, based on both the AIC and SABIC criteria, the multidimensional item response theory model (MIRT) exhibits a lower value compared to the bifactor model, suggesting superior suitability of the MIRT model for the dataset. This observation is further supported by the root mean square error of approximation (RMSEA), which indicates reduced error in the MIRT model. Additionally, the M2 index, analogous to the chi-square statistic, is extensively employed in validating item response theory models, where a lower M2 value signifies, a better model fit compared to alternative specifications. These findings collectively affirm that the previously presented MIRT model is the most appropriate choice according to all indices examined.

## 4.3 CHAPTER SUMMARY

Chapter 4 delves into the research findings and their implications, organized into three primary sections. Initially, it covers the eye-tracking user tests conducted in Brazil and Spain (Section 4.1), offering a thorough analysis of the results and their relevance to the research objectives, along with a critical examination of the insights gained and their broader implications. The results enhanced the understanding of online flow conditions in shopping apps through the participants' gaze. The Brazilian user test simulated shopping app usage, while the Spanish user test simulated shopping website usage. Both user tests highlighted crucial web design elements for online flow conditions, such as breadcrumb navigation, signposts, navigation bars, and sensory feedback, among others. Following this, the chapter discusses the evaluation of the scale designed to measure online flow conditions in shopping apps (Section 4.2), assessing its effectiveness, highlighting its strengths and limitations, and exploring its practical implications for future research and applications. The three dimensions were confirmed through factor analysis and confirmatory factor analysis, supported by a structural model, with all fit indexes being satisfactorily met. Subsequently, the results of the exploratory suitability check of the IRT modeling are presented. The Multidimensional model showed the best fit, leading to the analysis of its statistics. The final scale, consisting of 16 items, now aids in measuring the degree of online flow conditions in online shopping environments. Overall, it was found in this session that the analysis of dimensionality and the verification of item quality is fundamental for defining possible models that fit the data reality. The unidimensional model is not always the most suitable, just as the multidimensional or bifactor model is not always the best choice. It primarily depends on the nature of the items and the characteristics of the respondents, thus requiring an in-depth analysis of each model.

# 5 FINAL CONSIDERATIONS

In this chapter, the final considerations of the study are presented, highlighting conclusions, underscoring the research contributions, and limitations, in addition to providing recommendations for future studies.

## 5.1 RESEARCH CONCLUSION

This study represents a significant advancement in understanding and measuring online flow conditions within shopping applications. Central to this research is the development of a novel scale specifically designed to assess these conditions. The scale underwent rigorous validation using robust statistical methods to ensure its reliability and validity. Employing sophisticated multidimensional Item Response Theory (IRT) modeling techniques, the study successfully identified and validated sixteen specific items capable of effectively measuring online flow across three critical components: Demand-Skill Balance, Clear Goals, and Immediate Feedback within virtual shopping environments.

Furthermore, through a comparative analysis of shopping applications, distinct patterns of strengths in online flow conditions were revealed: applications available in Brazil exhibited notable proficiency in Demand-Skill Balance and Immediate Feedback, while applications available in Spain demonstrated superiority in Clear Goals. This comparative approach not only highlights regional differences in user experience priorities but also underscores the applicability of the developed scale across diverse cultural and market contexts.

Methodologically, the study was guided by a comprehensive exploration of both theoretical foundations and practical implications related to inducing the online flow state in shopping applications. This involved an extensive systematic review aimed at defining and conceptualizing online flow conditions, thereby addressing existing gaps in the literature and the inherent challenges of comparing various measurement instruments.

To operationalize the conceptual framework, it was developed a set of 47 meticulously crafted objective items tailored specifically for assessing online flow conditions within the unique contexts of shopping applications available in Brazil and Spain. These items were subsequently deployed across a diverse sample comprising 200 apps, establishing a robust empirical foundation for constructing a comprehensive measurement model. This model not only serves as a valuable tool for guiding app developers in optimizing user engagement and commercial outcomes but also enhances understanding of how technical and non-technical factors interact to influence online flow dynamics.

A critical aspect of the study's methodology was the exploration of dimensionality issues, shouted in Flow Theory and substantiated through rigorous statistical analyses. The findings supported the adoption of a sophisticated two-parameter multidimensional IRT model, which emphasized the importance of treating each dimension and item individually to accurately capture the nuanced aspects of online flow conditions.

This research makes a substantive contribution to advancing the field by proposing a refined multidimensional model based on IRT for measuring online flow conditions in shopping applications. The model's alignment with established theoretical frameworks provides valuable insights for both theoretical development and practical application, offering potential avenues for enhancing user experiences and refining interface design strategies within the dynamic landscape of digital commerce. The study not only enhances our understanding of online flow conditions but also provides a methodological plan for future research endeavors in similar domains

The study's findings have important implications for practitioners and researchers alike, offering actionable insights into enhancing user engagement and satisfaction in digital shopping environments. By identifying specific strengths and areas for improvement across different cultural contexts, the research facilitates targeted interventions to optimize online flow conditions and thereby maximize commercial outcomes. This nuanced approach acknowledges the complex interplay between user expectations, technological capabilities, and interface design in shaping the online shopping experience.

Moreover, the study emphasizes the significance of multidimensional approaches in assessing online flow conditions, challenging traditional unidimensional models by highlighting the differential impacts of distinct flow components. This methodological innovation not only enriches our theoretical understanding but also informs practical strategies for developing and refining digital platforms to foster optimal user experiences.

By advancing the knowledge of online flow conditions in shopping applications, this study contributes to the broader discourse on user experience in digital environments. It sets a precedent for future investigations seeking to enhance theoretical frameworks and methodological approaches in understanding and optimizing online flow conditions across diverse contexts and platforms.

## 5.2 RESEARCH CONTRIBUTIONS

This research is grounded in two foundational theoretical frameworks: Csikszentmihalyi's (1975) Flow Theory and the concept of online flow as initially proposed by Novak, Hoffman, and Yung (2000). The Flow Theory elucidates an optimal state of engagement where individuals experience heightened focus and immersion in activities. In this study, particular emphasis was placed on the conditions of flow in virtual environments, specifically addressing the first three components of Flow Theory – Demand-Skill Balance, Clear Goals, and Immediate Feedback – which are crucial in facilitating the flow state. The research underscored the significance of these elements in promoting flow through robust statistical modeling.

According to Flow Theory, clear goals, balanced demand and skill levels, and immediate feedback enhance the likelihood of achieving the flow state, thereby fostering engagement and concentration on tasks. Digital environments that effectively promote flow not only drive higher conversion rates but also enhance user perceptions and satisfaction with the online platform. This study aimed to identify web design elements that characterize these flow conditions to aid app developers in creating conducive online environments.

Understanding the elements that foster online flow is pivotal for maintaining user engagement. User-centered design, which prioritizes flow conditions, plays a critical role in promoting this cognitive state. Continual adaptation and updating of online environments are essential to align with evolving user preferences and behaviors, ensuring the efficacy of flow promotion strategies. Moreover, striking a balance between system demands and user skills is instrumental in stimulating online flow. Adapting digital environments to individual user preferences and incorporating interactive elements that enhance engagement and reward users are effective strategies in this regard.

The economic and marketing benefits for companies that sustain users in a flow state are substantial, including heightened customer loyalty, improved conversion rates, and increased user satisfaction. Extending Flow Theory to virtual environments enabled researchers to apply established theoretical concepts in digital contexts, resulting in the concept of online flow. This study exclusively focused on Online Flow Conditions, which serve as condition elements of the flow state and aid users in achieving this desired cognitive state.

The developed measurement scale in this research comprises statistically validated items that contribute significantly to understanding these conditions, offering valuable insights for the development of online environments conducive to flow. This scale enables app developers to quantitatively assess the presence and intensity of online flow conditions in virtual settings. This measurement aspect is crucial for app developers to gauge how effectively their platforms foster flow experiences, thereby enhancing user engagement and satisfaction.

The practical contribution of this research is evident in the developed measurement scale, which addresses a gap in the literature by focusing on determinants of online flow. Consisting of three dimensions to assess Online Flow Conditions, the scale includes 16 items that can guide the development of applications and virtual environments to enhance user engagement and retention.

Furthermore, this study evaluated the flow conditions of 200 shopping applications available in Brazil and Spain, using IRT modeling to highlight areas where each application lacks sufficient flow conditions. This analysis contributes significantly to the management of these online platforms, offering insights into areas needing improvement to provide a more flow-enhanced virtual environment for users.

Understanding the elements that promote online flow is crucial for sustaining user engagement. The developed scale provides concrete web design guidelines for creating digital environments that foster flow, optimizing user concentration, engagement, and platform usage time. Investing in promoting online flow not only enhances user experience but also provides competitive advantages, strengthening brands in the digital marketplace. Thus, this thesis contributes to understanding and applying these essential elements for creating virtual environments conducive to flow and optimizing user engagement and platform usage.

## 5.3 RESEARCH LIMITATIONS AND RECOMMENDATIONS

In the present study, the development of items and the preliminary identification of dimensions were based on a literature review, without conducting user tests or interviews to identify their satisfaction and/or preferences. Therefore, it represents an objective evaluation of the information system under analysis. The application of the methodology presented in this study in user tests or questionnaires to identify their preferences or satisfaction represents an unexplored opportunity and could signify a valuable advancement for the field, enabling comparisons between apps deemed proficient in facilitating flow and users' perceptions of whether they have experienced such fluidity.

Moreover, the purchasing process in apps can be seen as a much broader chain of processes than just the information system, as addressed in this study. Authors such as Rosenbaum, Otalora, and Ramírez (2021), among others, consider the online purchasing process from the initial contact with the app to post-sale, including negotiation and the physical organization of goods transportation. Therefore, expanding the approach of this study to encompass the entire online purchasing process is recommended, contributing not only to web design but also to organization and marketing disciplines.

One limitation of this study was the use of apps available only in Brazil and Spain. This limitation could be overcome in future research, allowing for the evaluation of other cultural and linguistic influences in the analysis, and assisting the way for possible application of the Differential Item Functioning (DIF) technique within Item Response Theory (IRT).

Future research avenues could also explore the longitudinal dynamics of online flow conditions, considering how user behaviors and technological advancements influence the sustainability of flow experiences over time. Additionally, comparative studies across broader geographical regions could provide further insights into global variations in online flow dynamics and their implications for digital commerce strategies.

Another boundary of the study was the exclusive use of shopping apps. However, a diverse arrangement of applications exists, each characterized by distinct features and goals, including those oriented towards indirect commercial objectives such as gaming applications. These apps may have flow condition concepts that differ from those addressed in this study, providing an alternative for applying the methodology presented.

The recommendations above suggest other possible approaches based on the methodology presented here. Furthermore, it is possible to develop new strategies or refine this methodology. For instance, the current approach uses a set of 16 items. Given that this involves evaluating information systems and data collection does not solely rely on users, this set of items could be expanded to assess other dimensions of online flow, focusing on the characteristics and impacts of this flow.

Moreover, increasing the number and quality of items can aid in developing an automated approach for data collection and applying Computerized Adaptive Tests (CAT) to evaluate and diagnose online flow conditions in apps. This approach is facilitated by the Item Response Theory methodology, among other techniques applicable to Information and Communication Technology (ICT) contexts.

## REFERENCES

AGGARWAL, D.; SHARMA, D.; SAXENA, A. B. Enhancing the online shopping experience of consumers through artificial intelligence. **International Journal of Information Technology & Computer Engineering (IJITC)**, v. 4, n. 02, p. 1–5, 2024.

Al-KHALIFA, H.S. The accessibility of Saudi Arabia government Web sites: an exploratory study, **Universal Access in the Information Society,** v.10, n.4, 2010.

AKAIKE, H. Information Theory and an Extension of the Maximum Likelihood Principle. In: Petrov, B.N. and Csaki, F., Eds., International Symposium on Information Theory, 267-281, 1973.

AMEEN, N.; TARHINI, A.; MAHMOOD, S.; MADICHIE, N. O. Going with the flow: smart shopping malls and omnichannel retailing. **Journal of Services Marketing**, v. 35, n. 3, p. 325-348, 2021.

ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. D. C. **Item Response Theory: Concepts** and Applications. São Paulo: SINAPE, 2000. 164p.

ANWYL-IRVINE, A. L.; MASSONNIÉ, J.; FLITTON, A.; KIRKHAM, N.; EVERSHED, J. K. Realistic precision and accuracy of online experiment platforms, web browsers, and devices. **Behavior Research Methods**, v. 51, n. 3, p.1146-1164, 2019.

AYADA, W. M.; HAMMAD, M. A. E. E. Design Quality Criteria for Smartphone Applications Interface and its Impact on User Experience and Usability. **International Design Journal**, v. 13, n. 4, p. 339-354, 2023.

BABBIE, E., **The Basics of Social Research**. Wadsworth Publishing, Belmont, CA, 2005.

BARTA, S.; FLAVIÁN, C.; GURREA, R. Managing consumer experience and online flow: Differences in handheld devices vs PCs. **Technology in Society**, v. 64, 2021.

BARTA, S.; GURREA, R.; FLAVIÁN, C. The role of flow consciousness in consumer regret. **Internet Research**, v. 32, n. 3, p. 875-89, 2022.

BARTHELMÄS, M.; KELLER, J. Antecedents, Boundary Conditions and Consequences of Flow. In: PEIFER, C.; ENGESER, S. (eds.). **Advances in Flow Research**. Cham: Springer, 2021.

BARTOLUCCI, F. A class of multidimensional IRT models for testing unidimensionality and clustering items. **Psychometrika**, v. 72, p.141–157, 2007.

BENN, A.; et al. Processos de compra em websites de supermercados. **International Journal of Retail & Distribution Management**, v. 43, n. 4/5, p. 332-350, 2015.

BERTONCINI, A. L. C.; SERAFIM, M.; BORBA, E. H. Escalas de mensuração e antecedentes motivacionais de ambientes virtuais de aprendizado fomentados pela inteligência artificial. **Revista Tecnologias Educacionais em Rede - RETER**, v. 4, p. 1-19, 2023.

BILGIHAN, A. Gen Y customer loyalty in online shopping: An integrated model of trust, user experience and branding. **Computers in Human Behavior**, v. 61, p. 103-113, 2016.

BILGIHAN, A.; OKUMUS, F.; NUSAIR, K..; BUJISIC, M. Online experiences: flow theory, measuring online customer experience in e-commerce and managerial for the lodging industry. **Inf Technol Tourism**, v. 14, n. 1, p. 49-71, 2013.

BIRNBAUM, A. Some latent trait models and their use in inferring an examinee's ability. In: LORD, F. M. and NOVICK, M. R. (Ed.). **Statistical theories of mental test scores**. Reading, MA: Addison-Wesley, 1968. chapter 17-20, p.395-479.

BLACKMON, M. H. Information scent determines attention allocation and link selection among multiple information patches on a webpage. **Behaviour & Information Technolog**y, 31(1), 3–15, 2011.

BOLEN, M.C.; OZEN, U. Understanding the factors affecting consumers' continuance intention in mobile shopping: the case of private shopping clubs. **Int. J. Mobile Communications**, v. 18, n. 1, 2020.

BOONE, H. N.; BOONE, D. A. Analyzing Likert data. **Journal of Extension**, v. 50, n. 2, p. 1-5, 2012.

BORBA, E. H.; TEZZA, R. Avaliação da ludicidade na intenção de uso de m-commerce de moda. **Revista de Ciências da Administração**, v. 25, p. 1-20, 2023.

BORBA, E. H.; TEZZA, R. Aplicação de escalas de teoria de resposta ao item na área de tecnologia da informação e comunicação. **Revista de tecnologia aplicada**, v.11, n.1, p. 21-33, 2022.

BORBA, E. H.; TEZZA, R. Analysis of users' perceptions on mobile electronic commerce of fashion. **Revista Alcance**, v. 28, p. 197-211, 2021 a.

BORBA, E. H.; TEZZA, R. Factors affecting the adoption of technologies applied to the fashion sector. **Revista de Sistemas e Computação**, v. 11, p. 23-32, 2021 b.

BOZDOGAN, H. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. **Psychometrika**, v. 52, n. 3, p. 345-370, 1987.

BROWN, T. A. **Confirmatory Factor Analysis for Applied Research**. New York: Guilford Press, 2015.

BRUNNER, M.; NAGY, G.; WILHELM, O. A tutorial on hierarchically structured constructs. **Journal of Personality**, v. 80, n. 4, p. 796-846, 2012.

BURRELL, G.; MORGAN, G. **Sociological Paradigms and Organizational Analysis**. London: Heinemann Educational Books.

CHAFFEY, D.; ELLIS-CHADWICK, F.; MAYER, R.; JOHNSTON, K. **Digital marketing: Strategy, implementation and practice**. 7. ed. Pearson Education Limited, 2020.

CHALMERS, R.P., mirt: A Multidimensional Item Response Theory Package for the R Environment, **Journal of Statistical Software**, v.48, n.6, p.1-29, 2012.

CHAN, S. S.; FANG, X.; BRZEZINSKI, J. R.; ZHOU, Y.; XU, S.; LAM, J. Usability for Mobile Commerce Across Multiple Form Factors**. J. Electron. Commer. Res.**, v. 3, p. 187-199, 2002.

CHEN, T.; HSU, T.; LU, Y. Impact of flow on mobile shopping intention. **Journal of Retailing and Consumer Services**, v. 41, p. 281-287, 2018.

CHERNYSHENKO, O. S., STARK, S., CHAN, K.-Y., DRASGOW, F., WILLIAMS, B. Fitting item response theory models to two personality inventories: Issues and insights. **Multivariate Behavioral Research**, v.36, p.563–562, 2001.

CHIOU, W.-C.; LIN, C.-C.; PERNG C., A strategic framework for website evaluation based on a review of the literature from 1995–2006. **Information & Management**, v.47, p. 282–290, 2010.

CHOI, B.C.F., KIRSHNER, S.N., WU, Y. Swiping vs. Scrolling in Mobile Shopping Applications. In: Nah, FH., Tan, CH. (eds) HCI in Business, Government, and Organizations: eCommerce and Innovation, HCIBGO, 2016. **Lecture Notes in Computer Science**, v. 9751. Springer, Cham.

CLARK, L. A.; WATSON, D. Constructing validity: Basic issues in objective scale development. **Psychological Assessment**, v. 7, n. 3, p. 309-319, 1995.

CNMC - Comisión Nacional de los Mercados y la Competencia. **CNMC Data**. 2022. Available in: <https://data.cnmc.es/comercio-electronico>. Access: 11 Jan. 2023.

CRESWELL, J. W. **Research Project:** qualitative, quantitative, and mixed methods. 2. Ed. Porto Alegre: Artmed, 2007. 248p.

CRONBACH, L. J. Coefficient alpha and the internal structure of tests. **Psychometrika**, v. 16, n. 3, p. 297-334, 1951.

CRONBACH, L. J. **Fundamentos da testagem psicológica**. Porto Alegre: Artes Médicas., 1996.

CSIKSZENTMIHALYI, M. **Beyond Boredom and Anxiety**. San Francisco: Jossey-Bass, 1975.

CSIKSZENTMIHALYI, M. **Finding Flow**: The Psychology of Engagement with Everyday Life. New York: Harper Collins Publishers, 1997.

CSIKSZENTMIHALYI, M. **Flow**: Thepsychology of Optimal Experience. New York: Harper and Row, 1990.

CSIKSZENTMIHALYI, M; CSIKSZENTMIHALYI, I. S. **Optimal Experience**: Psychological Studies of Flow Consciousness. New York: Cambridge University Press, 1988.

DAUD, K.A.M., KHIDZIR, N.Z., ISMAIL, A.R. AND ABDULLAH, F.A. Validity and reliability of instrument to measure social media skills among small and medium entrepreneurs at Pengkalan Datu River. **International Journal of Development and Sustainability**, v. 7, n. 3, p. 1026-1037, 2018.

DE AYALA, R. J. **The theory and practice of item response theory**. New York: The Guilford Press, 2009. 448p.

DEVELLIS, R. F. **Scale Development:** Theory and Applications. 4ed. Los Angeles: SAGE, 2017. 280p.

DHAR, S.; VAESHNEY, U. Challenges and business models for mobile location-based services and advertising. **Communications of the ACM**, v. 54, n. 5, p. 121-128, 2011.

DJAMASBI, S.; SIEGEL, M.; TULLIS, T. Improving visual behavior research in communication science: An overview, review, and reporting recommendations for using eye-tracking methods. **Communication Methods & Measures**, v. 13, n. 3, p. 239-256, 2019.

DOWNING, S. M. Item response theory: Applications of modern test theory in Medical Education. **Medical Education**, v.37 p.739–745, 2003.

DUCHOWSKI, A. T. **Eye Tracking Methodology:** Theory and Practice. 3 ed. Springer, 2017.

DUHAN, P.; SINGH, A. **M-COMMERCE**: Experiencing the Phygital Retail. Palm Bay: Apple Academic Press Inc., 2019.

EMBRETSON, S.; REISE, S. P. **Item Response Theory for Psychologists**. Psychology Press, 2013.

ESTEBAN-MILLAT, I; MARTINEZ-LOPEZ, J.J.; HUERTAS-GRACIA, R.; MESEGUER-ARTOLA, A.; RODRIGUEZ-ARDURA, I. Modelling students' flow experiences in an online learning environment. **Computers & Education**, v. 7, p. 111–123, 2014.

FIELD, Andy. **Discovering Statistics Using IBM SPSS Statistics**. 5. ed. Thousand Oaks: SAGE Publications, 2018.

FINNERAN, C.M.; ZHANG, P. A person–artefact–task (PAT) model of flow antecedents in computer-mediated environments. **International Journal of Human-Computer Studies**, v. 59, n. 4, p. 475–496, 2003.

FLECK, M. P.A., BOURDEL, M. C. Método de simulação e escolha de fatores na análise dos principais componentes, **Revista Saúde Pública**, v.32, n.3, p267-72, 1998.

FORD, S., MIUCCIO, M., SHIRLEY, L., DURUAKU, F., & WISNIEWSKI, P. Mobile Augmented Reality Design Evaluation. In: Proceedings of the **Human Factors and Ergonomics Society Annual Meeting**, v. 66, n. 1, p. 1275-1279, 2022.

FORNELL, C.; LARCKER, D. F. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. **Journal of Marketing Research**, v. 18, n. 1, p. 39-50, 1981.

FOSTER, G.C.; MIN, H.; ZICKAR, M. J. Review of Item Response Theory Practices in Organizational Research: Lessons Learned and Paths Forward. **Organizational Research Methods,** v. 20, n. 3, p. 465-486, 2017.

GAO, L.; WAECHTER, K.A.; BAI, X. Understanding consumers' continuance intention towards mobile purchase: A theoretical framework and empirical study – A case of China. **Computers in Human Behavior**, v. 53, p. 249–262, 2015.

GIBBONS, R. D., BOCK, R. D., HEDEKER, D., WEISS, D. J., SEGAWA, E., BHAUMIK, D. K., KUPFER, D. J., FRANK, E., GROCHOCINSKI, V. J., STOVER, A. Full-information item bi-factor analysis of graded response data. **Applied Psychological Measurement**, v.31, p.4–19, 2007.

GIBBONS, R. D.; HEDEKER, D. R. Full-information Item Bi-Factor Analysis. **Psychometrika**, v. 57, p. 423-436, 1992.

GIL, A. C. **How to develop research projects**. 4° ed São Paulo: Atlas, 2002.

GIOVANNINI, C. J. et al. The effects of trust transference, mobile attributes, and enjoyment on mobile trust. **Brazilian Administration Review,** v. 12, n. 1, p. 88-108, 2015.

GREWAL, D.; LEVY, M.; KUMAR, V. **Marketing** (6a ed.). New York: McGraw-Hill Education, 2018.

GUNASEKARAN, A.; NGAI, E.W.T. Editorial: Strategic Enterprise Information Systems for Global Supply Chain Competitiveness. **The Journal of Strategic Information Systems**, v. 20, n.3, p. 215-216, 2011.

GUO, X.; et al. Processos de compra em websites de moda. **Journal of Fashion Marketing and Management**, v. 19, n. 4, p. 375-391, 2015.

GUO, Y. M.; KLEIN, B. D. Beyond the Test of the Four Channel Model of Flow in the Context of Online Shopping. **Communications of the Association for Information Systems**, v. 24, n. 48, p. 837-856, 2009.

GUO, Y. M.; POOLE, M. S. Antecedents of flow in online shopping: a test of alternative models. **Information Systems Journal**, v. 19, n. 4, p. 369–390, 2009.

HAESNER, M.; et al. Captura da atenção do usuário para elementos de interface e imagens em plataformas de comércio eletrônico. **Electronic Commerce Research and Applications**, v. 32, p. 100908, 2018.

HAIR, J. F., BLACK, W. C., BABIN, B. J., ANDERSON, R. E. **Multivariate data analysis**. 8. ed. Cengage Learning: Boston, MA., 2019.

HAMBLETON, K. Emergence of Item Response Modeling in Instrument Development and Data Analysis. **Medical Care,** v.38 n.9, p. 60-65, 2000.

HAMBLETON, R. K., SWAMINATHAN, H., ROGERS, H.J. **Fundamentals of item response theory**. Newbury Park, CA: Sage, 1991.

HARTIG, J., HÖHLER, J. Multidimensional IRT models for the assessment of competencies. **Studies in Educational Evaluation**, v.35, p.57-63, 2009.

HERRANDO, C.; JIMENEZ-MARTINEZ, J.; HOYOS, M. J. M. Social Commerce Users’ Optimal Experience: Stimuli, Response and Culture. **Journal of Electronic Commerce Research,** v. 20, n. 4, 2019.

HERRANDO, C.; JIMENEZ-MARTINEZ, J.; HOYOS, M. J. M. The state of flow in social commerce: the importance of flow consciousness. **International Journal of Electronic Marketing and Retailing,** v. 9, n. 4, 2018.

HINMAN, R. **The Mobile Frontier: A Guide for Designing Mobile Experience**, Rosenfeld Media, Brooklyn, 2012.

HOFFMAN, D.L.; NOVAK, T. P. Flow online: lessons learned and future prospects. **J. Inter. Mark**., v. 23, p. 23–34, 2009.

HOLMQVIST, K.; NYSTRÖM, M.; ANDERSSON, R.; DEWHURST, R.; JARODZKA, H.; VAN DE WEIJER, J. **Eye Tracking**: A Comprehensive Guide to Methods and Measures. Oxford University Press, 2011.

DATAREPORTAL. **Digital 2024**: Global Overview Report. 561p. 2024. Available in: <https://datareportal.com/reports/digital-2024-global-overview-report>. Access: 06 Jun. 2024.

HUDDLESTON, E.; WOODS, D.; DUMAS, J. S. Designing eye-tracking experiments to reflect natural user interactions with minimal interruptions is crucial for analyzing user experiences in mobile apps. **Journal of Mobile Technology in Medicine**, v. 4, n. 2, p. 25-32, 2015.

HWANG, J.; LEE, C. Understanding user behavior within virtual spaces. **Computers in Human Behavior**, v. 76, p. 376-388, 2017.

JIANG, Z.; YIN, H.; LUO, Y.; GONG, J.; YANG, Y.; YANG, M. Quantitative Analysis of Mobile Application User Interface Design. In: IEEE 38th **International Performance Computing and Communications Conference** (IPCCC), London, UK, 2019, p. 1-8, 2019.

JOE, H., MAYDEU-OLIVARES, A. A general family of limited information goodness-of-fit statistics for multinomial data. **Psychometrika**, v.75, p.393-419, 2010.

JONKER, J.; PENNINK, B. **The essence of research methodology**. Heidelberg: Springer, 2010.

KAISER, H.F. The varimax criteria for analytical rotation in factor analysis. **Psychometrica**, v.23 p.141-51, 1958.

KHAJOUEI, R.; GOHARI, S. H.; MIRZAEE, M. Comparison of two heuristic evaluation methods for evaluating the usability of health information systems. **Journal of biomedical informatics**, v. 80, p. 37-42, 2018.

KIM, H. Effective organization of design guidelines reflecting designer's design strategies. **International Journal of Industrial Ergonomics**, v. 40, p. 669–688, 2010.

KIM, Y. J.; HAN, J. Y. Why smartphone advertising attracts customers: A model of Web advertising, flow, and personalization. **Computers in Human Behavior**, v. 33, p. 256–269, 2014.

KING, A. B. **Speed Up Your Site**: Web Site Optimization. 1. New Riders Publishing, 2003.

KING, A. J.; BOL, N.; CUMMINS, R. G.; JOHN, K. K. Improving visual behavior research in communication science: An overview, review, and reporting recommendations for using eye-tracking methods. **Communication Methods & Measures**, v. 13, n. 3, 2019.

KOCH, S.; KERSCHBAUM, M. Joining a smartphone ecosystem: Application developers' motivations and decision criteria. **Inf. Softw. Technol**., v. 56, p. 1423-1435, 2014.

KOUFARIS, M. Applying the Technology Acceptance Model and Flow Theory to Online Consumer Behavior. **Information Systems Research**, Vol. 13, No. 2, June 2002.

LADEIRA, W. J.; et al. Atenção visual influencia significativamente as avaliações de consumidores sobre produtos e serviços. **Journal of Retailing and Consumer Services**, v. 51, p. 297-305, 2019.

LI, X.; ZHAO, L.; LEE, S. H. Measuring user experiences with e-sourcing platforms: the development of the e-sourcing experience index. **Journal of Fashion Marketing and Management: An International Journal**, p. 1361-2026, 2020.

LOPEZ, M. F. B.; VIRTO, N. R.; SAN-MARTIN, S. Local Food Shopping: Factors Affecting Users' Behavioural E-Loyalty. **Adm. Sci**., v. 8, n. 47, 2018.

LORD, F. M. A theory of test scores. Psychometric Monograph No. 7, **Psychometric Society**, 1952.

LORD, F. M.; NOVICK, M. R. **Statistical Theories of Mental Test Score**. Reading: Addison-Wesley, 1968.

LUCAS, G. A.; LUNARDI, G. L.; DOLCI, D. B. From e-commerce to m-commerce: An analysis of the user's experience with different access platforms. **Electronic Commerce Research and Applications**, v. 58, 2023.

MAES, A.; VAN GEEL, A.; COZIJN, R. Signposts on the digital highway: The effect of semantic and pragmatic hyperlink previews. **Interacting with Computers**, v. 18, n. 2, p. 265–282, mar. 2006.

MAHNKE, R.; BENLIAN, A.; HESS, T. A Grounded Theory of Online Shopping Flow. **International Journal of Electronic Commerce**, v. 19, n. 3, pp. 54–89, 2015.

MASLOWSKA, E.; et al. Processamento de avaliações online. **Computers in Human Behavior**, v. 102, p. 14-24, 2020.

MAY, T. **Social Research:** Issues, methods, and process. 4th. ed. Berkshire, England: Open University Press. 354p.

MCDONALD, R. P. Describing the elephant: Structure and function in multivariate data. **Psychometrika**, v. 51, n. 4, p. 513-534, 1986.

MCLEAN, G.; OSEI-FRIMPONG, K.; AL-NABHANI, K.; MARRIOTT, H. Examining consumer attitudes towards retailers' m-commerce mobile applications – An initial adoption vs. continuous use perspective. **Journal of Business Research**, v. 106, p. 139-157, 2020.

MENDES-DA-SILVA, W. Contributions and limitations of narrative reviews and systematic reviews in the business area. **Revista de Administração Contemporânea,** v. 23, n.2, 2019.

MITSUTAKE, S.; AI SHIBATA; ISHII, K.; MIYAWAKI, R.; OKA, K. Associations of eHealth Literacy with Obtaining Knowledge about Colorectal Cancer among

Internet Users Accessing a Reputable Cancer Website: Internet-Based Survey Study. **Int. J. Environ. Res. Public Health,** v. 17, 2020.

MOHER, D.; LIBERATI, A.; TETSLAFF, J.; ALTMAN, D. G. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. **PLoS Med**, v. 6, n. 7, pp. 1-6, 2009.

MOLINARI, L. **Mobile application testing**: Quality and development in mobile applications. 1 ed. São Paulo: Saraiva, 2017. 296p.

MOREIRA JR., J.; TEZZA, R.; ANDRADE, D. F.; BORNIA, A. C. Algorithm of a computerized adaptive test based on the theory of response to the item for estimating the usability of e-commerce sites. **Production**, v. 23, n. 3, p. 525-536, 2013.

MORGAN, G. Paradigms, Metaphors, and Puzzle Solving in Organization Theory. **Administrative Science Quarterly**, v. 25, n. 4, p. 605-622, 1980.

NEIL, T. **Mobile Design Pattern Gallery**: UI Patterns for Smartphone Apps. Sebastpol: O'Reilly Media, 2014.

NGAI, E.W.T.; LI, C.L.; CHENG, T.C.E.; LUN, Y. H. V.; LAI, K.H.; CAO, J., M. C. M. Design, and development of an intelligent context-aware decision support system for realtime monitoring of container terminal operations. **International Journal of Production Research**, v. 49, n. 12, p. 3501-3526, 2011.

NIELSEN, J.; BUDIU, R. **Mobile Usability**. 1 ed. Rio de Janeiro: Elsevier, 2014. 203p.

NIQ E-BIT. **WebShoppers**. 49. ed. 34p. 2024. Available in: <https://nielseniq.com/global/pt/landing-page/ebit/nielseniq-ebit-brasil/webshoppers/>. Access: 06 Jun. 2024.

NOVAK, T.P.; HOFFMAN, D. L. Measuring the flow experience among Web users. **Interval Research Corporation**, Project 2000, Vanderbilt University, 1997.

NOVAK, T.P.; HOFFMAN, D.L.; YUNG, Y. F. Measuring the customer experience in online environments: A structural equation approach'. **Marketing Science**, v. 19, n. 1, p. 22-42, 2000.

NUNNALLY, J. C.; BERNSTEIN, I. R. **Psychometric theory**. 3 ed. McGraw-Hill, 1994.

ONTSI - Observatorio Nacional de Tecnología y Sociedad. **Compras Online en España:** Edición 2023 - Datos 2022. 164p. 2023. Available in: <https://www.ontsi.es/es/publicaciones/compras-online-espana-2021>. Access: 06 Jun. 2024.

OZKARAA, B.Y.; OZMENA, M.; KIM, J. W. Examining the effect of flow experience on online purchase: A novel approach to the flow theory based on hedonic and utilitarian value. **Journal of Retailing and Consumer Services**, v. 37, p. 119–131, 2017.

PAPPAS, I. O.; et al. Avaliação do apelo visual em ambientes de comércio eletrônico. **Journal of Business Research**, v. 85, p. 348-360, 2018.

PARÉ, G.; TRUDEL, M.-C.; JAANA, M.; KITSIOU, S. Synthesizing information systems knowledge: A typology of literature reviews. **Information & Management**, v. 52, n. 2, p. 225-239, 2015.

PASQUALI, L. **Psicometria**: teoria e aplicações. Brasília: Editora Universidade de Brasília, 1997.

PASQUALI, L. **Psychometrics**: theory of tests in psychology and education. Petropolis, RJ: Voices, 2003.

PEIFER, C.; ENGESER, S. (orgs.). **Advances in flow research**. Cham: Springer, 2021.

PELET, J. E.; ETTIS, S.; COWART, K. Optimal experience of flow enhanced by telepresence: evidence from social media use. **Information & Management**, v. 54, n. 1, pp.115–128, 2017.

PENG, L.; CUI, G.; CHUNG, Y.; LI, C. A multi-facet item response theory approach to improve customer satisfaction using online product ratings. **Journal of the Academy of Marketing Science**, v. 47, p. 960–976 2019.

PODSAKOFF, P. M., AHEARNE, M., MACKENZIE, S. B. Organizational citizenship behaviors and the quantity and quality of work group performance. **Journal of Applied Psychology**, v.82, p.262-270, 1997.

PODSAKOFF, P. M., AHEARNE, M., MACKENZIE, S. B. Organizational citizenship behaviors and the quantity and quality of work group performance. **Journal of Applied Psychology**, v.82, p.262-270, 1997.

PORTIGAL, S. **Interviewing users**: How to uncover compelling insights. O'Reilly Media, 2023.

PUNCHOOJIT, L.; HONGWARITTORRN, N. Usability Studies on Mobile User Interface Design Patterns: A Systematic Literature Review. **Advances in Human-Computer Interaction**, v. 2017, n. 1, 2017.

QUIVY, R.; CAMPENHOUDT, L. V. **Manual de Investigação em Ciências Sociais**. 2. Ed. Lisbon: Gradiva, 1998. 290p.

R Core Team - 2021. **A** language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at http://www.R-project.org/.

RAJI, M. A.; OLODO, H. B.; OKE, T. T.; ADDY, W. A.; OFODILE, O. C.; OYEWOLE, A. T. E-commerce and consumer behavior: A review of AI-powered personalization and market trends. **GSC Advanced Research and Reviews**, v. 18, n. 03, p. 066–077, 2024.

RAHMAN, T.; KIM, Y.S.; NOH, M.; LEE, C. K. Determinants of social commerce adoption in an emerging economy**. Service Business**, v. 14, p. 479–502, 2020.

RECKASE, M. D. **Multidimensional Item Response Theory**. East Lansing, MI: Springer, 2009. 353p.

RETTIE, R., An Exploration of Flow during Internet use. **Internet Research**, v. 11, n. 2, p. 103-113, 2001.

RIBEIRO JUNIOR, J. **What is positivism**? 2 ed. São Paulo: Brasiliense, 1996. 40 p.

RODRIGUEZ-ARDURA, I.; MESEGUER-ARTOLA, A. A PLS-neural network analysis of motivational orientations leading to Facebook engagement and the moderating roles of flow and age. **Frontiers in Psychology**, v. 11, n. 1869, 2020.

RODRIGUEZ-ARDURA, I.; MESEGUER-ARTOLA, A. E-Learning continuance: the impact of interactivity and the mediating role of imagery, presence, and flow. **Information & Management**, v. 53, n. 4, p. 504-516, 2016.

RODRIGUEZ-ARDURA, I.; MESEGUER-ARTOLA, A. Flow in e-learning: What drives it and why it matters. **British Journal of Educational Technology**, v. 48, n. 4, p. 899-915, 2017.

SAFFER, D. **Microinteractions**: Designing with Details. Sebastpol: O'Reilly Media, 2014.

SAFIEDDINE, M-commerce F. In: SEIFEDINE, K.; ALHAMI, A**. Innovations in E-Systems for Business and Commerce**. Waretown, USA: Apple Academic Press, 2016.

SALLEH, N.A.M.; HA, N.C.; KITCHEN, P. J. MOHD-ANY, A. A. Online purchasing: the role of web experience factors. **Int. J. Electronic Marketing and Retailing**, v. 10, n. 3, 2019.

SAMEJIMA, F. Estimation of latent ability using a response pattern of graded scores. **Psychometrika,** v. 34, n. S1, p. 1-97, 1969.

SCHNEIDER, W.A. **An analysis of users' behavior in relation to the sharing of private information with m-commerce providers.** 2019. 297p. (Doctoral thesis). Center for Administration and Socioeconomic Sciences, State University of Santa Catarina, Florianópolis, SC.

SECKLER, M.; OPWIS, K.; TUCH, A. N. Correlation with gaze behavior, which could empirically inform app design improvements. **International Journal of Human-Computer Studies**, v. 131, p. 13-24.

SHIH, H.; JIN, B. Driving Goal-Directed and Experiential Online Shopping**. Journal of Organizational Computing and Electronic Commerce**, v. 21, n. 2, p. 136-157, 2011.

SHITKOVA, M.; HOLLER, J.; HEIDE, T.; CLEVER, N.; BECKER, J. Towards Usability Guidelines for Mobile Websites and Applications. **Proceedings** […] Wirtschaftsinformatik Proceedings 2015, THOMAS, O.; TEUTEBERG, F. (Eds.), p. 107, 2015.

SIJTSMA, K. On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha. **Psychometrika**, v. 74, n. 1, p. 107-120, 2009.

SINA, A.S.; WU, J. Effects of 3D vs 2D interfaces and product-coordination methods. **International Journal of Retail & Distribution Management**, v. 47, n. 8, p. 855-871, 2019.

SINGH, J. Tackling measurement problems with Item Response Theory: Principles., characteristics, and assessment, with an illustrative example. **Journal of Business Research**, v. 57, p. 184-208, 2004.

SOUZA, M. W. D. **Measurement of perceived risk in the use of m-banking in the youth market, using the Item Response Theory**. 2016. 139p. (master's thesis). Center for Administration and Socioeconomic Sciences, State University of Santa Catarina, Florianópolis, SC.

STEINBERG, L., THISSEN, D.; WAINER, H. Validity. In H. Wainer (Ed), **Computerized Adaptive Testing**: A Primer. Mahwah, Jew Jersey: Lawrence Erlbaum. 2000.

STEPCHENKOVA S., et al. Benchmarking CVB website performance: Spatial and structural patterns, **Tourism Management,** v.31, p.611–620, 2010.

STOP, G.; TRUDEL, M.; JAAN, M.; KITSIOU, S. Synthesizing information systems knowledge: a typology of literature reviews. **Information & Management**, v. 52, n. 2, pp. 183-199, 2015.

SUPRIADI, O. A. User Interface Design of Mobile-based Commerce**. IOP Conference Series: Materials Science and Engineering**, v. 662, n. 2, p. 022047, 2019.

SYMPSON, J. B. A model for testing with multidimensional items. In: COMPUTERIZED ADAPTIVE TESTING CONFERENCE, 1977. **Proceedings** […] Proceedings of 1977 Computerized adaptive testing conference, WEISS, D. J. (Ed.), p. 82-98, 1978.

TABACHNICK, B. G.; FIDELL, L. S. **Using Multivariate Statistics.** 6 ed. Pearson Education, 2013.

TAVARES, H. R.; ANDRADE, D. F.; PEREIRA, C. A. Detection of determinant genes and diagnostic via item response theory. **Genetics and Molecular Biology**, v. 27, n. 4, p. 679-685, 2004.

TEZZA, R.; BORNIA, A.C.; ANDRADE, D. F. Measuring web usability using item response theory: Principles, features, and opportunities. **Interacting with Computers**, v. 23, p. 167–175, 2011.

TEZZA, R.; BORNIA, B.C.; ANDRADE; D.F.; BARBETTA, P. A. Multidimensional model to measure quality in e-commerce websites using item response theory. **Gest. Prod**., v. 25, n. 4, p. 916-934, 2018.

TEZZA, R.; BORNIA, A. C.; SPENASSATO, D. AND TRIERWEILLER, A. C. Empirical comparison of the multidimensional models of item response theory in e-commerce. **Pesquisa Operacional**, v. 36, n. 3, p. 503-532, 2016.

THANGAVEL, P.; CHANDRA, B. Two decades of m-commerce consumer research: A bibliometric analysis using R Biblioshiny. **Sustainability**, v. 15, n. 15, p. 11835, 2023.

TONKIN, E.; et al. Utilização de eye-tracking para investigar diferenças comportamentais em telas de diferentes tamanhos. **Journal of Eye Tracking, Visual Cognition and Emotion**, v. 3, n. 2, p. 45-56, 2011.

TRIBERTI, S.; DI NATALE, A. F.; GAGGIOLI, A. Flowing technologies: The role of flow and related constructs in human-computer interaction. In: PEIFER, C.; ENGESER, S. (orgs.). **Advances in flow research**. Cham: Springer, 2021.

TRIERWEILLER, A.C. et al. Measuring environmental management disclosure in industries in Brazil with Item Response Theory. **Journal of Cleaner Production,** v. 47, p. 298-305, 2013.

VAN DER VELDEN, M.; EMAM, K. E. “Not all my friends need to know”: a qualitative study of teenage patients, privacy, and social media. **Journal of the American Medical Informatics Association**, v. 20, n. 1, p. 16–24, 2013.

VENKATESH, V.; MORRIS, M. Why Don’t Men Ever Stop to Ask For Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behavior. **MIS Quarterly**, v. 24, n.1, p. 115-139, 2000.

VIDYAPU, S.; et al. Captura da atenção do usuário para elementos de interface e imagens em plataformas de comércio eletrônico. **Journal of Interactive Marketing**, v. 45, p. 27-41, 2019.

WANG, J. et al. Through the wall detection and localization of autonomous mobile device in indoor scenario. **IEEE Journal on Selected Areas in Communications**, v. 42, n. 1, p. 161-176, 2024.

WANG, L. et al. Impacto dos backgrounds de imagens de produtos na atenção do consumidor. **Journal of Consumer Research**, v. 45, n. 5, p. 1023-1040, 2019.

WEI, T. T., MARTHANDAN, G., CHONG, A. Y. L., OOI, K. B.; ARUMUGAM, S. What drives Malaysian m-commerce adoption? an empirical analysis. **Industrial Management & Data Systems**, v. 109, n.3, p. 370-388, 2009.

WILSON, M. **Constructing measures**: An item response modeling approach. New Jersey: Lawrence Erlbaum Associates, 2005. 248p.

WILSON, M., HOSKENS, M. **Multidimensional item responses**: Multimethodmultitrait perspectives. In S. Alagumalai et al. (Eds.), Applied Rasch measurement: A book of exemplars.(pp. 287-307). New York: Springer, 2005.

WONGTADA, N., RICE, G. Multidimensional latent traits of perceived organizational innovation: Differences between Thai and Egyptian employees. **Asia Pacific Journal of Management**, v.25, p.537–562, 2008.

WU, I.; CHIU, M.; CHEN, K. Defining the determinants of online impulse buying through a shopping process of integrating perceived risk, expectation-confirmation model, and flow theory issues. **International Journal of Information Management**, v. 52, 2020.

XIE, S.; YUAN, Y. An Empirical Study on the Impact of Online Travel Consumers' Brand Loyalty: The Mediating Effect of Flow Experience. In: INTERNATIONAL CONFERENCE ON ENVIRONMENTAL AND ENGINEERING MANAGEMENT, 253., 2021, Changsha. **Proceedings**[...]. Changsha, E3S Web Conf., 2021.

XU, H. et al. The personalization privacy paradox: An exploratory study of decision-making process for location-aware marketing. **Decision Support Systems,** v. 51, n. 1, p. 42-52, 2011.

ZARIMAN, N. F. M.; HUMAIDI, N.; ABD RASHID, M. H. Mobile commerce applications service quality in enhancing customer loyalty intention: mediating role of customer satisfaction. **Journal of Financial Services Marketing**, v. 28, p. 649–663, 2023.

ZILLICH, I.; KESSLER, S. Measuring selective exposure to online information: combining eye-tracking and content analysis of users' actual search behavior. In: PETER, C; NAAB, T; KÜHNE, R. **Measuring Media Use and Exposure**. Köln: Halem, 196-220, 2019.

ZIMANYI, Z., SCHÜLER, J. The Dark Side of the Moon. In: PEIFER, C., ENGESER, S. (eds) **Advances in Flow Research**. Springer, Cham, 2021.

# APPENDIX A – SYSTEMATIC SEARCH

A systematic literature review was chosen to be conducted due to its structured approach in selecting papers based on explicit and replicable methods of search, selection, and analysis (MENDES-DA-SILVA, 2019). Among various types of literature reviews, theoretical reviews share similarities with this study, particularly in terms of their comprehensive strategy and broad scope (PARÉ et al., 2015).

In the context of factors influencing the online flow conditions concept, no systematic reviews covering the past decade were identified. To synthesize the existing literature and contribute to future research agendas, this study adopted aspects recommended by the PRISMA model (MOHER et al., 2009), focusing on eligibility criteria, analysis methods, and communication standards applicable to systematic reviews.

This systematic review adhered to the criteria set forth by Moher et al. (2009), including eligibility criteria, information sources, search strategy, study selection, data collection procedures, assessment of the risk of bias, summary measures, synthesis of results, assessment of bias across studies, and supplementary analyses.

The inclusion criteria for studies were defined as follows:

a) Full articles published in Portuguese or English.

b) Articles published in peer-reviewed journals or presented at scientific conferences, without date restrictions, utilizing comprehensive databases.

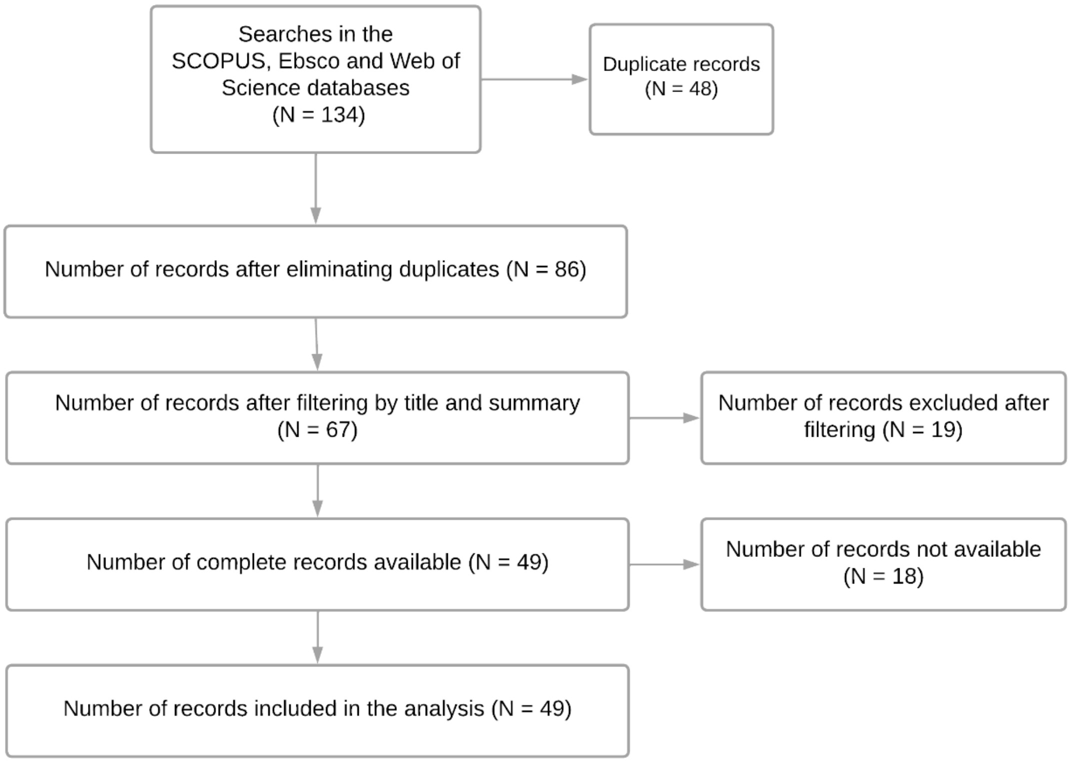
c) Databases and information sources included SCOPUS (complete collection), EBSCO (complete collection), and Web of Science (complete collection, including SCIELO).

d) Articles applying Flow Theory in an online context.

The search was conducted in August 2020, focusing on search terms in titles, abstracts, and keywords. Following the definition of inclusion criteria, the next step involved devising an appropriate search strategy using precise terms to construct the search query.

The search strategy was formulated as follows: “flow theory” AND (”online purchase” OR “online shopping” OR “electronic commerce” OR “mobile commerce” OR “*-commerce”). Figure 3 illustrates the flow of the selection process, adhering to the PRISMA model by Moher et al. (2009). Utilizing Mendeley software, this approach yielded a total of 134 articles. After removing 48 duplicate studies, 86 unique studies remained for analysis in the databases reviewed.

Figure 11 – Records selection process




Source: Prepared by the author based on Moher et al. (2009).

Following the initial search, a meticulous filtering process was conducted based on the relevance of article titles and abstracts to the theme. This led to the selection of 67 articles that specifically addressed Flow Theory within the context of online environments. Nineteen articles were excluded for reasons such as not focusing on Flow Theory in online shopping contexts, addressing Flow Theory in unrelated contexts (e.g., online gaming, autonomous vehicles, instant messaging), not treating Flow as a variable in the analyzed model, or discussing unrelated flow theories (e.g., Two-Step Flow Theory, Carbon Emission Flow Theory). Additionally, 18 out of the remaining 67 articles were inaccessible for reading due to privacy or subscription constraints. Consequently, the final sample for this systematic review comprised 49 articles.

To synthesize and summarize the findings, an analysis was conducted on the distribution of publications across journals over the years, the affiliations of researchers by country, citations received by prominent authors, and the frequency of keywords used. Together, these selected articles provide comprehensive insights into the application of Flow Theory in online contexts, offering a nuanced understanding of the topic.

The systematic review identified and analyzed 49 articles that explored Flow Theory in various online scenarios. Despite challenges such as inaccessible articles and diverse applications of Flow Theory, the selected literature contributes significantly to understanding how Flow Theory is applied in the context of online activities. The findings emphasize the importance of Flow Theory in enhancing user experiences and engagement in digital environments, suggesting avenues for future research to further elucidate its implications across different domains.

All the articles in the sample assess online flow from the user's perspective, focusing on behavior and experiences during the use of digital platforms. However, none of the studies directly evaluate the virtual environments themselves, instead concentrating predominantly on user perceptions and interactions. This gap in the literature highlights the need for further research into how the characteristics and design of virtual environments might influence flow and user experience. The absence of studies directly addressing virtual environments indicates an opportunity to explore how various elements of these environments impact the state of flow, providing valuable insights for the development and optimization of digital platforms.

# APPENDIX B – RESEARCH APP SAMPLES

Chart 8 – Sample of 100 shopping apps available in Brazil

| | | |
|---|---|---|
| Mercado Livre | Gazin: Compras Online | FARFETCH - Compre moda de luxo |
| Shopee | Sam's Club Brasil | Compras Paraguai |
| SHEIN | FARM | Superclube Supermarket |
| O Boticário | Cobasi: pet shop online | Mateus Mais |
| Magalu | Pão de Açúcar Mais | Amobeleza |
| Amazon Shopping | Petlove: petshop e saúde pet | Kalunga |
| AliExpress Shopping App | Drogarias Pacheco | CASA&VÍDEO |
| OLX - Compra e venda online | Petz: Tudo que seu pet precisa | AREZZO |
| C&A Loja Online: Comprar Roupa | ZARA | CSSBuy |
| Nike | Reserva | Prezunic |
| Lojas Renner - Comprar Roupas | Melissa Oficial | Anacapri |
| Vale Bonus | Pichau | Farmácia Preço Popular |
| Netshoes: Loja de Esportes | Privalia: Moda, Decoração e + | Pernambucanas |
| Casas Bahia: Comprar Online | KaBuM! Compras Online | Temu:Compre como um bilionário |
| App comercial Alibaba.com | Leroy Merlin - Brasil | Zattini: ofertas de roupas |
| Beleza na Web - Cosméticos | Evino | Tenda Atacado |
| Méliuz: Cashback, Cartão e + | Chilli Beans | Banggood Global Online Shop |
| MadeiraMadeira: compras online | Multi: Promoções dos Shoppings | Ponto: Comprar Produtos Online |
| Americanas: Compras Online | Decathlon | Clube Extra |
| enjoei: comprar e vender roupa | Submarino: Compras Online | Youcom: comprar roupas no app |
| Centauro: loja de esportes | Vans | Camicado |
| Riachuelo – Comprar roupas | Tok&Stok: Móveis e Decoração | Loja do Mecânico |
| SEPHORA | Lojas Colombo | Moda Center |
| O Boticário Revendedor | Hering - O Básico do Brasil | eBay: Buy & Sell Marketplace |
| Buscapé | Prime Gourmet 5.0 | Apple Store |
| OFF Premium | Pechinchou Promoções e Ofertas | Lojas Torra: Comprar Roupas |
| Shultz | Elo7 | glam: beleza em primeiro lugar |
| HAVAN | Farmácias São João | Zee.Now - Pet Shop Delivery |
| Drogaria São Paulo | Marisa | JOHN JOHN |
| Época Cosméticos: Maquiagens | market4u | St Marche |
| Meu Carrefour | Max Atacadista | Maria Filó |
| Dafiti | Crocs | ZZ MALL |
| Natura: perfumes e cosméticos | Avon APP | Ferreira Costa |
| Droper | | |

Source: Prepared by the author, 2024.

Chart 9 – Sample of 100 shopping apps available in Spain

| | | |
|---|---|---|
| Temu: Compra como millonario | CIDER - Ropa y moda | Massimo Dutti: Moda online |
| SHEIN - Compras Online | JD Sports | Fashion Nova |
| Miravia: App de compras online | Tezenis | Privalia - Outlet de marcas |
| Vinted: Ropa de segunda mano | ManoMano | OYSHO: Tienda De Moda Online |
| Amazon | Costco Wholesale Spain | Saigu Cosmetics |
| Wallapop - Compra y vende | Lefties - Moda Online | Rituals Cuerpo y hogar |
| AliExpress Shopping App | Lidl - Tienda online - Ofertas | Alcampo tienda online |
| Vestiaire Collective | ASOS: la mejor moda en línea | Lyst: Shop Fashion Brands |
| Milanuncios: Segunda mano | Private Sport Shop - Outlet | ABOUT YOU Fashion Online Shop |
| Supermercados DIA | AUTODOC — Recambios de Coche | BROWNIE - Moda online |
| Hacoo - sara lower price mart | Club Alcampo | Cortefiel |
| Mercadona | Mundo Consum - Compra online | Naturitas: Salud Natural |
| LEROY MERLIN | Etsy: manualidades y regalos | Action |
| Zalando: Tienda De Moda Online | eBay: Comprar y Vender de Todo | Myprotein: Fitness y Compras |
| Back Market, compra online | DHgate tienda online mayorista | C&A Tienda online de moda |
| PandaBuy | Mi Store | KIKO MILANO - Makeup & beauty |
| ZARA | Zara Home | Foot Locker |
| El Corte Inglés | Springfield | BM Supermercados |
| Mi Carrefour | Alvaro Moreno | Kiwoko - Todo para tu mascota |
| Privé by Zalando | StockX: sneakers y ropa | GLOWRIAS |
| IKEA | TOUS | Veepee - Outlet Online |
| Prendas y zapatillas Nike | Women'secret | Hollister So Cal Style |
| H&M - nos encanta la moda | La Sirena Congelados | Levi's |
| App de comercio B2B Alibaba | ALDI Supermercados | Catawiki - Subastas Online |
| Decathlon: Tienda de deporte | PcComponentes | Casa del Libro |
| Fnac: Compras Online | UNIQLO ES | HagoBuy |
| PULL&BEAR | SCALPERS | Showroomprive - Venta privada |
| Perfumerías DRUNI | Parfois | SUGARGOO "From China to Global" |
| Bershka | SNIPES - sneaker & streetwear | Fulanita |
| Sephora - Maquillaje, Belleza | Sprinter | Instant Gaming |
| Stradivarius - Ropa de mujer | Apple Store | Forum Sport |
| MANGO - Online fashion | FARFETCH - Compra moda de lujo | adidas CONFIRMED |
| Perfumerías Primor | idealo - App de compras online | La Roca Village |
| KIABI l'app mode à petits prix | | |

Source: Prepared by the author, 2024.

## APPENDIX C – ITEM EVALUATION FORM AND REFERENCES

Dear Professor,

I am currently engaged in a study under the guidance of my two thesis advisors, Professor Rafael Tezza (UDESC/Brazil) and Professor Antoni Meseguer-Artola (UOC/Spain), aimed at developing a scale to assess the degree of Online Flow Conditions in shopping applications. This scale will subsequently be applied to a sample of 200 applications.

The study draws upon Csikszentmihalyi's Flow Theory, originally formulated in the 1970s, which explores how individuals can deeply engage with tasks and achieve a state of optimal experience. In this state, individuals leverage their skills to overcome challenges that match their abilities. With the proliferation of the Internet, this theory has been adapted to online contexts, leading to the concept of online flow, which is the focal point of our investigation.

Online Flow refers to a cognitive state experienced during internet browsing where individuals become fully immersed in their online activities, experiencing a deep sense of engagement that captivates their attention. Moreover, in virtual environments, the flow state alters individuals' perception of time, leading them to perceive time passing more swiftly.

Our study identifies three primary domains of Online Flow: (1) Online Flow Conditions, (2) Online Flow Characteristics, and (3) Online Flow Consequences. Specifically, our research concentrates on Area 1 (Online Flow Conditions), which encompasses inherent elements within the application interface that facilitate the emergence of flow states in users. This assessment is conducted directly by researchers within the application interface, rather than relying on users' self-reported perceptions of experiencing flow.

Subsequently, based on the literature, Area 1 is further delineated into three elements: (1) Actions requiring skill, (2) Clear goals, and (3) Immediate feedback. These elements serve as the basis for defining items that may form part of the measurement scale.

At this time, we are in the process of submitting these items for evaluation by expert judges, and we kindly seek your collaboration in this regard. The items are formulated in a dichotomous manner, indicating whether each element is present or absent in the observed shopping application. We request that you assess each item based on its Clarity (clarity of wording without ambiguity) and Relevance (pertinence to the element being analyzed), assigning a score ranging from 1 to 3: 1 (low), 2 (moderate), or 3 (high). Additionally, if necessary, there is an opportunity to provide comments alongside each item. Suggestions and general recommendations are also greatly appreciated and can be provided after the questionnaire.

Below are the items presented for evaluation.

Thank you in advance for your collaboration!

| **Online Flow Conditions dimensions and items** | | **CLARITY** | | | **RELEVANCE** | | | **References** |
|---|---|---|---|---|---|---|---|---|
| | | Low (1) | Moderate (2) | High (3) | Low (1) | Moderate (2) | High (3) | |
| Demand-Skill Balance | | | | | | | | |
| 1 | The application has onboarding, an initial step-by-step that indicates how its use works. | | | | | | | King (2003); Rettie (2001); Guo and Poole (2009); Higgins (2020) |
| 2 | The application features an adjustable interface, enabling users to tailor it to their preferences. | | | | | | | King (2003); Guo and Poole (2009) |
| 3 | The application allows its use without the need for user registration. | | | | | | | King (2003); Guo and Poole (2009) |
| 4 | Users can access a list that compiles the history of viewed items within the application. | | | | | | | Choi, Kirshner and Wu (2016); Supriadi (2019) |
| 5 | The application shows personalized product recommendations based on the user's history. | | | | | | | Kim, (2010); Mahnke, Benlian and Hess (2015); Ayada and Hammad (2023) |
| 6 | Users can select their preferred method for receiving notifications from the application, such as via email, SMS, WhatsApp, or other available channels. | | | | | | | Ayada and Hammad (2023) |
| 7 | The application does not display a loading icon during use. | | | | | | | Ayada and Hammad (2023) |
| 8 | It is possible to use zoom to enlarge product images. | | | | | | | Punchoojit and Hongwarittorrn (2017) |
| 9 | The application has a horizontal layout and vertical scrolling. | | | | | | | Chan et al. (2002); Jiang et al. (2019) |
| 10 | The application has infinitely scrollable product sections. | | | | | | | Mahnke, Benlian and Hess (2015) |
| 11 | The application enables modification of the user's profile, including inserting a personal | | | | | | | King (2003); Ayada and Hammad (2023) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | photo, creating a username, and other related features. | | | | | | | |
| 12 | The application facilitates customization of its components within the account settings section. | | | | | | | King (2003); Ayada and Hammad (2023) |
| 13 | There is a version of the application for iOS and Android. | | | | | | | Koch and Kerschbaum (2014) |
| Clear Goals | | | | | | | | |
| 14 | The application allows users to save items to user-created lists. | | | | | | | Choi, Kirshner and Wu (2016) |
| 15 | The application enables users to add items to a favorites list. | | | | | | | Choi, Kirshner and Wu (2016) |
| 16 | The application allows users to share lists of items with other users. | | | | | | | Choi, Kirshner and Wu (2016) |
| 17 | The application has breadcrumb trails that allow the user to identify which part of the application they are in, referring to a general structure. | | | | | | | King (2003); Shitkova et al. (2015) |
| 18 | The application provides a section where users can answer questions, such as a FAQ (Frequently Asked Questions) section or direct contact with customer service. | | | | | | | King (2003); Maes, Van Geel e Cozijn (2006) |
| 19 | The application features “you are here” style location landmarks, symbols, or highlights so that the user can easily form a mental model of the app. | | | | | | | King (2003) |
| 20 | The application has a navigation bar to access its main information, for example, user profile, main page, etc. | | | | | | | King (2003) |
| 21 | The application enables users to apply various filters to the lists of available products. | | | | | | | Mahnke, Benlian and Hess (2015) |
| 22 | The application’s homepage features announcements, such as special birthday promotions, Black Friday promotions, etc. | | | | | | | Mahnke, Benlian and Hess (2015) |
| 23 | After completing an action, the application displays motivational information, such as promotions and most purchased products, to encourage continued user engagement. | | | | | | | Mahnke, Benlian and Hess (2015) |
| 24 | The application includes a visible button for accessing the shopping cart. | | | | | | | Kim (2010) |
| 25 | Users can obtain the number of items in the shopping cart without needing to access it directly. | | | | | | | Saffer (2014); Neil (2014) |
| 26 | Users have the option to qualify for free shipping within the application. | | | | | | | Kim (2010); Ayada and Hammad (2023) |
| 27 | The application has a customer support section. | | | | | | | Ayada and Hammad (2023) |
| 28 | The application allows more than two payment methods | | | | | | | Wu; Chiu and Chen (2020); Ayada and Hammad (2023) |
| 29 | The application uses the same font, color scheme, and layout on different pages. | | | | | | | Mahnke, Benlian and Hess (2015); Ayada and Hammad (2023) |
| 30 | The application provides a visible button to go back to the previous page or state. | | | | | | | Supriadi (2019) |
| Immediate Feedback | | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 31 | The application provides feedback in the form of navigation widgets, such as menus, toolbars, etc. | | | | | | | King (2003) |
| 32 | The application provides feedback in the form of performance variables, such as loading server icons, memory usage amount, etc. | | | | | | | King (2003) |
| 33 | The application has a progress bar, for example, in the cases of displaying the loading of a page, how much is left to get free shipping, etc. | | | | | | | King (2003) |
| 34 | The application presents a confirmation message before executing specific actions, such as deleting items from the favorites list or shopping cart. | | | | | | | King (2003) |
| 35 | The application has visual differentiation between visited and unvisited sessions. | | | | | | | King (2003) |
| 36 | The application delivers feedback audibly. | | | | | | | Ayada and Hammad (2023) |
| 37 | The application provides feedback in the form of symbols that express information, such as emojis, numbers, etc. | | | | | | | Ayada and Hammad (2023) |
| 38 | The application provides feedback in the form of a banner notification. | | | | | | | Ayada and Hammad (2023) |
| 39 | When entering incorrect input data, the application informs you that there is an error. | | | | | | | Saffer (2014) |
| 40 | Some buttons react to user actions by changing nomenclature, such as “add to cart” and “add another to cart”, “buy now” and “buy another now”, etc. | | | | | | | Saffer (2014) |
| 41 | The application offers more than one method for product search, such as a search bar, sorting and filtering options, and voice commands, among others. | | | | | | | Mahnke, Benlian and Hess (2015) |
| 42 | The application has an automatic entry feature in the search bar, that is, as something is typed, possible product options are already provided. | | | | | | | Mahnke, Benlian and Hess (2015) |
| 43 | The application provides alternative methods of logging in, such as authentication via social media accounts like Facebook, Google, and others. | | | | | | | Kim (2010) |
| 44 | The application offers generic product recommendations on the homepage. | | | | | | | Kim, (2010); Mahnke, Benlian and Hess (2015); Ayada and Hammad (2023) |
| 45 | The application allows users to read and submit product reviews. | | | | | | | Mahnke, Benlian and Hess (2015) |
| 46 | The application allows users to submit images, audio, or videos of the products in reviews. | | | | | | | Mahnke, Benlian and Hess (2015) |
| 47 | The application sends messages of a personal nature to the user, such as welcome, good morning, good evening, etc. | | | | | | | Saffer (2014); Neil (2014) |

Source: Prepared by the author, 2024.

# APPENDIX D – EXPERTS' ASSESSMENT

Chart 10 – Experts' assessment

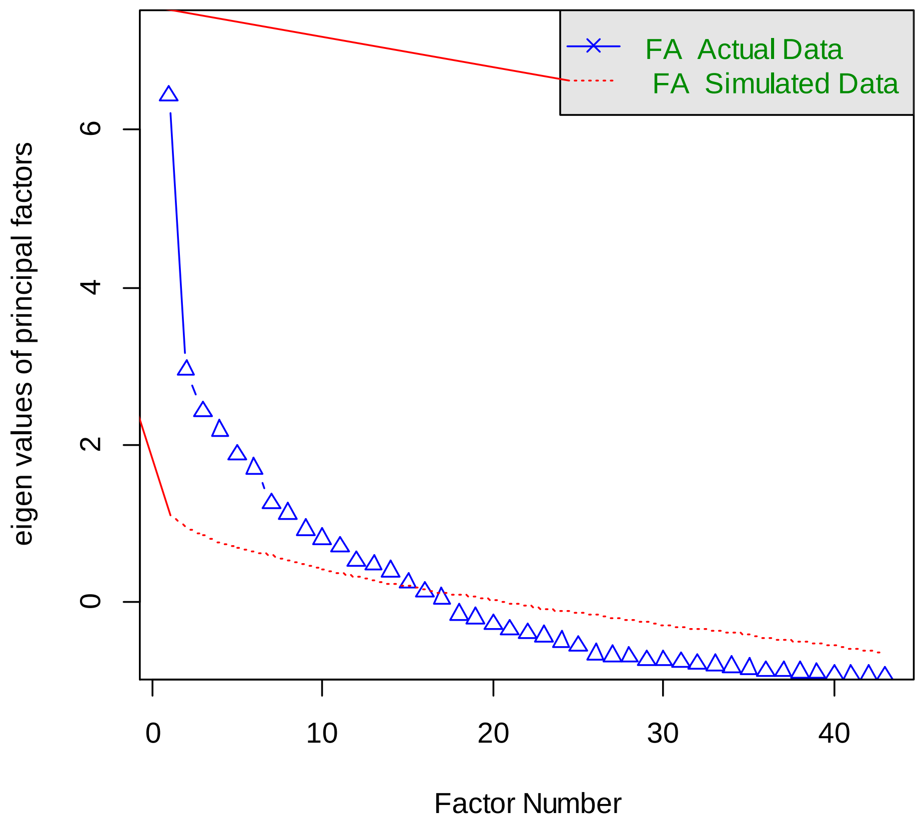
| Online Flow Conditions dimensions and items | | CLARITY | | | | | | RELEVANCE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Expert 1 | Expert 2 | Expert 3 | Expert 4 | Expert 5 | Median | Expert 1 | Expert 2 | Expert 3 | Expert 4 | Expert 5 | Median |
| **Demand-Skill Balance** | | | | | | | | | | | | | |
| 1 | The application has onboarding, an initial step-by-step that indicates how its use works. | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 2 | 3 | 1 | 2 |
| 2 | The application features an adjustable interface, enabling users to tailor it to their preferences. | 1 | 2 | 2 | 2 | 1 | 2 | 3 | 1 | 3 | 3 | 3 | 3 |
| 3 | The application allows its use without the need for user registration. | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 3 | 1 | 2 |
| 4 | Users can access a list that compiles the history of viewed items within the application. | 3 | 3 | 2 | 3 | 1 | 3 | 3 | 3 | 2 | 3 | 3 | 3 |
| 5 | The application shows personalized product recommendations based on the user's history. | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 |
| 6 | Users can select their preferred method for receiving notifications from the application, such as via email, SMS, WhatsApp, or other available channels. | 3 | 3 | 3 | 2 | 2 | 3 | 1 | 3 | 2 | 3 | 3 | 3 |
| 7 | The application does not display a loading icon during use. | 1 | 2 | 1 | 2 | 3 | 2 | 1 | 3 | 2 | 2 | 1 | 2 |
| 8 | It is possible to use zoom to enlarge product images. | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 3 | 3 |
| 9 | The application has a horizontal layout and vertical scrolling. | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 2 | 3 | 3 |
| 10 | The application has infinitely scrollable product sections. | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 2 | 3 | 3 |
| 11 | The application enables modification of the user's profile, including inserting a personal photo, creating a username, and other related features. | 2 | 3 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 |
| 12 | The application facilitates customization of its components within the account settings section. | 2 | 3 | 2 | 2 | 3 | 2 | 3 | 2 | 3 | 3 | 3 | 3 |
| 13 | There is a version of the application for iOS and Android. | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 1 | 3 |
| **Clear goals** | | | | | | | | | | | | | |
| 14 | The application allows users to save items to user-created lists. | 3 | 1 | 3 | 2 | 3 | 3 | 1 | 1 | 2 | 3 | 3 | 2 |
| 15 | The application enables users to add items to a favorites list. | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 16 | The application allows users to share lists of items with other users. | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 17 | The application has breadcrumb trails that allow the user to identify which part of the application they are in, referring to a general structure. | 1 | 3 | 1 | 1 | 3 | 1 | 2 | 3 | 2 | 3 | 3 | 3 |
| 18 | The application provides a section where users can answer questions, such as a FAQ (Frequently Asked Questions) section or direct contact with customer service. | 3 | 2 | 3 | 2 | 3 | 3 | 2 | 2 | 1 | 3 | 3 | 2 |
| 19 | The application features "you are here" style location landmarks, symbols, or highlights so that the user can easily form a mental model of the app. | 1 | 3 | 1 | 3 | 3 | 3 | 1 | 3 | 3 | 3 | 3 | 3 |
| 20 | The application has a navigation bar to access its main information, for example, user profile, main page, etc. | 3 | 2 | 2 | 1 | 3 | 2 | 2 | 2 | 2 | 3 | 3 | 2 |
| 21 | The application enables users to apply various filters to the lists of available products. | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 22 | The application's homepage features announcements, such as special birthday promotions, Black Friday promotions, etc. | 3 | 2 | 2 | 3 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 |
| 23 | After completing an action, the application displays motivational information, such as promotions and most purchased products, to encourage continued user engagement. | 3 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 3 |
| 24 | The application includes a visible button for accessing the shopping cart. | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 25 | Users can obtain the number of items in the shopping cart without needing to access it directly. | 3 | 3 | 1 | 1 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 |
| 26 | Users have the option to qualify for free shipping within the application. | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 |
| 27 | The application has a customer support section. | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 28 | The application allows more than two payment methods | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 |
| 29 | The application uses the same font, color scheme, and layout on different pages. | 3 | 3 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 3 |
| 30 | The application provides a visible button to go back to the previous page or state. | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 |
| **Immediate feedback** | | | | | | | | | | | | | |
| 31 | The application provides feedback in the form of navigation widgets, such as menus, toolbars, etc. | 1 | 2 | 1 | 2 | 1 | 1 | 3 | 2 | 3 | 3 | 3 | 3 |
| 32 | The application provides feedback in the form of performance variables, such as loading server icons, memory usage amount, etc. | 1 | 3 | 2 | 1 | 1 | 1 | 1 | 3 | 3 | 2 | 3 | 3 |
| 33 | The application has a progress bar, for example, in the cases of displaying the loading of a page, how much is left to get free shipping, etc. | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 | 3 |
| 34 | The application presents a confirmation message before executing specific actions, such as deleting items from the favorites list or shopping cart. | 2 | 2 | 3 | 3 | 3 | 3 | 2 | 2 | 3 | 3 | 3 | 3 |
| 35 | The application has visual differentiation between visited and unvisited sessions. | 1 | 2 | 1 | 1 | 3 | 1 | 2 | 2 | 2 | 2 | 3 | 2 |
| 36 | The application delivers feedback audibly. | 3 | 3 | 2 | 2 | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 |
| 37 | The application provides feedback in the form of symbols that express information, such as emojis, numbers, etc. | 3 | 3 | 2 | 2 | 1 | 2 | 3 | 3 | 2 | 3 | 3 | 3 |
| 38 | The application provides feedback in the form of a banner notification. | 3 | 2 | 2 | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 | 3 |
| 39 | When entering incorrect input data, the application informs you that there is an error. | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 40 | Some buttons react to user actions by changing nomenclature, such as "add to cart" and "add another to cart", "buy now" and "buy another now", etc. | 3 | 3 | 3 | 1 | 2 | 3 | 3 | 3 | 2 | 3 | 3 | 3 |
| 41 | The application offers more than one method for product search, such as a search bar, sorting and filtering options, and voice commands, among others. | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| 42 | The application has an automatic entry feature in the search bar, that is, as something is typed, possible product options are already provided. | 3 | 2 | 1 | 3 | 1 | 2 | 3 | 2 | 2 | 3 | 2 | 2 |
| 43 | The application provides alternative methods of logging in, such as authentication via social media accounts like Facebook, Google, and others. | 3 | 3 | 3 | 3 | 2 | 3 | 2 | 3 | 3 | 3 | 3 | 3 |
| 44 | The application offers generic product recommendations on the homepage. | 3 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 3 |
| 45 | The application allows users to read and submit product reviews. | 3 | 2 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 | 3 |
| 46 | The application allows users to submit images, audio, or videos of the products in reviews. | 3 | 2 | 3 | 3 | 3 | 3 | 3 | 2 | 3 | 3 | 3 | 3 |
| 47 | The application sends messages of a personal nature to the user, such as welcome, good morning, good evening, etc. | 3 | 3 | 1 | 1 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 |

Source: Prepared by the author, 2024.

**APPENDIX E – PARALLEL ANALYSIS (43-ITEM SET)**

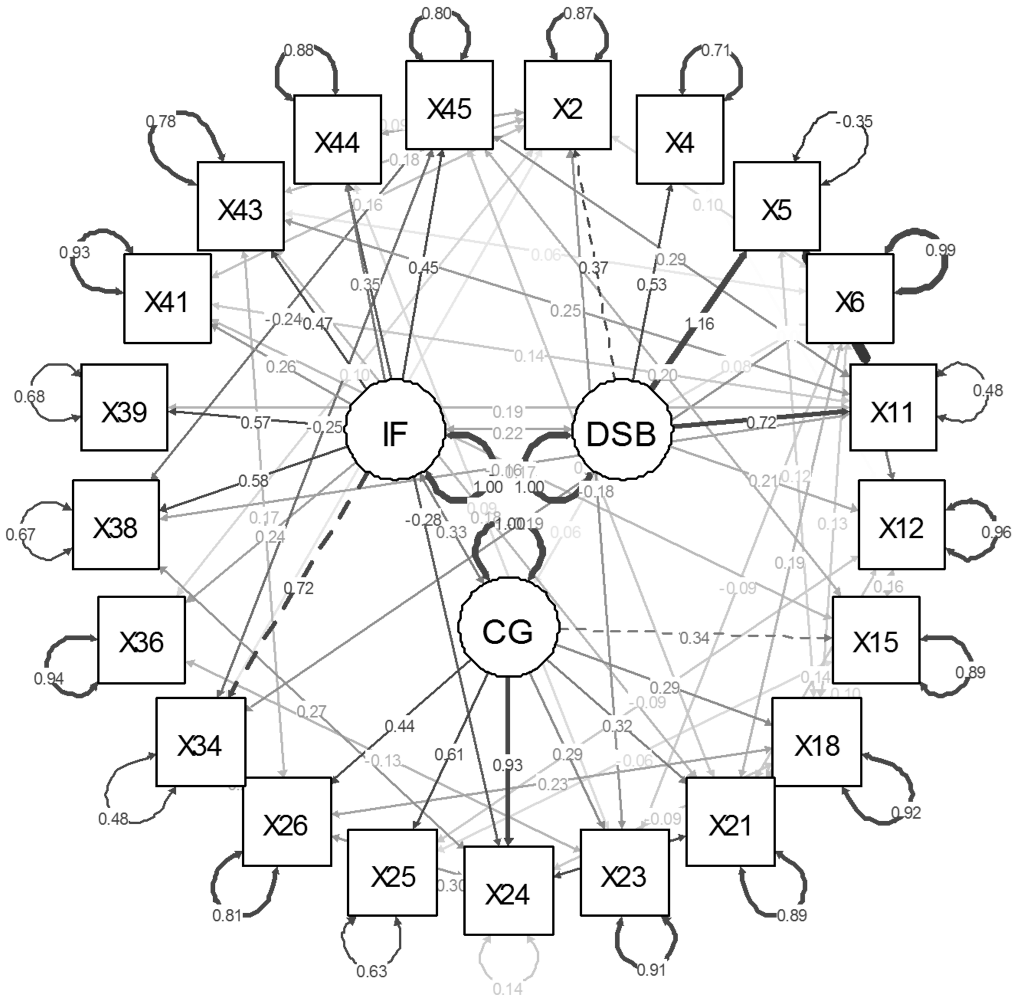
Figure 12 – Parallel Analysis scree plot (43-item set)




Source: Results from the R software prepared by the author.

## APPENDIX F – STRUCTURAL MODEL

Figure 13 – Structural model output




Note: DBS – Demand-Skill Balance; CG – Clear Goals; IF – Immediate Feedback; X – Item.
Source: Results from the R software prepared by the author.

## APPENDIX G – ONLINE FLOW CONDITIONS LEVEL

Table 22 – Online flow conditions proficiency level for each dimension

| n | Country | App | Demand-Skill Balance | Clear Goals | Immediate Feedback |
|---|---|---|---|---|---|
| 1 | BR | Mercado Livre | 1.608 | -0.009 | 0.077 |
| 2 | BR | Shopee | 1.418 | -0.389 | 0.372 |
| 3 | BR | SHEIN | 0.915 | 0.716 | 0.564 |
| 4 | BR | O Boticário | 0.831 | -0.420 | 0.797 |
| 5 | BR | Magalu | 0.831 | -0.420 | 0.797 |
| 6 | BR | Amazon | 0.976 | 1.113 | 0.057 |
| 7 | BR | AliExpress | 0.871 | 0.862 | 0.085 |
| 8 | BR | OLX | 0.536 | 0.626 | -0.415 |
| 9 | BR | C&A | 0.915 | 0.716 | 0.564 |
| 10 | BR | Nike | 1.354 | 0.368 | -0.168 |
| 11 | BR | Lojas Renner | 1.540 | -0.340 | 0.383 |
| 12 | BR | Vale Bonus | 0.377 | 0.388 | 0.219 |
| 13 | BR | Netshoes | 0.724 | 0.300 | 0.868 |
| 14 | BR | Casas Bahia | 0.724 | 0.300 | 0.868 |
| 15 | BR | App comercial Alibaba | 1.564 | 0.963 | -0.291 |
| 16 | BR | Beleza na Web | 0.831 | -0.420 | 0.797 |
| 17 | BR | Méliuz | 0.629 | 0.075 | -0.085 |
| 18 | BR | MadeiraMadeira | 0.724 | 0.300 | 0.868 |
| 19 | BR | Americanas | 0.724 | 0.300 | 0.868 |
| 20 | BR | enjoei | 0.642 | 0.426 | -0.237 |
| 21 | BR | Centauro | 0.832 | -1.134 | 0.741 |
| 22 | BR | Riachuelo | 0.519 | -1.166 | 0.586 |
| 23 | BR | SEPHORA | 1.128 | -0.895 | -0.246 |
| 24 | BR | O Boticário Revendedor | 0.324 | -0.112 | -0.556 |
| 25 | BR | Buscapé | 0.362 | 0.234 | -0.132 |
| 26 | BR | OFF Premium | 0.404 | 0.243 | 0.703 |
| 27 | BR | Shultz | 0.519 | -1.166 | 0.586 |
| 28 | BR | HAVAN | 0.632 | -0.735 | -0.474 |
| 29 | BR | Drogaria São Paulo | -0.342 | -1.703 | 0.114 |
| 30 | BR | Época Cosméticos | 0.719 | -0.467 | 0.791 |
| 31 | BR | Meu Carrefour | 0.516 | 0.115 | -0.402 |
| 32 | BR | Dafiti | 0.404 | 0.243 | 0.703 |
| 33 | BR | Natura | 0.294 | -0.215 | 0.601 |
| 34 | BR | Gazin | -0.393 | -0.034 | 0.628 |
| 35 | BR | Sam's Club Brasil | -0.163 | 0.529 | 0.610 |
| 36 | BR | FARM | -0.233 | 0.732 | 0.198 |
| 37 | BR | Cobasi | 0.804 | -0.554 | 0.423 |
| 38 | BR | Pão de Açúcar Mais | -0.163 | 0.529 | 0.610 |
| 39 | BR | Petlove | 1.435 | 0.392 | 0.459 |

| | | | | | |
|---|---|---|---|---|---|
| 40 | BR | Drogarias Pacheco | -0.342 | -1.703 | 0.114 |
| 41 | BR | Petz | 0.519 | -1.166 | 0.586 |
| 42 | BR | ZARA | -0.707 | 0.219 | 0.065 |
| 43 | BR | Reserva | 0.519 | -1.166 | 0.586 |
| 44 | BR | Melissa Oficial | 0.832 | -1.134 | 0.741 |
| 45 | BR | Pichau | 0.797 | -0.984 | 0.244 |
| 46 | BR | Privalia | -0.136 | -1.255 | -0.080 |
| 47 | BR | KaBuM! | 0.962 | -0.051 | -0.139 |
| 48 | BR | Leroy Merlin | 0.676 | -0.465 | 0.155 |
| 49 | BR | Evino | -0.177 | 0.510 | -0.001 |
| 50 | BR | Chilli Beans | -0.393 | -0.034 | 0.628 |
| 51 | BR | Mercado OXXO | -0.928 | 0.233 | 0.551 |
| 52 | BR | Decathlon | -0.023 | -1.339 | 0.060 |
| 53 | BR | Submarino | -0.110 | -0.710 | 0.715 |
| 54 | BR | Vans | 0.519 | -1.166 | 0.586 |
| 55 | BR | Tok&Stok | -0.335 | -1.578 | -0.035 |
| 56 | BR | Farmácias São João | -0.396 | -0.751 | 0.574 |
| 57 | BR | Hering | -0.295 | -0.705 | 0.572 |
| 58 | BR | Prime Gourmet 5.0 | 0.000 | 0.737 | -0.944 |
| 59 | BR | Pechinchou | 0.883 | -0.811 | -1.176 |
| 60 | BR | Elo7 | 0.337 | -0.548 | -1.251 |
| 61 | BR | Lojas Colombo | -0.340 | -1.401 | 0.524 |
| 62 | BR | Marisa | -0.403 | -0.040 | 0.020 |
| 63 | BR | market4u | -2.110 | -0.959 | -2.028 |
| 64 | BR | Max Atacadista | -0.381 | -0.856 | -2.625 |
| 65 | BR | Crocs | -0.657 | -1.646 | -0.041 |
| 66 | BR | Avon APP | -0.300 | -1.396 | 0.536 |
| 67 | BR | FARFETCH | 1.142 | 0.927 | -0.960 |
| 68 | BR | Compras Paraguai | -0.875 | -0.987 | -0.569 |
| 69 | BR | Supermarket | -0.920 | -1.281 | -1.363 |
| 70 | BR | Mateus Mais | -0.972 | -0.436 | -0.230 |
| 71 | BR | Amobeleza | -0.102 | -1.932 | -1.158 |
| 72 | BR | Kalunga | -0.008 | -0.664 | 0.715 |
| 73 | BR | CASA&VÍDEO | -0.403 | -0.040 | 0.020 |
| 74 | BR | AREZZO | 0.450 | -0.394 | -0.126 |
| 75 | BR | CSSBuy | -0.468 | -0.851 | -0.798 |
| 76 | BR | Prezunic | -0.257 | -0.001 | -0.878 |
| 77 | BR | Anacapri | 0.519 | -1.166 | 0.586 |
| 78 | BR | Farmácia Preço Popular | 0.371 | 0.229 | 0.077 |
| 79 | BR | Pernambucanas | -0.408 | -0.742 | -0.029 |
| 80 | BR | Temu | 0.879 | 0.544 | 0.191 |
| 81 | BR | Zattini | 0.831 | -0.420 | 0.797 |
| 82 | BR | Tenda Atacado | -0.956 | -0.002 | -0.589 |
| 83 | BR | Banggood | 0.879 | 0.544 | 0.191 |
| 84 | BR | Ponto | -0.022 | -0.519 | 0.230 |
| 85 | BR | Clube Extra | -0.314 | -0.817 | -0.360 |

| | | | | | |
|---|---|---|---|---|---|
| 86 | BR | Youcom | -0.058 | -0.461 | -0.292 |
| 87 | BR | Camicado | -0.399 | -0.173 | 0.272 |
| 88 | BR | Loja do Mecânico | 0.744 | -1.300 | -0.457 |
| 89 | BR | Moda Center | -1.768 | -1.010 | -1.806 |
| 90 | BR | eBay | 0.802 | 0.899 | -0.443 |
| 91 | BR | Apple Store | 1.359 | -0.303 | -2.506 |
| 92 | BR | Lojas Torra | -0.682 | -0.291 | -0.926 |
| 93 | BR | glam | -0.946 | 0.107 | 0.156 |
| 94 | BR | Zee.Now | -0.736 | -1.006 | -0.475 |
| 95 | BR | JOHN JOHN | -0.433 | -0.093 | -0.295 |
| 96 | BR | St Marche | -0.445 | -1.044 | -0.432 |
| 97 | BR | Maria Filó | -0.233 | 0.732 | 0.198 |
| 98 | BR | ZZ MALL | 0.654 | -0.097 | -0.283 |
| 99 | BR | Ferreira Costa | 0.371 | 0.229 | 0.077 |
| 100 | BR | Droper | 0.501 | 0.376 | -0.773 |
| 101 | SP | Temu | 0.879 | 0.544 | 0.191 |
| 102 | SP | SHEIN | 0.915 | 0.716 | 0.564 |
| 103 | SP | Miravia | 0.915 | 0.716 | 0.564 |
| 104 | SP | Vinted | 0.301 | 1.289 | -0.463 |
| 105 | SP | Amazon | 0.976 | 1.113 | 0.057 |
| 106 | SP | Wallapop | 0.278 | 1.255 | -1.027 |
| 107 | SP | AliExpress | 0.871 | 0.862 | 0.085 |
| 108 | SP | Vestiaire Collective | 0.820 | 0.977 | -1.125 |
| 109 | SP | Milanuncios | 0.769 | 1.045 | -0.877 |
| 110 | SP | Supermercados DIA | -0.367 | -0.807 | -1.237 |
| 111 | SP | Hacoo | 0.936 | -0.191 | -0.480 |
| 112 | SP | Mercadona | -0.366 | -1.794 | -0.762 |
| 113 | SP | LEROY MERLIN | -0.598 | 0.360 | 0.134 |
| 114 | SP | Zalando | -0.393 | -0.034 | 0.628 |
| 115 | SP | Back Market | 0.724 | 0.300 | 0.868 |
| 116 | SP | PandaBuy | 0.769 | 1.045 | -0.877 |
| 117 | SP | ZARA | -0.709 | 0.360 | -0.377 |
| 118 | SP | El Corte Inglés | -0.216 | 0.351 | 0.317 |
| 119 | SP | Mi Carrefour | 0.516 | 0.115 | -0.402 |
| 120 | SP | Privé by Zalando | 0.062 | -0.679 | 0.120 |
| 121 | SP | IKEA | -0.235 | 0.477 | -0.708 |
| 122 | SP | Nike | 0.496 | 0.089 | -0.267 |
| 123 | SP | H&M | -0.247 | 0.556 | -0.072 |
| 124 | SP | B2B Alibaba | 1.564 | 0.963 | -0.291 |
| 125 | SP | Decathlon | -0.023 | -1.339 | 0.060 |
| 126 | SP | Fnac | 0.556 | 0.794 | -0.069 |
| 127 | SP | PULL&BEAR | 0.465 | -0.261 | 0.221 |
| 128 | SP | Perfumerías DRUNI | -0.005 | -1.360 | 0.675 |
| 129 | SP | Bershka | -0.770 | 0.814 | -0.696 |
| 130 | SP | Sephora | -0.618 | 0.422 | 0.204 |
| 131 | SP | Stradivarius | -1.419 | 0.625 | 0.019 |
| 132 | SP | MANGO | -0.960 | 0.162 | -0.354 |

| | | | | | |
|---|---|---|---|---|---|
| 133 | SP | Perfumerías Primor | -0.653 | 0.292 | 0.689 |
| 134 | SP | CIDER | 0.915 | 0.716 | 0.564 |
| 135 | SP | JD Sports | -1.154 | -0.612 | -0.017 |
| 136 | SP | Tezenis | -0.143 | -0.202 | 0.169 |
| 137 | SP | ManoMano | 0.724 | 0.300 | 0.868 |
| 138 | SP | Costco Wholesale | -1.828 | -0.210 | -1.882 |
| 139 | SP | Lefties | -1.227 | 1.038 | -0.266 |
| 140 | SP | Lidl | -0.128 | 0.010 | 0.151 |
| 141 | SP | ASOS | -0.163 | 0.529 | 0.610 |
| 142 | SP | Private Sport Shop | -1.473 | -0.697 | -0.474 |
| 143 | SP | AUTODOC | -0.733 | -0.306 | -1.024 |
| 144 | SP | Club Alcampo | -1.790 | -0.014 | -1.397 |
| 145 | SP | Mundo Consum | -2.086 | 0.134 | -0.951 |
| 146 | SP | Etsy | 0.275 | 0.979 | -0.149 |
| 147 | SP | eBay | 0.802 | 0.899 | -0.443 |
| 148 | SP | Dhgate | 0.306 | 1.012 | 0.461 |
| 149 | SP | Mi Store | 0.879 | 0.544 | 0.191 |
| 150 | SP | Zara Home | 0.052 | 0.505 | 0.118 |
| 151 | SP | Springfield | -0.687 | 0.160 | -0.004 |
| 152 | SP | StockX | -0.216 | 0.859 | -0.805 |
| 153 | SP | Alvaro Moreno | -0.300 | -1.396 | 0.536 |
| 154 | SP | TOUS | -0.881 | -0.458 | 0.479 |
| 155 | SP | Women'secret | -1.217 | 0.405 | -0.621 |
| 156 | SP | La Sirena Congelados | -0.924 | -1.429 | -0.955 |
| 157 | SP | ALDI Supermercados | -0.953 | 0.413 | -1.146 |
| 158 | SP | PcComponentes | 0.125 | 0.590 | 0.762 |
| 159 | SP | UNIQLO ES | -0.758 | 0.743 | 0.218 |
| 160 | SP | SCALPERS | -1.754 | -1.251 | -0.235 |
| 161 | SP | Parfois | -1.495 | 0.870 | -0.046 |
| 162 | SP | SNIPES | -0.114 | -0.344 | 0.251 |
| 163 | SP | Sprinter | -1.148 | 0.246 | -0.046 |
| 164 | SP | Apple Store | 1.359 | -0.303 | -2.506 |
| 165 | SP | FARFETCH | 1.142 | 0.927 | -0.960 |
| 166 | SP | idealo | 0.078 | -0.260 | -0.370 |
| 167 | SP | Massimo Dutti | -0.290 | 0.964 | -0.441 |
| 168 | SP | Fashion Nova | 0.915 | 0.716 | 0.564 |
| 169 | SP | Privalia | -1.031 | 0.502 | -0.798 |
| 170 | SP | OYSHO | -0.330 | 0.936 | -0.792 |
| 171 | SP | Saigu Cosmetics | -0.958 | -1.460 | 0.031 |
| 172 | SP | Rituals Cuerpo y hogar | -0.949 | 0.436 | 0.158 |
| 173 | SP | Alcampo | -0.508 | -0.466 | -0.059 |
| 174 | SP | Lyst | -0.187 | 0.652 | -0.445 |
| 175 | SP | ABOUT YOU | 0.013 | 0.942 | 0.311 |
| 176 | SP | BROWNIE | 0.404 | 0.243 | 0.703 |
| 177 | SP | Cortefiel | -0.176 | 0.674 | 0.139 |
| 178 | SP | Naturitas | -0.653 | 0.292 | 0.689 |

| | | | | | |
|---|---|---|---|---|---|
| 179 | SP | Action | -1.435 | -0.192 | -0.657 |
| 180 | SP | Myprotein | -0.931 | 0.361 | -0.481 |
| 181 | SP | C&A | 0.853 | 0.932 | 0.158 |
| 182 | SP | KIKO MILANO | -1.255 | 0.271 | 0.378 |
| 183 | SP | Foot Locker | -0.850 | -0.580 | 0.148 |
| 184 | SP | BM Supermercados | -1.280 | -0.058 | -1.414 |
| 185 | SP | Kiwoko | -0.566 | -0.532 | 0.273 |
| 186 | SP | GLOWRIAS | -0.980 | 0.218 | -0.691 |
| 187 | SP | Veepee | -0.189 | 0.232 | -0.517 |
| 188 | SP | Hollister So Cal Style | -0.932 | 0.220 | -0.045 |
| 189 | SP | Levi's | -0.678 | 0.498 | 0.288 |
| 190 | SP | Catawiki | -0.848 | -0.310 | -0.537 |
| 191 | SP | Casa del Libro | -0.559 | -0.402 | 0.625 |
| 192 | SP | HagoBuy | -1.027 | -0.699 | -1.705 |
| 193 | SP | Showroomprive | -0.280 | 0.993 | 0.142 |
| 194 | SP | SUGARGOO | -0.837 | -0.455 | -1.634 |
| 195 | SP | Fulanita | -0.508 | 0.440 | -1.785 |
| 196 | SP | Instant Gaming | -0.463 | 1.154 | 0.154 |
| 197 | SP | Forum Sport | -1.423 | 0.595 | -1.263 |
| 198 | SP | adidas CONFIRMED | -1.543 | 0.921 | -1.649 |
| 199 | SP | La Roca Village | -0.075 | -0.192 | -0.069 |
| 200 | SP | KIABI | 0.749 | -0.354 | 0.015 |

Note: BR – Brazil; SP – Spain.
Source: Results from the R software prepared by the author.