
**Estudo exploratório sobre a tolerância de
usuários a diferentes níveis de confiabilidade de
software**

Matheus Bastos Neves



Universidade Federal de Uberlândia
Faculdade de Computação
Programa de Pós-Graduação em Ciência da Computação

Uberlândia
2024

Matheus Bastos Neves

**Estudo exploratório sobre a tolerância de
usuários a diferentes níveis de confiabilidade de
software**

Dissertação de mestrado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Prof. Dr. Rivalino Matias Júnior

Uberlândia
2024

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

N518
2024

Neves, Matheus Bastos, 1985-
Estudo exploratório sobre a tolerância de usuários a diferentes níveis de confiabilidade de software [recurso eletrônico] / Matheus Bastos Neves. - 2024.

Orientador: Rivalino Matias Júnior.
Dissertação (Mestrado) - Universidade Federal de Uberlândia, Pós-graduação em Ciência da Computação.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.di.2024.679>
Inclui bibliografia.

1. Computação. I. Matias Júnior, Rivalino ,1971-, (Orient.). II. Universidade Federal de Uberlândia. Pós-graduação em Ciência da Computação. III. Título.

CDU: 681.3

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em Ciência da
Computação

Av. João Naves de Ávila, 2121, Bloco 1A, Sala 243 - Bairro Santa Mônica, Uberlândia-MG,
CEP 38400-902

Telefone: (34) 3239-4470 - www.ppgco.facom.ufu.br - cpgrafacom@ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciência da Computação				
Defesa de:	Dissertação, 40/2024, PPGCO				
Data:	25 de setembro de 2024	Hora de início:	14:00	Hora de encerramento:	16:30
Matrícula do Discente:	12022CCP008				
Nome do Discente:	Matheus Bastos Neves				
Título do Trabalho:	Estudo exploratório sobre a tolerância de usuários a diferentes níveis de confiabilidade de software				
Área de concentração:	Ciência da Computação				
Linha de pesquisa:	Engenharia de Software				
Projeto de Pesquisa de vinculação:	-----				

Reuniu-se por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Lucio Borges de Araujo - IME/UFU, Paulo Romero Martins Maciel - CIn/UFPE e Rivalino Matias Júnior - FACOM/UFU, orientador do candidato.

Os examinadores participaram desde as seguintes localidades: Paulo Romero Martins Maciel - Recife/PE. Os outros membros da banca e o aluno participaram da cidade de Uberlândia.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. Rivalino Matias Júnior, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação da Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir ao candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

Aprovado

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Rivalino Matias Júnior, Professor(a) do Magistério Superior**, em 27/09/2024, às 10:44, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Lucio Borges de Araujo, Professor(a) do Magistério Superior**, em 27/09/2024, às 12:51, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Paulo Romero Martins Maciel, Usuário Externo**, em 22/10/2024, às 11:19, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5687010** e o código CRC **3E67BDA5**.

Este trabalho é dedicado todos aqueles que não desistiram e conseguiram forças para seguir na luta e superar a COVID19 e as adversidades.

Agradecimentos

Agradeço a Ayla Bafile Elian por me incentivar, apoiar e suportar na realização de um sonho, que é este trabalho. Agradeço ao Prof. Rivalino Matias Jr. pelo conhecimento compartilhado e a dedicação ao longo desta jornada. Agradeço ao Caio Augusto Rodrigues dos Santos pelo apoio no grupo de pesquisas. Agradeço também aos demais professores e colaboradores da PPGCO/FACOM/UFU pelo apoio neste projeto.

*"Me and yo' daughter, got's this thing going on
(We got a special kind of thing going on)
You say it's puppy love
We say it's full grown
Hope that we feel this, feel this way forever
You can plan a pretty picnic
But you can't predict the weather, Ms. Jackson"
(Outkast)*

*"Failure is success in progress"
(Albert Einstein)*

Resumo

A confiabilidade do software vai além da simples ausência de falhas, sendo um dos principais fatores que influenciam a satisfação dos usuários. A percepção de confiabilidade pelos usuários é afetada por fatores como o contexto de uso, expectativas e a gravidade das falhas. A tolerância, por sua vez, representa a disposição dos usuários em aceitar essas falhas e é condicionada por diversos elementos, como a importância da tarefa, a disponibilidade de alternativas, os custos envolvidos, a frequência das falhas e o impacto causado. O principal objetivo desta pesquisa foi identificar os fatores determinantes da tolerância a falhas de software em diferentes cenários e analisar a relação entre as variáveis de aceitabilidade, satisfação e qualidade percebida da experiência. A pesquisa utilizou um questionário que apresentou diferentes cenários de falhas, investigando fatores relacionados à tolerância e coletando dados sobre a satisfação, aceitação e percepção geral de qualidade da experiência. A análise dos dados foi feita por meio da técnica de análise de variância (ANOVA) não paramétrica. Além disso, uma escala Likert foi desenvolvida para quantificar a tolerância, considerando os fatores identificados ao longo do estudo e estabelecendo um conceito de tolerância. A pesquisa demonstrou que o nível de tolerância a falhas pode ser mensurado de forma objetiva e sistemática, permitindo estabelecer uma mediana como referência para avaliar o impacto das falhas na qualidade da experiência (QoE). Foram identificados fatores humanos, técnicos e econômicos que influenciam a tolerância, e ficou clara a relação positiva entre maior tolerância e maior satisfação do usuário. Essas descobertas são fundamentais para o aprimoramento contínuo de produtos de software, com o objetivo de aumentar a satisfação, a lealdade dos usuários e melhorar a qualidade da experiência.

Palavras-chave: Confiabilidade de Software, Confiabilidade do serviço percebida pelo usuário, Qualidade da Experiência (QoE), Falhas de Software, Satisfação.

Abstract

Software reliability goes beyond the mere absence of failures and is one of the main factors influencing user satisfaction. Users' perception of reliability is affected by factors such as the context of use, expectations, and the severity of failures. Tolerance, in turn, represents the users' willingness to accept these failures and is conditioned by various elements, such as the importance of the task, availability of alternatives, costs involved, frequency of failures, and the impact caused. The main objective of this research was to identify the determining factors of tolerance to software failures in different scenarios and to analyze the relationship between the variables of acceptability, satisfaction, and perceived quality of experience. The research employed a questionnaire presenting different failure scenarios, investigating factors related to tolerance and collecting data on satisfaction, acceptance, and overall perception of experience quality. Data analysis was performed using the technique of non-parametric analysis of variance (ANOVA). Additionally, a Likert scale was developed to quantify tolerance, considering the factors identified throughout the study and establishing a concept of tolerance. The research demonstrated that the level of tolerance to failures can be measured objectively and systematically, allowing the establishment of a median as a reference to assess the impact of failures on Quality of Experience (QoE). Human, technical, and economic factors influencing tolerance were identified, and the positive relationship between higher tolerance and greater user satisfaction became clear. These findings are fundamental for the continuous improvement of software products, aiming to increase user satisfaction, user loyalty, and enhance the quality of the experience.

Keywords: Software Reliability, User Perceived Service Reliability, Quality of Experience (QoE), Software Failure, Satisfaction.

Lista de ilustrações

Figura 1 – Fatores relacionados à tolerância	32
Figura 2 – Diagrama ServQual <i>Gap 5</i> (Zona de Tolerância)	54
Figura 3 – Fatores relacionados ao serviço desejado e insatisfatório (Zona de Tolerância).....	58
Figura 4 – Percepção do consumidor sobre a qualidade e a satisfação do cliente.	59
Figura 5 – Percepção do consumidor sobre as funcionalidades do produto ou serviço.	62
Figura 6 – Resumo do modelo teórico da relação entre satisfação e a qualidade da experiência.....	63
Figura 7 – Modelo conceitual para qualidade dos serviços eletrônicos	77
Figura 8 – Protocolo de pesquisa	107
Figura 9 – Soma dos valores atribuídos para cada fator	123
Figura 10 – Desconfirmação dos valores atribuídos para cada fator	124
Figura 11 – <i>Box plot</i> comparando a desconfirmação entre os cenários.....	125
Figura 12 – <i>Box plot</i> comparando as medidas de Qualidade, Satisfação e Aceitabilidade entre os cenários	126
Figura 13 – <i>Box plot</i> comparando as medidas da Pontuação da escala de Tolerância entre os cenários de sucesso e falha	127
Figura 14 – Gráfico de declividade (<i>scree plot</i>).....	142
Figura 15 – Mensagem de divulgação com o convite de participação.....	173

Lista de tabelas

Tabela 1 – Resumo dos trabalhos relacionados sobre confiabilidade percebida . . .	74
Tabela 2 – Resumo dos estudos relacionados sobre Qualidade de Serviço (SQ) . .	81
Tabela 3 – Afirmativas positivas e negativas das dimensões de Tolerância.....	87
Tabela 4 – Cenários de pesquisa	94
Tabela 5 – Questionário variáveis dependentes.....	101
Tabela 6 – Questionário de controle.....	103
Tabela 7 – Questionário percepção da funcionalidade	104
Tabela 8 – Questionário perfil.....	105
Tabela 9 – Perfil dos participantes.....	116
Tabela 10 – Visão sobre funcionalidades de software e finalidade de uso	118
Tabela 11 – Análise descritiva da medida de Satisfação.....	119
Tabela 12 – Análise descritiva da medida de Qualidade.....	119
Tabela 13 – Análise descritiva da medida de Aceitabilidade.....	120
Tabela 14 – Análise descritiva da medida de Tolerância.....	121
Tabela 15 – Análise descritiva da medida de Desconfirmação	122
Tabela 16 – Resultados do teste de Friedman	128
Tabela 17 – Resultados do teste de Kruskal-Wallis	129
Tabela 18 – Resultados das comparações em pares do teste de Kruskal-Wallis	130
Tabela 19 – Resultados do teste de Kruskal-Wallis para a medida de Desconfirmação	131
Tabela 20 – Resultados das comparações em pares do teste de Kruskal-Wallis para a medida de Desconfirmação	132
Tabela 21 – Resultados da decisão de troca em cenários de falha	134
Tabela 22 – Análise descritiva dos atributos da escala de Tolerância	134
Tabela 23 – Matriz de correlação da Aceitabilidade com cada atributo.....	135
Tabela 24 – Matriz de correlação entre atributos da escala de Tolerância	136
Tabela 25 – Cargas fatoriais da análise fatorial exploratória.....	137
Tabela 26 – Resumo dos fatores da análise fatorial exploratória.....	139
Tabela 27 – Correlação entre fatores da análise fatorial exploratória	139

Tabela 28 – Medidas de aderência da análise fatorial exploratória	140
Tabela 29 – Teste de esfericidade de Bartlett.....	140
Tabela 30 – Medida de adequação da amostragem de Kaiser-Meyer-Olkin (KMO) .	141
Tabela 31 – Autovalores iniciais (<i>Eigenvalues</i>)	142
Tabela 32 – Análise descritiva do escore dos fatores da análise fatorial exploratória	143
Tabela 33 – Resultados do teste de Kruskal-Wallis dos escores dos fatores	144
Tabela 34 – Resultados das comparações em pares para escores dos fatores.....	145

Lista de siglas

API *Application Programming Interface*

ANOVA Análise de Variância

AFE Análise Fatorial Exploratória

AaaS *Anything as a Service*

BSOD *Blue Screen of Death*

BTS Teste de Bartlett

BIC *Bayesian Information Criterion*

CEP Comitê de Ética em Pesquisa

CONEP Comissão Nacional de Ética em Pesquisa

df Graus de Liberdade

HCI Interfaces homem-máquina

IC Intervalo de Confiança

KMO Teste de Kaiser-Meyer-Olkin

LGPD Lei Geral de Proteção de Dados

MPC *Multi-Party Chain*

MOS Opinião média dos usuários (*Mean Opinion Score*)

MTBF Tempo médio entre falhas (*Mean time between failures*)

MTTF Tempo médio até a falha (*Mean time to failure*)

PDF *Portable Document Format*

QoS Qualidade de Serviço

QoE Qualidade da Experiência

QoSE Qualidade de Serviço Experimentado

RMSEA *Root Mean Square Error of Approximation*

SQ Qualidade de Serviço

SO Sistema Operacional

SaaS *Software as a Service*

SLA *Service Level Agreement*

TCLE Termo de Consentimento Livre e Esclarecido

TI Tecnologia da Informação

TIC Tecnologia da Informação e Comunicação

TLI *Tucker-Lewis Index*

UFU Universidade Federal de Uberlândia

UX *User Experience*

WEB *World Wide Web*

ZoT Zona de Tolerância

Sumário

1	INTRODUÇÃO	25
1.1	Contextualização	25
1.2	Justificativa	29
1.3	Objetivos	35
1.3.1	Objetivo geral.....	35
1.3.2	Objetivos específicos	35
1.4	Perguntas de Pesquisa e Hipóteses	35
1.5	Metodologia	37
2	REFERENCIAL TEÓRICO	41
2.1	Introdução	41
2.2	Fundamentos	42
2.2.1	Confiabilidade de software	42
2.2.2	Taxonomia.....	46
2.2.3	Métricas de confiabilidade de software.....	47
2.2.4	Qualidade de serviço.....	51
2.2.5	Qualidade da Experiência (QoE).....	65
3	TRABALHOS RELACIONADOS	69
3.1	Confiabilidade percebida pelo usuário	69
3.2	Estudos relacionados a Qualidade de Serviços	75
4	MATERIAL E MÉTODOS	83
4.1	Introdução	83
4.2	Material	84
4.2.1	Ambiente de execução	85
4.2.2	Conceito de tolerância	86
4.2.3	Cenários de pesquisa	91

4.2.4	Questionário	101
4.2.5	Tamanho da amostra	106
4.3	Métodos	106
4.3.1	Procedimento da pesquisa.....	108
4.3.2	Procedimento para determinar o tamanho da amostra.....	109
4.3.3	Procedimento para realizar a análise descritiva.....	110
4.3.4	Procedimento para validar diferenças significativas em medidas de cenários pareados	110
4.3.5	Procedimento para validar diferenças significativas em medidas diferentes.....	111
4.3.6	Procedimento para analisar a correlação e realizar a análise fatorial exploratória	112
4.3.7	Extração das cargas fatoriais	113
5	RESULTADOS	115
5.1	Introdução	115
5.2	Perfil dos participantes	116
5.3	Análise descritiva	118
5.3.1	Satisfação	118
5.3.2	Qualidade	119
5.3.3	Aceitabilidade	120
5.3.4	Tolerância	121
5.3.5	Desconfirmação.....	122
5.3.6	Atributos de Tolerância.....	123
5.4	Comparação entre cenários	125
5.5	Comparação entre medidas	129
5.5.1	Medidas de Qualidade, Satisfação e Aceitabilidade	129
5.5.2	Medidas de desconfirmação	131
5.6	Análise das decisões de troca em cenários de falha	133
5.7	Análise fatorial exploratória e Correlação	134
5.7.1	Nomeação dos fatores.....	138
5.7.2	Estatísticas dos fatores	139
5.7.3	Aderência do modelo.....	140
5.7.4	Verificações de suposições	140
6	CONCLUSÃO	147
6.1	Conclusões da Pesquisa	147
6.1.1	Introdução	147
6.1.2	Respostas para as perguntas de pesquisa	148
6.1.3	Validação do modelo teórico.....	152
6.1.4	Associação com o ServQual.....	153

6.1.5	Associação com o paradoxo da recuperação	154
6.1.6	Conclusão	155
6.1.7	Dificuldades Encontradas.....	156
6.2	Ameaças à Validade.....	157
6.2.1	Validade interna	157
6.2.2	Validade externa	157
6.2.3	Validade de constructo.....	158
6.3	Trabalhos Futuros.....	158
6.3.1	Exploração de diversos contextos de uso	158
6.3.2	Análise longitudinal	158
6.3.3	Investigação de intervenções para melhorar a Tolerância	158
6.3.4	Exploração de fatores culturais e demográficos.....	159
6.3.5	Modelagem preditiva de tolerância a falhas	159
	REFERÊNCIAS.....	161

APÊNDICES

165

APÊNDICE A	- DADOS COLETADOS.....	167
APÊNDICE B	- ANEXOS	173
B.0.1	Convite ao participante.....	173
B.0.2	Parecer CEP	174
B.0.3	Questionário.....	175

Introdução

1.1 Contextualização

Os serviços proporcionados pelos softwares tornaram-se uma parte essencial da sociedade moderna, fornecendo soluções para uma ampla gama de necessidades e funções. No entanto, a crescente dependência da sociedade nesses sistemas também aumenta a vulnerabilidade a falhas, que podem ter graves consequências financeiras, também colocando em perigo vidas humanas e afetando serviços básicos (LAPRIE, 1995). À medida que as consequências dessas falhas impactam cada vez mais o desempenho das organizações e a vida cotidiana, a confiabilidade de software torna-se uma prioridade cada vez mais crítica. Assim, é importante garantir que os sistemas de software sejam confiáveis, seguros e capazes de lidar com possíveis falhas de forma adequada.

A construção de software confiável é relatada por Pressman (2010) como um desafio complexo, que envolve diversos fatores técnicos e de gerenciamento. Embora seja possível validar perfis operacionais¹ e cenários de uso para garantir a confiabilidade do software, isso pode ser custoso e demandar um tempo significativo. Além disso, a garantia da qualidade do software pode ser afetada pelas restrições orçamentárias do projeto, como o custo e o prazo necessários para a entrega do produto, necessitando assim de um processo sistemático para priorização dos esforços de desenvolvimento. No entanto, é importante ressaltar que a confiabilidade não é o único atributo de qualidade do software. Outros fatores, como a clareza dos requisitos, o design robusto e a implementação cuidadosa também podem contribuir para a confiabilidade e a qualidade do software (LYU, 2007). Em última análise, fornecer produtos de qualidade, que possam garantir a confiança dos usuários, é essencial para o sucesso do software e do projeto em si.

A qualidade do software pode ser definida como a capacidade do produto de atender às necessidades de seus usuários em termos de atributos e características específicos. Pressman (2010) relacionou vários atributos de qualidade que são relevantes para a engenharia de software, incluindo funcionalidade, usabilidade, confiabilidade, desempenho

¹ Utiliza-se aqui o conceito de perfil operacional de Musa (1993).

e capacidade de suporte. Embora todos esses atributos sejam importantes para a qualidade do software, alguns deles, como a confiabilidade, podem ser notados pelos usuários (PRESSMAN, 2010). Por exemplo, um software confiável pode garantir a segurança dos dados dos usuários e evitar perda de informações críticas em caso de falha. Desse modo, é essencial garantir que o software seja desenvolvido com todos esses atributos, visando atender às expectativas dos usuários e garantir o sucesso do projeto.

A confiabilidade de software é uma medida quantitativa baseada na probabilidade de sucesso do sistema em desempenhar sua função prevista dentro dos limites de projeto especificados (PHAM, 2007). Em suma, a confiabilidade é a capacidade do software de funcionar corretamente sem falhas. No entanto, a construção de software confiável é um desafio para as equipes de desenvolvimento, pois é necessário encontrar um equilíbrio entre os parâmetros críticos do projeto, de maneira a atender as necessidades dos clientes e usuários. Pressman (2010) destaca esse desafio e enfatiza a importância de uma abordagem sistemática para garantir a confiabilidade do software. Para suprir essa necessidade, a indústria está cada vez mais empenhada na aplicação de técnicas de engenharia de confiabilidade de software (LYU, 2007). Essas técnicas incluem prevenção, remoção, tolerância de faltas e previsão de falhas, e são desenvolvidas através da análise de confiabilidade e outras abordagens, o que auxilia na melhora da confiabilidade dos produtos e processos de software (LYU, 2007). Apesar do desafio na aplicação dessas técnicas, as mesmas podem resultar em produtos de software de maior qualidade e melhor experiência do usuário.

A busca por serviços em “nuvem” tornou-se um modelo de solução cada vez mais popular para empresas que desejam flexibilidade e agilidade em suas operações. A nuvem é uma metáfora para uma rede global de serviços remotos que opera como um ecossistema único, conectado à Internet, e ela é composta por infraestrutura, servidores, armazenamento e outros componentes que permitem a oferta de serviços. Essas empresas, que valorizam a praticidade, têm optado por demandar serviços em vez de produtos de prateleira, com o objetivo de desonerar o custo com operação de TI. A utilização de serviços em nuvem é uma estratégia comum de terceirização adotada por muitas empresas, visando aliviar o trabalho de gestão de TI e melhorar a qualidade de suas operações por meio de atendimento especializado (MEIRELLES, 2022).

Em sua pesquisa anual acerca do mercado de tecnologia, Meirelles (2022) indica alguns exemplos de softwares de sucesso, como o ERP Salesforce, que é disponibilizado em nuvem, e relata casos de conversão de diversos produtos tradicionais para esse modelo. A oferta de serviços, em vez de produtos, evoluiu para um modelo conhecido como *Software as a Service* (SaaS). Esse modelo é constituído pela oferta de softwares através de serviços disponibilizados na nuvem, e a interação do usuário é normalmente realizada através de interfaces WEB ou APIs. O modelo baseado na oferta de serviços desencadeou uma tendência de que tudo pode ser oferecido como serviço, e assim passou a ser referência, intitulado como *Anything as a Service* (AaaS). Esse modelo foi assim chamado por se

tratar de um modelo bastante flexível, que possibilitou o fornecimento não somente de serviços de softwares, mas de uma gama de serviços de qualquer tipo, desde que o recurso possa ser disponibilizado para consumo através da nuvem.

Um exemplo real é o reposicionamento do produto Microsoft Office, que passou a disponibilizar a sua versão completamente na nuvem, o Microsoft Office 365, fazendo frente a outras opções já disponíveis no mercado. Essa modalidade em nuvem também tem a vantagem de possibilitar ao usuário o uso de qualquer dispositivo sem a necessidade de realizar uma nova instalação.

Uma vez que o foco é objetivo de negócio, o risco da operação dos recursos de Tecnologia da Informação (TI) é transferido aos seus fornecedores, tornando a dependência um risco gerenciado através do estabelecimento de uma relação baseada em acordos de nível de qualidade de serviços (*Service Level Agreement (SLA)*), que é avaliada através das métricas de qualidade de serviço.

Nesse contexto, é importante ressaltar que o estabelecimento de métricas de qualidade de serviços passou a ser uma premissa na prestação de serviços por meio de softwares. O termo Qualidade de Serviço (QoS) surgiu na área de telecomunicações e passou a ser um assunto estudado em vários contextos, além de amplamente aplicado na indústria. Entidades como ITUT (2008) buscam padronizar o entendimento de alguns termos amplamente utilizados, entre eles a qualidade. Para o ITUT (2008) define a qualidade como a totalidade das características de uma entidade que influenciam sua capacidade de satisfazer necessidades declaradas e implícitas, desde que as características possam ser observáveis e/ou mensuráveis. A QoS foi estabelecida a partir do entendimento de que a qualidade deve ser algo mensurado diretamente através da observação da performance do produto. A QoS portanto é definida como a totalidade das características de um serviço de telecomunicações que condicionam a sua capacidade para satisfazer as necessidades declaradas e implícitas do usuário do serviço (ITUT, 2008). As características são percebidas como propriedades que diferenciam indivíduos uns dos outros dentro de uma determinada população, podendo ser mensurada quantitativamente (por variáveis) ou qualitativamente (por atributos).

A QoS pode ser estimada, por exemplo, a partir de métricas de confiabilidade. Essas métricas normalmente são calculadas a partir da taxa de falhas, Tempo médio até a falha (*Mean time to failure*) (MTTF) e Tempo médio entre falhas (*Mean time between failures*) (MTBF) (CLULEY, 2003). A QoS pode ser estabelecida por um conjunto variado e por outros tipos de métricas. Alguns exemplos são: *bandwidth* (largura de banda), *delay* (latência), *jitter* (eventos de gargalo de rede), *packet loss* (perda de pacotes), *bit-rate* (volume de dados por segundo), *throughput* (vazão), *encoding* (formato em que o arquivo é codificado), entre outras. Cada uma dessas métricas contribui para avaliar a qualidade da transmissão e garantir a confiabilidade e eficiência dos serviços oferecidos em uma rede, especialmente em contextos de tempo real, como o tráfego de voz e vídeo

(BULGARELLA et al., 2019).

É importante também notar que a QoS é uma medida extraída da observação direta da performance do produto ou serviço e é utilizada para o acompanhamento dos níveis de qualidade de serviço acordados, podendo variar de acordo com o tipo de serviço prestado e que, em alguns casos, pode ser necessário utilizar outras métricas específicas para avaliar a qualidade do serviço.

Möller e Raake (2014) relatam que a indústria comumente assumiu que o alto desempenho de uma QoS leva a uma alta aceitação dos respectivos serviços providos. Entretanto, apontam uma série de exemplos de serviços que, apesar de apresentarem altos níveis de QoS, são percebidos pelos usuários como tendo uma performance insatisfatória. Nesse sentido, a comunidade passou a buscar entender quais desses fatores são relevantes para o usuário, em relação às métricas clássicas de QoS. Com o objetivo de dar um peso às métricas clássicas de qualidade que geram impacto direto na percepção do usuário, o ITUT (2008) definiu um tipo de QoS que avalia a performance do produto quando experimentado e percebido pela opinião de seus usuários (Qualidade de Serviço Experimentado (QoSE)). A métrica proposta pelo é definida como uma declaração que expressa o nível de qualidade que os usuários acreditam ter experimentado, através da classificação de sua opinião. A métrica é composta por dois componentes principais: quantitativo e qualitativo. O componente quantitativo pode ser influenciado pelos efeitos providos do sistema (como a infraestrutura de rede, por exemplo) e o componente qualitativo pode ser influenciado pelas expectativas do utilizador, condições ambientais, fatores psicológicos, contexto de aplicação etc. (ITUT, 2008).

Desde então, a indústria percebeu que investir apenas na melhoria da QoS não é suficiente para a aceitação de um produto por parte de seus usuários, pois a QoS não contempla fatores externos, como a percepção do usuário, que é também influenciada por fatores emocionais e suas reações ao interagir com o produto ou serviço. Com isso, métricas focadas em medir a qualidade sob o ponto de vista dos seres humanos, conhecidas como Qualidade da Experiência (QoE), passam a se tornar populares na comunidade.

Diferentemente da QoS, a QoE é uma medida que leva em conta fatores humanos, expectativas e o contexto. A QoE é definida por como o grau de prazer ou aborrecimento de uma pessoa ao utilizar uma aplicação, serviço ou sistema. Portanto, a QoE resulta da avaliação do usuário acerca do cumprimento de suas expectativas e necessidades com relação a um determinado serviço ou software, tendo como critérios a utilidade e/ou prazer, o contexto, a personalidade e o estado de espírito da pessoa no momento da utilização (KUIPERS et al., 2010).

Nesse sentido, a Opinião média dos usuários (*Mean Opinion Score*) (MOS) é um exemplo típico de métrica QoE, e é aferida a partir da resposta típica à pergunta “Como você avalia a sua experiência durante o uso do aplicativo?”. Dentro desse contexto, a pergunta pode variar a depender de uma característica do produto ou serviço, como a

qualidade do áudio ou a qualidade do vídeo, como relatado por Kuipers et al. (2010).

A relação entre QoS e QoE é um assunto estudado, sobretudo em alguns cenários, como os de produtos que fazem a transmissão de vídeos na nuvem, conforme relatado por Kuipers et al. (2010). Quando algumas métricas de QoS estão relacionadas à qualidade da conectividade (como, por exemplo, caso o usuário se depare com latências, lentidão e perda de sincronismo), a QoE do usuário é diretamente afetada e causa uma frustração no espectador, devido às falhas percebidas. No entanto, há casos em que essa relação não é direta.

Varela, Skorin-Kapov e Ebrahimi (2014) relatam casos de softwares que prestavam um nível baixo de QoS e eram percebidos com um nível alto de QoE para seus usuários. Um exemplo desse tipo de software é o Sistema Operacional (SO) Windows, em que o mesmo é conhecido pelos erros frequentes, entre eles, a “tela azul” (*Blue Screen of Death* (BSOD)). Entretanto o Windows é o sistema operacional mais utilizado do mercado (MEIRELLES, 2022). A aceitação do produto, mesmo como a presença de falhas, é relatada e supostamente explicada, pelo fato de a equipe de desenvolvimento do sistema operacional ter como objetivo, além da melhoria contínua da confiabilidade, a implementação da lista de desejos de seus usuários (MURPHY; LEVIDOW, 2000).

Dentro desse cenário, a confiabilidade é um fator chave na percepção da qualidade de um produto, e entender como os usuários avaliam a confiabilidade do software é crucial para o sucesso de um produto ou serviço. No entanto, há casos em que os usuários toleram falhas de software, como no exemplo do sistema operacional Windows, que apesar de apresentar falhas frequentes, continua sendo amplamente utilizado. Isso demonstra que a percepção de confiabilidade dos usuários pode ser diferente da análise técnica da confiabilidade do software.

Nesse contexto, o objetivo deste trabalho é investigar os fatores que influenciam a percepção do usuário sobre a confiabilidade do software e compreender por que os usuários podem tolerar falhas. Abordar esse desafio de pesquisa é fundamental para desenvolver produtos e serviços que atendam às expectativas dos usuários.

1.2 Justificativa

A indústria já adota práticas que visam satisfazer as necessidades dos usuários, através de funcionalidades de software que os agradem (PRESSMAN, 2010). A experiência tem demonstrado que investir em estratégias que levam em consideração a percepção do usuário pode aumentar a atratividade e a vantagem econômica do produto (ZEITHAML; BITNER; GREMLER, 2017). Ainda que as falhas em produtos sejam consideradas como o fator determinante na relação de confiança entre o usuário e o produto, é importante notar que nem todas as falhas são percebidas da mesma forma pelos usuários (YABE et al., 2018).

Alguns exemplos são bem conhecidos na comunidade, tais como: i) a falha de quatro horas do software WhatsApp que, em 2014, motivou os seus usuários a temporariamente utilizar o produto concorrente, o Telegram (HAMBURGER, 2014); ii) a falha da Fiat Chrysler que impossibilitava o desligamento do sistema Cruise Control em seus carros, situação em que a empresa precisou realizar atualizações do software embarcado em seus veículos para garantir a segurança dos usuários (SHEPARDSON, 2018); iii) a falha do banco HSBC, que deixou os usuários sem acesso às suas contas bancárias por dois dias (FLINDERS, 2016).

Assim como os exemplos anteriores, muitos produtos de software, incluindo sistemas operacionais amplamente utilizados, apresentam taxas variadas de falhas (DOS SANTOS; MATIAS JR; TRIVEDI, 2021). Relatórios anuais, como o do Uptime Institute (2023, 2021, 2020), destacam falhas causadas na infraestrutura necessária para a execução efetiva de serviços de software em nuvem, reportando interrupções ocorridas e os impactos causados. Ainda assim, muitos desses produtos continuam sendo largamente usados. O exemplo citado, do sistema operacional Windows, que detém a maior fatia de mercado para desktops (MEIRELLES, 2022), reflete o impacto relatado, pois apresenta diferentes tipos de falhas e ainda é amplamente adotado (MURPHY; LEVIDOW, 2000; DOS SANTOS; MATIAS JR, 2016).

Nesse sentido, uma pergunta importante a se fazer é: o que leva usuários a utilizar produtos de software que sabidamente apresentam falhas? Que fatores influenciam a tolerância dos usuários quanto ao uso de produtos de software que falham? Essas e outras perguntas relacionadas são importantes para entender como as falhas de software são percebidas pelos usuários

Os exemplos mencionados ilustram a capacidade dos usuários de tolerar falhas e continuar utilizando os produtos e serviços oferecidos por essas empresas. Por essa razão, compreender por que os usuários toleram falhas e identificar os fatores relacionados à percepção do usuário em relação à qualidade do software são questões de pesquisa relevantes e ainda não respondidas que orientam este estudo, cujo objetivo é investigar o conceito de tolerância dos usuários em relação a falhas de software, examinando os diversos fatores que influenciam esse comportamento. Mais especificamente, buscamos compreender como as falhas de software impactam a tolerância do usuário em relação à prestação de um serviço realizado por um produto de software. Além disso, pretende-se identificar os fatores mais significativos em termos de seus efeitos sobre a tolerância do usuário. Os resultados deste estudo têm o propósito de contribuir para uma melhor compreensão da relação entre a qualidade da experiência do usuário e a confiabilidade do software.

A maioria dos modelos de confiabilidade se baseiam na probabilidade de falha como métrica para representar a não confiabilidade, porém, essa abordagem não leva em conta os fatores que afetam a percepção do usuário sobre a qualidade, em especial, a confiabilidade do software. Em um estudo conduzido por Yabe et al. (2018), foram identificados

vários fatores que influenciam a percepção do usuário sobre a confiabilidade, como a consequência da falha, a relação com prazos apertados, formato da mensagem de erro, possibilidade de recuperação da falha, expectativa sobre a solução, tipo e frequência da falha, e familiaridade com o evento de falha. Esses resultados mostram que outros fatores devem ser considerados na análise da confiabilidade percebida pelos usuários.

A avaliação da experiência do usuário é geralmente baseada em fatores emocionais e é obtida por meio de perguntas diretas aos usuários. A média das respostas dos usuários, conhecida como *Mean Opinion Score* - (MOS), é uma métrica de QoE que pode ser usada para avaliar a qualidade de diferentes aspectos do produto ou serviço, como qualidade de áudio e vídeo. A QoE é uma métrica que leva em consideração diversos fatores, como contexto do usuário, satisfação percebida e aspectos emocionais. Embora a MOS inclua os fatores identificados por Yabe et al. (2018), ela não permite identificar diretamente o fator qualitativo mais relevante na avaliação. O estudo desses fatores pode fornecer subsídios para o desenvolvimento de novas métricas de QoE que possam ser aplicadas na avaliação da confiabilidade percebida pelos usuários.

A qualidade é um tema abordado por diversas áreas do conhecimento e pode ser entendida a partir de diferentes perspectivas. Normalmente, a qualidade é analisada pelos atributos e características, pela identificação de variações no processo de produção e pela comparação de experiências. No caso da avaliação da qualidade de um serviço prestado por um software, é crucial considerar a percepção do usuário, suas expectativas e as situações em que há uma quebra das promessas durante a interação com o produto. A avaliação da qualidade de serviços é realizada de várias maneiras, e muitas dessas formas de avaliação são baseadas no estudo de Parasuraman, Zeithaml e Berry (1988) que fundamentou a técnica de desconformação, que confronta a relação entre expectativa e percepção para determinar a qualidade de um serviço. Apesar da técnica ter originalmente sido criada para serviços prestados por humanos, ainda há poucos casos na literatura em que a mesma foi adaptada para avaliações de softwares.

A natureza do software como um produto que oferece um serviço aos usuários permite a aplicação de técnicas clássicas de avaliação de qualidade de serviço. Zeithaml, Bitner e Gremler (2017) argumentam que “tudo é um serviço” e citam a Apple e a Samsung como exemplos de empresas que incorporam serviços a seus produtos, como smartphones e tablets, oferecendo soluções que são consideradas essenciais no dia a dia dos usuários. Essa noção de produto e serviço é crucial ao relacionar duas áreas diferentes do conhecimento que têm o mesmo objetivo: criar produtos que satisfaçam as necessidades dos usuários.

A pesquisa de embasamento teórico deste trabalho revela que a “Zona de Tolerância,” conforme definida por Parasuraman, Zeithaml e Berry (1988), sustenta a ideia de que um produto é de qualidade quando é considerado aceitável, ou seja, quando atende aos requisitos mínimos, padrões ou expectativas estabelecidos, embora essa aceitação possa variar dependendo do contexto. Em outras palavras, tudo aquilo que é considerado de

qualidade, é aceitável quando está dentro dos limites razoáveis ou toleráveis. Além disso, relatos de Lovelock, Wirtz e Hemzo (2011), Zeithaml, Bitner e Gremler (2017) e as demais fontes relevantes deste trabalho foram empregados como base para desdobrar e categorizar os fatores que contribuem para a formação do conceito de tolerância.

Na Figura 1, apresenta-se uma árvore hierárquica de fatores, que representa de forma estruturada os elementos relevantes relacionados à análise dos fatores relacionados com a tolerância do usuário no que tange a relação com falhas de software, especialmente no contexto em que o software oferece serviços aos seus usuários. O propósito dessa representação visual é possibilitar a categorização e avaliação das interações entre esses fatores, bem como a visualização organizada de sua importância em relação aos demais fatores. No contexto da pesquisa, o diagrama ilustra as variáveis independentes, representadas como fatores, e as variáveis dependentes, que são os resultados esperados.

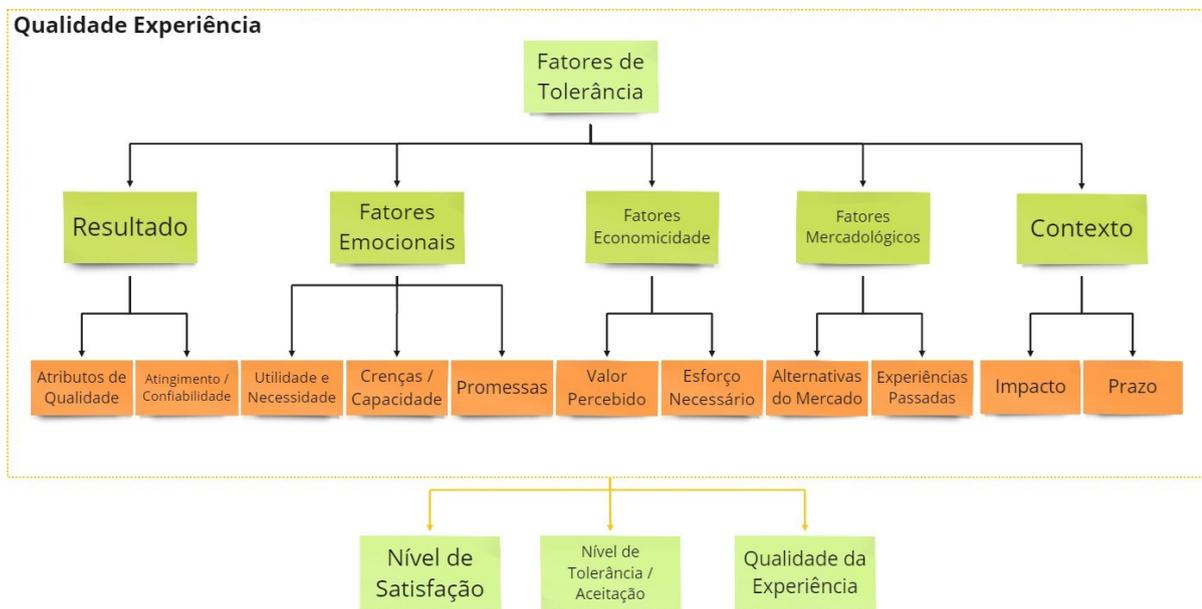


Figura 1 – Fatores relacionados à tolerância

Neste trabalho, avaliamos variáveis relacionadas com a qualidade da experiência (QoE), satisfação, tolerância e aceitação dos usuários em relação ao serviço proporcionado por um software. Usamos uma escala de zero a dez para obter a média da opinião dos usuários com base nas funções do software que experimentaram. A QoE foi medida perguntando aos usuários sobre a qualidade geral de sua experiência, que variou de Muito Negativa a Muito Positiva. A satisfação foi avaliada perguntando aos usuários o quanto estavam satisfeitos com o resultado do software, variando de Muito Insatisfeito(a) a Muito Satisfeito(a). Quanto à tolerância, perguntamos aos usuários o quanto o software atendeu às suas necessidades, variando de Totalmente Inaceitável a Totalmente Aceitável. Pesquisas anteriores mostraram que a experiência do usuário pode ser avaliada por meio da média de suas opiniões coletadas após o uso do aplicativo. Além disso, a tolerância pode ser

obtida usando o modelo de desconfirmação, assumindo que a expectativas de falha é indesejada em qualquer nível, que compara as expectativas dos usuários com a sua percepção em cenários onde ocorrem variações ou impactos na entrega do serviço fornecido pelo software.

A seguir, serão apresentados os grupos e fatores considerados pelos usuários ao lidar com situações de falhas descritos na Figura 1. Em relação aos resultados, quando um serviço cumpre sua proposta, seja alcançando objetivos ou apresentando características como usabilidade, responsividade, segurança e integridade, os usuários tendem a considerá-lo favoravelmente. Em termos emocionais, os fatores estão relacionados às expectativas de benefícios que o serviço pode proporcionar ou à prevenção de sentimentos negativos, como a frustração. Nesse contexto, também há uma influência das crenças limitantes, da capacidade cognitiva necessária e da expectativa de satisfação de necessidades básicas. Os fatores de economicidade estão ligados ao valor percebido do serviço em relação ao que o usuário considera justo, bem como ao esforço pessoal exigido para concluir as tarefas relacionadas à utilização do serviço. Fatores mercadológicos também desempenham um papel, pois os usuários podem ter expectativas com base em serviços similares que já experimentaram, e sua tolerância pode variar dependendo das alternativas disponíveis no mercado. O contexto em que o serviço é utilizado também é um fator relevante, uma vez que uma falha pode ter consequências de impacto baixo ou catastrófico, de acordo com a situação. Além disso, o prazo para conclusão de uma tarefa é outro fator que influencia a tolerância, pois está diretamente relacionado à ansiedade dos usuários e a impactos mais severos.

A avaliação da Experiência do Usuário é um assunto pouco explorado, pelo menos no que tange aos fatores relacionados à sua percepção diante de possíveis falhas, o que proporciona uma compreensão abrangente do impacto dessas falhas na satisfação do usuário e na qualidade do serviço. O termo Tolerância refere-se à capacidade de lidar com desafios, adversidades, erros ou falhas de forma equilibrada, paciente e compreensiva. É a disposição para aceitar, suportar ou lidar de maneira positiva com eventos ou circunstâncias que possam causar desconforto, incerteza ou insatisfação. O termo Aceitável é usado para descrever algo que atende aos requisitos mínimos, padrões ou expectativas estabelecidos, embora possa variar dependendo do contexto. Ou seja, aceitável descreve algo que está dentro dos limites aceitos, razoáveis ou toleráveis. No contexto de um serviço realizado por alguém, quando dizemos que o trabalho foi aceitável, significa que atendeu aos requisitos básicos ou às expectativas mínimas estabelecidas para essa tarefa. Já Consequência refere-se aos resultados ou efeitos que ocorrem como resultado de uma ação, evento ou decisão. É uma noção ampla que abrange os desdobramentos ou resultados diretos e indiretos. As consequências podem ser positivas (benefícios) ou negativas (prejuízos) na percepção do usuário. Já o termo Impacto refere-se ao efeito significativo ou alteração notável resultante de uma ação, evento ou decisão. O impacto está associado a mudanças

substanciais, influência perceptível ou repercussões importantes, e indica a magnitude ou intensidade dos efeitos produzidos.

1.3 Objetivos

1.3.1 Objetivo geral

O objetivo desta pesquisa é explorar a inter-relação entre os fatores que impactam a percepção dos usuários sobre as falhas de software, identificar os fatores determinantes da tolerância a essas falhas em diferentes cenários, e analisar a relação entre as variáveis de aceitabilidade, satisfação e qualidade percebida da experiência.

1.3.2 Objetivos específicos

De maneira mais detalhada, este estudo pretende:

- ❑ Identificar os principais fatores que influenciam a tolerância dos usuários em relação a falhas de software;
- ❑ Avaliar a experiência percebida em um cenário de falha, em uma funcionalidade de software, avaliando aspectos da tolerância e visando compreender como esses fatores podem ser utilizados para minimizar os efeitos negativos causados pelas falhas, possibilitando assim o aprimoramento da experiência do usuário;
- ❑ Estabelecer uma relação entre a tolerância e a QoE baseada na satisfação, através dos principais fatores relacionados à tolerância.

1.4 Perguntas de Pesquisa e Hipóteses

Q1: **Pergunta:** O nível de tolerância a falhas de software é uma medida observável de forma sistemática entre os usuários de software?

Hipótese: Conjectura-se que seja possível medir de maneira sistemática o nível de tolerância que os usuários de software têm em relação a falhas. Em outras palavras, busca-se investigar se é possível quantificar e observar de forma objetiva o grau de tolerância dos usuários em relação ao mau funcionamento encontrados em um software.

Q2: **Pergunta:** É possível estabelecer um valor padrão (exemplo: média/mediana) para o nível de tolerância, abaixo do qual a QoE não é afetada (ou acima do qual a QoE é afetada) negativamente de forma significativa?

Hipótese: Conjectura-se que seja possível estabelecer um valor de referência para o nível de tolerância, no qual, abaixo desse valor, a QoE não seria afetada de forma significativa, e acima desse valor, a QoE seria afetada negativamente de maneira significativa.

Q3: Pergunta: Quais os principais fatores (humanos, técnicos, economicidade, etc.) associados ao nível de tolerância a falhas de software observado em usuários? Como são essas associações?

Hipótese: Conjectura-se que categorias de fatores, como fatores humanos, técnicos, economicidade entre outros, podem influenciar a forma como os usuários lidam com as falhas, mostrando a existência de correlações que implicam o nível de tolerância a falhas de software. Por exemplo, pode-se analisar se fatores como a familiaridade com a tecnologia, o conhecimento técnico, a importância da tarefa realizada no software, o custo envolvido ou a disponibilidade de alternativas impactam a tolerância das pessoas em relação a falhas.

Q4: Pergunta: Qual o tipo de relação entre o nível de tolerância a falhas de software e a satisfação?

Hipótese: Conjectura-se investigar a existência de uma correlação entre o nível de tolerância a falhas e a satisfação geral do usuário com a experiência de uso do software. Ou seja, a QoE medida por sua aproximação com relação à satisfação. Dessa forma, é possível avaliar se os usuários que possuem uma maior tolerância a falhas tendem a ter uma satisfação mais elevada ou se há uma relação inversa, em que usuários com baixa tolerância a falhas tendem a ter menor satisfação.

1.5 Metodologia

A metodologia proposta para esta pesquisa envolve a coleta de dados primários por meio de um questionário online destinado a participantes anônimos. A elaboração da metodologia de pesquisa foi alinhada aos requisitos estabelecidos pelo CEP/CONEP, incorporando os pontos essenciais para a construção do Termo de Consentimento Livre e Esclarecido (TCLE) conforme as diretrizes especificadas pelo CEP/UFU. O questionário foi desenvolvido para capturar as opiniões dos participantes diante de cenários que envolvem falhas ao utilizar o software em tarefas comuns realizadas em computadores pessoais. Ele abrangerá diferentes contextos e incluirá perguntas destinadas a explorar:

- O contexto de uso do software;
- A gravidade das falhas;
- As expectativas do usuário sobre o software;
- A tolerância do usuário a falhas;
- A satisfação geral do usuário com a experiência de uso do software.

A revisão da literatura deste trabalho identificou os principais fatores relacionados ao conceito de tolerância, conforme descrito na Figura 1. Para avaliar os fatores relacionados à tolerância do usuário a falhas, serão apresentados aos participantes uma série de cenários de falha realistas. Posteriormente, por meio do questionário disponível na subseção 4.2.4, serão coletadas as opiniões e percepções dos participantes em relação ao efeito percebido da falha. Os cenários foram elaborados considerando diferentes níveis de fatores relacionados às falhas conforme referenciado na Figura 1.

A formulação do experimento foi fundamentada na metodologia aplicada no estudo de Tversky e Kahneman (1981), que destacou a influência do enquadramento das opções na tomada de decisões humanas. Esse tipo de pesquisa explora como as escolhas são moldadas pela apresentação das informações e como as pessoas avaliam alternativas em contextos diversos. Os autores conduziram uma série de experimentos mentais para investigar essa dinâmica. A metodologia empregada nos experimentos consistiu na criação de cenários fictícios, nos quais os participantes enfrentaram escolhas entre opções idênticas em termos de utilidade objetiva, mas apresentadas de maneiras diferentes. Assim como nos cenários desenvolvidos por Tversky e Kahneman (1981) que abrangeram várias situações, incluindo questões médicas e financeiras, os participantes são convidados a realizar escolhas baseadas em enquadramentos, que em casos positivos, enfatizavam vidas salvas, em comparação com enquadramentos negativos que destacavam vidas perdidas.

Os resultados desses experimentos evidenciaram a suscetibilidade das escolhas humanas ao viés de enquadramento (TVERSKY; KAHNEMAN, 1981). Isso implica que as

decisões são influenciadas não somente pela utilidade objetiva das opções, mas também pela maneira como estas são apresentadas. A conclusão do artigo sublinha a importância desse viés na compreensão das decisões e destaca seu impacto em campos como economia comportamental, saúde e política. Em síntese, o estudo de Tversky e Kahneman (1981) demonstra que a forma como as alternativas são apresentadas desempenha um papel fundamental nas escolhas humanas. Além disso, ele ressalta a utilidade dos experimentos mentais como ferramentas valiosas para compreender como as pessoas reagem a diferentes enquadramentos, o que realça a complexidade da tomada de decisões e a relevância da psicologia cognitiva nesse âmbito.

Nesta pesquisa, será empregada uma combinação de experimento mental e a técnica de pesquisa com múltiplos cenários, também conhecida como pesquisa com manipulação de cenários. O conceito de experimento mental tem suas raízes na filosofia e foi amplamente utilizado por filósofos como Galileu Galilei no século XVII e mais tarde por René Descartes. Essa abordagem se mostra eficaz para coletar informações detalhadas sobre as preferências ou percepções dos respondentes, no que tange a diferentes condições ou contextos de uso de software. No âmbito deste estudo específico, que visa compreender a percepção da tolerância, frente ao impacto das consequências das falhas de software, será utilizado um único questionário para apresentar aos respondentes uma série de cenários variados em termos de combinações de fatores relacionados a tolerância. Cada cenário representa uma combinação única, introduzindo um novo elemento.

- Cenário 1: O usuário consegue editar o documento normalmente;
- Cenário 2: Ocorre uma falha ao editar o texto e aplicativo reinicializa;
- Cenário 3: Ressalta as promessas de uso do aplicativo;
- Cenário 4: Ressalta a perda do prazo para envio;
- Cenário 5: Ressalta as opções concorrentes do mercado;
- Cenário 6: Ressalta a atualização do aplicativo e o esforço em melhorar as falhas;
- Cenário 7: Ressalta avaliação do usuário frente à reincidência do mesmo problema.

Os respondentes serão solicitados a avaliar suas preferências em relação a cada cenário, utilizando uma escala de avaliação pertinente às variáveis dependentes em análise. Essa estratégia permite coletar dados visando à avaliação da percepção dos respondentes em diversas condições experimentais, eliminando a necessidade de aplicar questionários separados para cada cenário.

A escala Triádica de Likert será utilizada para medir o conceito de tolerância, enquanto a satisfação geral, qualidade da experiência e aceitabilidade dos usuários em relação ao cenário apresentado serão avaliadas por meio de uma escala de opinião de 0 a 10.

A avaliação desse conceito será conduzida por meio da técnica da escala triádica de Likert, também conhecida como escala de Likert de três pontos. Essa abordagem é amplamente empregada como uma técnica de medição em pesquisas sociais e comportamentais. Introduzida por Likert (1932), a escala foi desenvolvida para capturar opiniões dos respondentes de maneira abrangente. A escala triádica de Likert é utilizada para avaliar atitudes, percepções e preferências dos participantes. Cada fator em análise será acompanhado por três afirmações: uma positiva, uma negativa e uma neutra. Essa técnica possibilita uma avaliação mais equilibrada das opiniões dos respondentes, ao oferecer uma opção neutra que permite expressar incerteza ou falta de conhecimento em relação ao fator em análise. A escala triádica de Likert é comumente empregada em pesquisas de satisfação do cliente, avaliação de produtos, estudos de opinião pública e outras áreas de pesquisa social. Essa abordagem demonstra ser flexível e eficaz na obtenção de percepções e atitudes dos participantes.

O nível de tolerância em cada cenário será avaliado por meio de perguntas diretas feitas aos usuários, que contribuirão para a consolidação da Média de Opinião do Usuário (MOS). Conforme os cenários avançam, ocorre uma variação incremental de fatores, e à medida que os participantes são imersos nos cenários propostos na experiência, as variáveis de pesquisa são coletadas por meio de questionários. A expectativa do usuário é representada pelo cenário de sucesso. Assim, a diferença entre os cenários é empregada para comparar as variações entre eles. A MOS é utilizada para medir a opinião média em relação à adequação à necessidade de uso, à satisfação e à tolerância. Essas medidas compõem as observações da QoE. A seguir, todos os cenários propostos e as variáveis coletadas em cada etapa estão descritos:

O objetivo do desfecho primário é identificar os fatores que influenciam a tolerância dos usuários a falhas de software em diferentes cenários. Para isso, a pesquisa medirá a tolerância dos usuários por meio de um questionário online. O questionário incluirá cenários com tipos de falhas de software comumente encontradas e buscará compreender como os participantes reagiram às falhas de software e o que os participantes acreditam que influencia a aceitabilidade dos usuários a falhas de software. Esse desfecho é relevante para os objetivos da pesquisa, pois fornece conhecimentos importantes sobre os fatores que influenciam a tolerância dos usuários a falhas de software.

O desfecho secundário busca analisar a relação entre a tolerância dos usuários a falhas de software e a satisfação do usuário. Será utilizada uma escala de Likert para medir a tolerância, e os dados coletados serão analisados para identificar possíveis relações entre os fatores ligados à Qualidade da Experiência (QoE) e a satisfação do usuário em cenários de falhas. Essas informações poderão ser aplicadas para aprimorar a experiência do usuário e a qualidade dos serviços oferecidos por meio de softwares mais confiáveis.

É esperado que a amostra de participantes seja diversificada, abrangendo usuários com diferentes níveis de experiência no uso de softwares. Essa abordagem visa assegurar

a representatividade da população-alvo e será monitorada por meio da seção de perfil no questionário.

Os dados coletados serão submetidos a análises estatísticas para identificar padrões e relações significativas da tolerância, incluindo a relevância dos principais fatores que influenciam a tolerância do usuário e sua relação com a satisfação. Os resultados serão apresentados em relatórios detalhados, oferecendo implicações práticas para o desenvolvimento de software, visando aprimorar a Qualidade da Experiência (QoE) e considerar os fatores relacionados à tolerância do usuário às falhas.

A determinação da quantidade de amostras necessárias considerou um efeito moderado de 0,35, um nível de confiança de 95%, poder de teste de 80% e sete cenários, o tamanho amostral resultou em um total de 86 sujeitos de pesquisa.

Referencial teórico

2.1 Introdução

A confiabilidade tem sua base na ciência, onde um experimento é considerado confiável quando pode ser repetido e produzir o mesmo resultado. O estudo da confiabilidade se desenvolveu após a Segunda Guerra Mundial, impulsionado pela necessidade do exército americano de adquirir equipamentos com qualidade aceitável e a capacidade de prever a taxa de falhas ao longo do ciclo de vida. Um conceito amplamente utilizado nesse contexto é conhecido como a ‘curva da banheira’ e consiste em uma função estatística que divide o ciclo de vida do produto em três fases distintas: uma fase inicial com taxas de falhas variadas, uma fase estável com baixa taxa de falhas e uma fase final com aumento das falhas, gerando um gráfico em forma de banheira. Essa curva tem sido efetivamente aplicada na previsão da confiabilidade de diversos produtos. Para aprimorar a qualidade, foram desenvolvidos vários processos industriais voltados para a melhoria contínua na fabricação e manutenção de equipamentos, com o uso da metrificação das falhas. No entanto, o ciclo de vida de um produto de software é caracterizado por múltiplas versões consecutivas a cada incremento, o que resulta em variações e comportamentos diferentes dos modelos tradicionais. O software possui particularidades únicas que diferem da confiabilidade básica. Nos próximos capítulos, serão apresentados os fundamentos da engenharia de confiabilidade, juntamente com as principais técnicas e termos utilizados na engenharia de software.

Um processo de melhoria eficaz requer a utilização de métricas como base para estabelecer pontos de monitoramento e controle. Tradicionalmente, a curva de confiabilidade tem sido uma métrica empregada para avaliar a qualidade de produtos e serviços por meio do registro e análise das falhas observadas. No entanto, no caso específico do software, a qualidade é considerada quando ele é capaz de satisfazer de forma eficiente e confiável as necessidades dos usuários dentro de prazos e custos aceitáveis. A literatura apresenta uma variedade de atributos de qualidade relevantes, incluindo segurança, compreensibilidade, portabilidade, proteção, testabilidade, usabilidade, confiabilidade, adaptabilidade,

reusabilidade, resiliência, modularidade, eficiência, robustez, complexidade e facilidade de aprendizado. A metrificação da qualidade do software enfrenta desafios, uma vez que muitos desses atributos não podem ser diretamente extraídos da observação do produto, pois estão intrinsecamente ligados à percepção dos usuários. A garantia da satisfação das necessidades dos usuários é uma combinação de fatores externos que se relacionam à experiência dos usuários ao utilizar o software. Nos próximos capítulos, serão apresentados fundamentos que enfocam a qualidade da relação entre os usuários e o sistema, conhecida como experiência do usuário, e como a qualidade pode ser avaliada por meio de métricas que se concentram na percepção dos usuários.

2.2 Fundamentos

2.2.1 Confiabilidade de software

A engenharia de confiabilidade é uma disciplina de grande interesse na indústria, cujo objetivo é melhorar a confiabilidade de produtos e processos por meio da aplicação de técnicas eficazes (LYU, 2007). Ela abrange uma variedade de atividades, incluindo o projeto, desenvolvimento, teste, operação e manutenção de sistemas confiáveis, com ênfase na prevenção, detecção e correção de falhas. De acordo com Laprie (1995), a confiabilidade de software é definida como “a capacidade de um software executar corretamente suas funções especificadas sob condições operacionais determinadas, durante um período de tempo definido”. Essa definição é semelhante à fornecida pelo IEEE (1990), que descreve a confiabilidade como “a capacidade de um sistema ou componente executar suas funções necessárias sob condições estabelecidas durante um período de tempo especificado”, e que é amplamente utilizado. Além disso, mais recentemente, Avizienis et al. (2004), complementam de forma mais objetiva que a confiabilidade está relacionada à continuidade da prestação do serviço correto. Lyu (2007) destaca que, entre os atributos de qualidade de software, a confiabilidade é o principal fator, uma vez que quantifica as falhas do software que podem comprometer o funcionamento do sistema. Uma falha no software pode impedir que os usuários utilizem as funções projetadas para atender às suas necessidades, prejudicando assim a experiência do usuário.

Pham (2007) cita o IEEE (1990) e explica que quando um programador ou projetista comete um equívoco, lapso, uma falta é introduzida no código. Comumente, a falta é chamada de defeito, bug ou erro de forma generalizada, mas seu uso mais específico é preferível. SWEBOOK et al. (2014) destacam que a causa de um mau funcionamento é considerada uma falta. O erro é visto como o mau funcionamento do sistema (IEEE, 1990). De acordo com Avizienis et al. (2004), o serviço prestado pelo sistema é uma sequência de estados externos, e uma falha de serviço ocorre quando pelo menos um desses estados externos se desvia do estado correto. Avizienis et al. (2004) também enfatizam que essa

divergência é chamada de erro, e a causa presumida ou hipotética de um erro é chamada de falta. Laprie (1995) explica que as faltas são dormentes até o momento de sua ativação, o que resulta em um erro. O erro tem a característica de se propagar, podendo causar outros erros. Quando esses erros ultrapassam a fronteira da aplicação, ocorre uma falha. Em um cenário em que o resultado de um componente de uma aplicação é a entrada de um componente subsequente, a ativação de uma nova falta pode ocorrer, gerando assim um novo erro (LAPRIE, 1995). Esse fenômeno é definido por Avizienis et al. (2004) como a cadeia fundamental da dependabilidade.

A definição de um serviço correto representa um desafio na engenharia de confiabilidade de sistemas, pois a presença de falhas relacionadas à especificação ainda é uma possibilidade. Uma especificação incorreta pode resultar em um software que não atende às necessidades do usuário, o que é uma preocupação central nesta área. Laprie (1995) destaca que um serviço correto é aquele que implementa a função do sistema, enquanto uma falha ocorre quando o sistema não está em conformidade com a especificação ou quando esta não descreve adequadamente a função esperada. Quando falhas na especificação surgem, o serviço entregue pelo sistema é questionado pelos usuários (AVIZIENIS et al., 2004). Avizienis et al. (2004) ressaltam que as falhas relacionadas à especificação frequentemente envolvem omissões ou erros, como interpretações equivocadas, suposições injustificadas, inconsistências e erros tipográficos. Nessas circunstâncias, a identificação de uma falha como um evento indesejado só pode ocorrer após sua ocorrência e por meio de suas consequências. Isso evidencia que as falhas podem ser subjetivas e discutíveis, exigindo julgamento para sua identificação e caracterização, segundo Avizienis et al. (2004). Assim, garantir um serviço correto na busca pela confiabilidade dos sistemas envolve aprimoramentos na elaboração das especificações, visando atender adequadamente às expectativas dos usuários.

Existem outros fatores identificados como causas de falhas de desenvolvimento que afetam a confiabilidade dos sistemas, e muitos estão relacionados com aspectos da engenharia de software, como a gestão de custos (tanto para operação quanto para projetos) e o prazo necessário para executá-los. Avizienis et al. (2004) enfatizam que especificações incompletas ou defeituosas, um grande número de alterações solicitadas pelos usuários, design inadequado em relação à funcionalidade e metas de desempenho, faltas recorrentes durante o desenvolvimento, capacidade insuficiente de remoção de erros, previsão inadequada de confiabilidade ou segurança e estimativas incorretas dos custos de desenvolvimento são fatores-chave para a baixa confiabilidade. Segundo Avizienis et al. (2004), esses problemas muitas vezes surgem devido a uma subestimação da complexidade do sistema a ser desenvolvido e destacam a importância de abordar técnicas eficientes para enfrentar os desafios no desenvolvimento de sistemas confiáveis.

Como a busca pela melhoria da qualidade de software é um dos principais objetivos da engenharia de confiabilidade e conforme apontado por SWEBOK et al. (2014), a definição

de qualidade de software passou por uma evolução ao longo do tempo. Antigamente, a qualidade era associada à “conformidade com os requisitos” estabelecidos. No entanto, atualmente, é entendida como a “capacidade do produto de software de satisfazer as necessidades explícitas e implícitas sob certas condições”. Em outras palavras, a qualidade refere-se ao grau em que um produto de software atende aos requisitos estabelecidos, desde que estes representem com precisão as necessidades, desejos e expectativas das partes interessadas. SWEBOK et al. (2014) ressaltam que ambas as definições enfatizam a conformidade com os requisitos iniciais. Além disso, a engenharia de software considera a qualidade do software como um esforço para maximizar o valor entregue às partes interessadas, equilibrando as restrições de custo e cronograma. Essas perspectivas destacam a importância de adotar técnicas para garantir a qualidade do software, que vão além da garantia da conformidade com os requisitos. Deming Edwards W., um dos autores mais referenciados sobre o tema da melhoria contínua da qualidade, também resalta que zero defeitos não é sinônimo de um produto de qualidade, reconhecendo que a qualidade é um fator muito mais complexo. Portanto, é necessário compreender e atender às necessidades e expectativas das partes interessadas, buscando maximizar o valor entregue pelo produto de software, levando em consideração as restrições impostas ao desenvolvimento da solução.

Pressman (2010) destaca a crescente demanda por maior qualidade de software na comunidade, devido à sua integração cada vez maior em todas as atividades humanas. O autor resalta que bilhões de dólares são desperdiçados anualmente em software que não atende às características e funcionalidades prometidas. Segundo Pressman (2010), a qualidade é um conceito complexo e multifacetado, que pode ser descrito a partir de diferentes perspectivas. Pressman (2010) menciona a visão do usuário, a visão do fabricante e a visão do produto como exemplos dessas perspectivas. A visão baseada no usuário considera a qualidade do software com base na satisfação do cliente e nas características inerentes ao produto. Já a visão baseada no fabricante enfatiza a entrega do software dentro do orçamento e prazo previstos. Além disso, cita-se a visão de DeMarco e Lister (2013), que defendem que a qualidade de um produto de software está relacionada à sua capacidade de transformar positivamente o mundo, mesmo que ocasionalmente este apresente problemas de confiabilidade ou desempenho. Pressman (2010) discute a definição de qualidade de software como forma de viabilizar uma gestão efetiva da qualidade, com o objetivo de criar um produto útil e de valor mensurável para os envolvidos na sua produção e uso.

A comunidade de engenharia de software enfrenta desafios ao tentar desenvolver medidas precisas para a qualidade de software, muitas vezes frustrada pela natureza subjetiva da atividade. Essa busca contínua por medidas e padrões de qualidade reflete o compromisso da comunidade em aprimorar a confiabilidade e o desempenho do software, a fim de atender às expectativas dos usuários e proporcionar benefícios tangíveis. Desta forma,

vários modelos e propostas foram criados com este objetivo. Dentre os modelos mencionados por Pressman (2010), estão as dimensões de qualidade propostas por Garvin, os fatores de qualidade de McCall e os fatores de qualidade ISO 9126, nos quais a confiabilidade é considerada um pilar essencial. Em resumo, o autor ressalta a dificuldade, e às vezes a impossibilidade, de desenvolver medidas diretas para esses fatores de qualidade devido à complexidade intrínseca relacionada ao entendimento dos fatores originais que geram o impacto percebido como falhas.

Pressman (2010) destaca o dilema da qualidade discutido por Bertrand Meyer, que ressalta as dificuldades enfrentadas na busca pelo equilíbrio entre qualidade e viabilidade econômica no desenvolvimento de software. Meyer argumenta que se um sistema de software for produzido com péssima qualidade, ninguém o comprará, enquanto um software absolutamente perfeito exigiria um tempo e esforço infinitos, além de altos custos de produção. Portanto, é necessário encontrar um meio-termo que atenda às necessidades do mercado sem comprometer excessivamente tempo e recursos.

A característica do software de poder ser incrementado ao longo do seu ciclo de vida, traz uma importante vantagem em relação à confiabilidade tradicional, pois a cada incremento a confiabilidade pode ser melhorada continuamente. Pressman (2010) menciona que algumas empresas deliberadamente lançam software com erros conhecidos, reconhecendo que a primeira versão pode não ter a melhor qualidade. Elas planejam melhorias para versões subsequentes, cientes de que alguns clientes podem fazer reclamações. No entanto, elas entendem que colocar o produto no mercado o mais rápido possível é uma estratégia eficaz em termos de qualidade, desde que o software fornecido seja “suficientemente bom”. Isso significa oferecer funções e características de alta qualidade que os usuários desejam, mesmo que acompanhadas de outras mais específicas ou menos conhecidas, e ainda contendo erros conhecidos. O fornecedor de software espera que a grande maioria dos usuários ignore os erros devido à satisfação geral com as outras funcionalidades oferecidas pela aplicação. Essa abordagem reconhece a importância de entregar um produto funcional e útil dentro de um prazo razoável, considerando que a perfeição absoluta pode ser inatingível ou impraticável. Ao equilibrar a qualidade com a capacidade de resposta ao mercado, as empresas buscam atender às demandas dos usuários, enquanto também planejam melhorias contínuas e ajustes para versões futuras. Essa abordagem pragmática reflete a complexidade inerente à busca pela qualidade no desenvolvimento de software, considerando as restrições de tempo, recursos e as expectativas dos usuários.

2.2.2 Taxonomia

Diversos termos foram utilizados no desenvolvimento desta pesquisa, a fim de descrever conceitos relevantes para o trabalho. Abaixo, apresenta-se uma listagem desses termos, com suas respectivas definições:

- ❑ **Qualidade:** o grau em que um sistema, componente ou processo atende aos requisitos especificados ou às necessidades ou expectativas do cliente ou usuário;
- ❑ **Requisito:** uma condição ou capacidade implementada em um sistema ou parte dele, com o objetivo de resolver um problema ou alcançar um objetivo, satisfazendo um contrato, padrão, especificação ou outros documentos formalmente descritos;
- ❑ **Função:** um objetivo definido ou ação característica de um sistema ou componente;
- ❑ **Confiabilidade:** a capacidade de um sistema ou componente executar suas funções requeridas em condições estabelecidas, durante um período de tempo específico;
- ❑ **Falha:** a incapacidade de um sistema ou componente executar suas funções exigidas dentro dos requisitos de desempenho esperados;
- ❑ **Falha leve (*soft failure*):** uma falha que permite a operação contínua de um sistema com capacidade operacional parcial;
- ❑ **Tolerância a faltas (*fault tolerance*):** a capacidade de um sistema ou componente continuar operando normalmente, mesmo na presença de falhas;
- ❑ **Recuperação (*recovery*):** a restauração de um sistema, programa, banco de dados ou outro recurso do sistema a um estado no qual possa executar as funções necessárias;
- ❑ **Mascaramento de faltas (*fault masking*):** uma condição na qual uma falha é impedida de ser detectada por outros componentes;
- ❑ **Erro:** a manifestação de uma falta que causa uma diferença entre um valor calculado, observado ou medido, e o valor ou condição verdadeiro, especificado ou teoricamente correto;
- ❑ **Falta ou causa do erro:** um defeito físico em um equipamento ou outro componente dependente, ou um procedimento, passo ou dado estruturado incorretamente em um programa, também conhecido como defeito ou *bug*;
- ❑ **Injeção de erros (*error seeding*):** o processo usado para inserir intencionalmente faltas conhecidas em um componente, a fim de gerar erros que possam ser monitorados para avaliar a taxa de detecção e tratamento, e a resposta do sistema a diferentes modos de falhas;

- ❑ Correção ou corretude: o grau em que um sistema ou componente está livre de faltas em sua especificação, projeto e implementação, e o grau em que atende aos requisitos, necessidades e expectativas do usuário, especificados ou não.

2.2.3 Métricas de confiabilidade de software

Com o aumento da importância das funcionalidades de software, nossa dependência delas também cresce. Pham (2007) destaca que praticamente todas as pessoas ao redor do mundo são afetadas, direta ou indiretamente, pelos sistemas de computador. Uma única falha em um sistema pode ter impacto significativo, afetando milhões de pessoas. Nos últimos anos, os custos relacionados ao desenvolvimento de software e as penalidades por falhas se tornaram despesas consideráveis. As falhas de software podem resultar em perdas, danos ou até mesmo a destruição de ativos, além de representar riscos para a vida humana e acarretar altos custos monetários. Pham (2007) relata que maneiras de quantificação e previsão da confiabilidade de sistemas de computador em ambientes operacionais complexos têm sido aplicadas como forma de mitigar esses riscos. Estimar a confiabilidade do software é um aspecto essencial na engenharia de confiabilidade, buscando avaliar e prever a probabilidade de falhas durante a operação dos sistemas.

Ao contrário do hardware, o software apresenta características distintas, que influenciam a forma como sua confiabilidade é estimada. Enquanto as falhas em hardwares são frequentemente causadas por desgaste, fadiga ou problemas físicos, as falhas no software estão mais relacionadas a erros de programação, lógica incorreta, condições imprevistas ou interações complexas entre os componentes do sistema. Além disso, o software não se degrada ou se desgasta com o tempo, mas sua confiabilidade pode ser afetada pelas alterações realizadas durante a manutenção, como correção de defeitos, modificações de código e adaptação a diferentes ambientes. Portanto, as técnicas e abordagens utilizadas para estimar a confiabilidade do software devem levar em consideração essas diferenças e considerar as mudanças que ocorrem ao longo do tempo devido às atividades de manutenção e atualização. Entretanto, falhas do hardware podem ser propagadas para o software, podendo gerar novas falhas.

Pham (2007) destaca que, à medida que os aplicativos críticos aumentam em tamanho e complexidade, as falhas relacionadas à lógica no projeto de software tornam-se mais sutis e desafiadoras pois passam de falhas catastróficas para falhas mais sutis, como possíveis erros de lógica. Um aspecto relevante na análise de confiabilidade é a consideração da confiabilidade sistêmica, que envolve a decomposição do sistema de software em componentes individuais. Essa abordagem permite avaliar a confiabilidade de cada componente e entender sua contribuição para a confiabilidade global do sistema. Para isso, é necessário modelar as interações entre os componentes e avaliar o impacto das falhas individuais no desempenho geral do sistema.

Os aplicativos críticos estão se tornando cada vez maiores e mais complexos, sendo compostos por diversos componentes. Nesse contexto, Pham (2007) apresenta a análise de confiabilidade sistêmica como uma técnica utilizada para avaliar a confiabilidade desses sistemas de software. Essa abordagem consiste na decomposição do sistema em seus componentes individuais, permitindo a análise da confiabilidade de cada um e sua contribuição para a confiabilidade global do sistema. Para isso, é necessário modelar as interações entre os componentes e avaliar o impacto das falhas individuais no desempenho geral do sistema. Pham (2007) defende que com este processo é possível identificar áreas críticas e adotar medidas de melhoria para aumentar a confiabilidade do software e gerar um impacto positivo no sistema como um todo. O modelo em série considera que a falha de qualquer componente resulta em uma falha no sistema como um todo. Por outro lado, o modelo em paralelo permite que o sistema continue operando mesmo se um dos componentes falhar, oferecendo uma base para a implementação de redundâncias no sistema. Já o modelo paralelo-série é uma combinação de ambos, em que subsistemas operam em paralelo, sendo compostos por componentes idênticos em série. Esses modelos auxiliam na compreensão do comportamento e na identificação de estratégias adequadas para aumentar a confiabilidade dos sistemas de software.

Diversos modelos estatísticos têm sido propostos para estimar a confiabilidade de software, e a maioria deles se baseia na observação de falhas percebidas por monitoramento ou aquelas que foram reportadas. Esses modelos requerem uma quantidade significativa de dados de falhas para realizar previsões de confiabilidade precisas (PHAM, 2007). Para desenvolver um modelo confiável de confiabilidade de software, é essencial compreender como o software é produzido e testado, os tipos de erros e como eles são introduzidos, além de considerar fatores ambientais e outras variáveis relevantes para a aplicação em cenários reais. Pham (2007) destaca que, à medida que os aplicativos críticos se tornam maiores e mais complexos, as falhas relacionadas à lógica no projeto de software se tornam mais sutis e desafiadoras, pois diferentemente das falhas catastróficas os erros mais sutis são difíceis de serem detectados.

A estimativa de confiabilidade de software utiliza modelos estatísticos para quantificar a probabilidade de falhas ao longo do tempo. Esses modelos se baseiam em dados históricos de falhas, testes de confiabilidade e análise estática de código, entre outros. Eles permitem avaliar a confiabilidade esperada do software em condições específicas de operação. Pham (2007) propõe uma abordagem que modela a confiabilidade como uma função estatística, definindo-a como a probabilidade de sucesso do sistema em executar sua função pretendida dentro dos limites de projeto estabelecidos. A confiabilidade é a probabilidade de o produto ou componente operar corretamente ao longo de um período de tempo determinado, sem falhas, considerando as condições operacionais definidas. Matematicamente, a confiabilidade $R(t)$ é a probabilidade de sucesso do sistema no intervalo de tempo de 0 a t , enquanto a falta de confiabilidade $F(t)$ é a probabilidade de falha no

tempo t . Essas definições são fundamentais para a análise e modelagem da confiabilidade em sistemas de software (PHAM, 2007).

$$R(t) = P(T > t) \quad t \geq 0$$

$$F(t) = P(T \leq t) \quad t \geq 0$$

A modelagem da confiabilidade de software faz uso de diferentes distribuições estatísticas que desempenham um papel importante. Distribuições como a exponencial, Weibull e Poisson são aplicadas para representar o comportamento das falhas em diferentes estágios do ciclo de vida do software. Essas distribuições auxiliam na identificação de padrões de falha e na estimativa de parâmetros relevantes, como taxas de falha e tempo médio entre falhas (MTBF). A função de distribuição do tempo de falha $F(t)$ é utilizada quando a variável aleatória $F(t)$ representa o tempo até a falha, permitindo a previsão da probabilidade de ocorrência de falhas. É importante ressaltar que a confiabilidade é uma função do tempo de missão (por exemplo, de 24 horas) e perde seu significado quando o tempo não é considerado. O tempo médio até a falha (MTTF) representa o tempo esperado para a ocorrência de uma falha enquanto um componente está em operação com sucesso, sendo derivado da taxa de falhas observadas. A disponibilidade é uma medida utilizada em sistemas reparáveis, que considera tanto o tempo de operação quanto o tempo de reparo. No entanto, em sistemas não reparáveis, a disponibilidade é igual à confiabilidade, pois não há reparo. O tempo médio entre falhas (MTBF) é uma medida importante em sistemas reparáveis, levando em consideração o tempo gasto com o reparo. Nos sistemas não reparáveis, em que o tempo de reparo tende a zero, apenas o MTTF é considerado.

Diversos modelos analíticos têm sido propostos e estudados para avaliar a confiabilidade de sistemas de software. Pham (2007) relata que esses modelos são fundamentados em suposições sobre o processo de desenvolvimento, ambiente de teste, ambiente de operação e ciclo de vida do sistema. É essencial que usuários e profissionais de software estejam familiarizados com essas variáveis relevantes para construir modelos confiáveis e tomar decisões embasadas sobre a confiabilidade (PHAM, 2007). Um aspecto crucial na estimativa de confiabilidade de software é o perfil operacional do sistema. O perfil operacional descreve as condições de uso do software, como carga de trabalho, volume de dados, frequência de operação e ambiente de execução. Essas informações são indispensáveis para determinar a confiabilidade, uma vez que o comportamento do software pode variar dependendo das condições em que é utilizado. Consequentemente, compreender e considerar o perfil operacional é fundamental para obter estimativas confiáveis e precisas da confiabilidade do sistema de software.

Uma vez que essas premissas são compreendidas e modeladas, técnicas estatísticas são empregadas para calcular a probabilidade de sucesso ou falha. O teste de aderência é uti-

lizado para determinar qual das distribuições teóricas é mais adequada, sendo exemplos disso os testes de Goodness-of-Fit e Kolmogorov-Smirnov, entre outros. Pham (2007) apresenta diversas distribuições frequentemente utilizadas na modelagem estatística. A distribuição binomial é aplicada em situações em que um evento pode resultar em sucesso ou falha, sendo uma distribuição discreta. A distribuição de Poisson, similar à distribuição binomial, é usada quando o tamanho da amostra é desconhecido. A distribuição exponencial é ideal para modelar taxas de falhas constantes e é empregada na análise do tempo de vida de componentes e sistemas eletrônicos e elétricos. A distribuição normal, devido ao Teorema do Limite Central, é importante na estatística clássica e é utilizada em medições de suscetibilidade do produto e estresse externo, como desgaste. A distribuição lognormal é um modelo flexível que se ajusta a diversos tipos de dados de falha e permite modelar incertezas nas informações de taxa de falha, sendo adequada para sistemas reparáveis. A distribuição Weibull é uma generalização da distribuição exponencial e é comumente usada para representar a vida útil e a fadiga de componentes. Sua flexibilidade permite modelar funções de taxa de risco variáveis e se aplica a diversas aplicações. Além das distribuições, as cadeias de Markov são utilizadas na construção de modelos estocásticos para descrever a operação de um sistema ao longo do tempo.

O teste de software é um processo crucial para identificar erros em um programa. No entanto, é impossível testar todas as combinações de entradas e caminhos possíveis, uma vez que isso exigiria um tempo impraticável, devido à sua enorme quantidade de possibilidades (PHAM, 2007). Durante a fase operacional, certos estados de entrada são mais frequentemente executados do que outros, e é possível atribuir uma probabilidade a cada estado de entrada para formar o perfil operacional do programa. Esse perfil operacional é utilizado para construir o modelo de confiabilidade de software. Um modelo de crescimento de confiabilidade é utilizado para estimar o nível atual de confiabilidade e determinar o tempo e os recursos necessários para atingir a meta desejada de confiabilidade. Após o lançamento do software, fatores como adição de novos módulos, remoção de módulos antigos, correção de erros identificados, combinação de código novo e existente, alteração do ambiente do usuário, mudanças de hardware e envolvimento do gerenciamento de software devem ser considerados na avaliação da confiabilidade do software. Esses aspectos dinâmicos são fundamentais para garantir que a confiabilidade do software seja constantemente monitorada e melhorada ao longo do tempo.

As estimativas de confiabilidade de software consideram a diferença entre hardware e software, utilizando modelos estatísticos, distribuições, análise de componentes e informações do perfil operacional para quantificar a probabilidade de falhas em sistemas de software e medir e prever a confiabilidade desejada. Para avaliar a confiabilidade, é necessário ter dados de falhas de software e assim produzir estimativas sobre os erros remanescentes. O usuário percebe uma falha quando o programa não atende aos requisitos ou às expectativas quanto à operação do software, e isso pode ser devido a diferentes

critérios, que nem sempre são bem definidos. A verificação e a validação são fases importantes para a construção de um modelo de confiabilidade, pois nessa etapa do projeto o que se questiona é se o produto está sendo construído corretamente e se é o produto certo. Ambas as etapas são essenciais para a elaboração de um modelo relevante, que guiará o desenvolvimento de um produto confiável.

2.2.4 Qualidade de serviço

A qualidade de serviços é um tema amplamente estudado e documentado na literatura. Lovelock, Wirtz e Hemzo (2011) e Zeithaml, Bitner e Gremler (2017) são autores de referência que contribuíram significativamente para o campo do marketing de serviços, oferecendo insights valiosos sobre a qualidade de serviços. Por outro lado, Oliver (2010) é amplamente reconhecido por suas contribuições no estudo da satisfação do cliente. Essas três referências são consideradas fundamentais e fornecem uma base sólida de conhecimento para este trabalho. Apesar da colisão de termos em relação a métricas QoS, a qualidade de serviço, esta tem uma aplicação mais ampla. Embora grande parte dos estudos sobre qualidade de serviços tenha se concentrado na interação entre pessoas na prestação de serviços, a evolução da nossa sociedade tem resultado na substituição de pessoas por máquinas e softwares em certos serviços. Essa modernização tem desafiado os conceitos tradicionais e aberto novas perspectivas de análise e compreensão da qualidade de serviços, aspectos estes que serão abordados nesta pesquisa. Como o objetivo é avaliar a qualidade dos serviços prestados por software, os conceitos e *frameworks* propostos por Lovelock, Wirtz e Hemzo (2011), e Oliver (2010) serão utilizados como base teórica.

A palavra “qualidade” possui significados distintos, dependendo do contexto em que é aplicada. Lovelock, Wirtz e Hemzo (2011) exploram as diferentes perspectivas relacionadas a esse termo. A visão transcendente da qualidade é associada à excelência intrínseca, caracterizada por padrões rigorosos e um alto nível de realização. Esse ponto de vista é frequentemente aplicado nas artes dramáticas, visuais e no setor de serviços de luxo, e pressupõe que as pessoas aprendem a reconhecer a qualidade por meio da exposição repetida. No entanto, essa abordagem prática não é muito útil, pois é difícil determinar quando os clientes reconhecerão a qualidade ao encontrá-la.

A abordagem voltada para a manufatura se preocupa em aplicar práticas de engenharia e operações que visam à conformidade com especificações internas, muitas vezes orientadas por metas de produtividade e contenção de custos. Por outro lado, as definições de qualidade baseadas no usuário partem do pressuposto de que a qualidade está nos olhos de quem vê e a equiparam com máxima satisfação. Essa perspectiva subjetiva, centrada na demanda, reconhece que diferentes clientes têm desejos e necessidades diferentes. A visão baseada em “valor” estabelece a qualidade em termos de relação entre valor e preço. Considerando o *trade-off* entre desempenho (ou conformidade) e preço, a qualidade é definida como “a excelência possível com os recursos disponíveis”. Essas

diversas visões da qualidade explicam os eventuais conflitos entre definições e enfatizam a natureza multifacetada do conceito.

No entanto, pesquisadores argumentam que a natureza dos serviços demanda uma abordagem distinta para definir e medir a qualidade. Os serviços são caracterizados por sua natureza intangível e multifacetada, o que torna desafiador avaliar sua qualidade em comparação com bens tangíveis. Lovelock, Wirtz e Hemzo (2011) explicam que, dentre as múltiplas facetas da qualidade, a qualidade de serviço difere porque o cliente está completamente envolvido no processo de entrega, o que é conhecido como qualidade funcional, e no resultado efetivo do serviço, chamado de qualidade técnica, de acordo com Christian Grönroos.

Segundo Grönroos, a qualidade percebida resulta de um processo em que os clientes comparam suas percepções da entrega do serviço e seu resultado com aquilo que esperavam. Dessa forma, a qualidade de serviço, do ponto de vista do usuário, é definida como aquilo que atende ou supera as expectativas dos clientes. A satisfação do cliente está diretamente relacionada à percepção de que o serviço entregue corresponde ou excede suas expectativas, enquanto a frustração ocorre quando o serviço não atende às expectativas estabelecidas. Essa abordagem ressalta a importância de compreender as expectativas dos clientes e buscar atendê-las de forma consistente para garantir a qualidade do serviço.

Atualmente, muitas empresas combinam a oferta de serviços com a venda de produtos, como é o caso dos *smartphones*. Segundo Lovelock, Wirtz e Hemzo (2011), a maioria das empresas de bens agora se considera principalmente provedora de serviços, buscando oferecer uma solução completa para as necessidades dos clientes. O produto de serviço é uma combinação de produto e serviço, sendo predominantemente composto pelo componente intangível do serviço. Essa abordagem reflete uma mudança na perspectiva das empresas, que reconhecem a importância de agregar valor por meio dos serviços oferecidos. Integrar serviços com produtos tangíveis permite aprimorar a experiência do cliente, oferecer suporte pós-venda, personalizar serviços e estabelecer relacionamentos duradouros. Essa evolução destaca que a oferta de serviços vai além do simples fornecimento de um produto físico, com os clientes valorizando cada vez mais a qualidade e a excelência dos serviços prestados, o que impacta diretamente sua percepção da marca e sua satisfação como consumidores.

Com o avanço da tecnologia, os serviços eletrônicos têm se tornado cada vez mais relevantes, abrangendo uma ampla gama de atividades online, como compras, atendimento ao cliente e transações bancárias. A qualidade desses serviços desempenha um papel crucial na garantia de uma experiência positiva para o cliente e na construção de sua confiança. Lovelock, Wirtz e Hemzo (2011) destacam que os consumidores utilizam uma variedade de serviços e realizam operações cotidianas, como fazer ligações telefônicas, realizar compras com cartão de crédito e sacar dinheiro em caixas eletrônicos, que muitas vezes passam despercebidas, a menos que algo dê errado. No entanto, os clientes nem sempre estão sa-

tisfeitos com a qualidade e o valor dos serviços que recebem, frequentemente reclamando de promessas não cumpridas, falta de valor em relação ao preço pago, falta de compreensão de suas necessidades, atendimento inadequado por parte dos funcionários, horários inconvenientes, burocracia, perda de tempo, máquinas de autosserviço com defeito, sites complexos e uma série de outros problemas. É fundamental que os provedores de serviços eletrônicos se esforcem para atender e superar as expectativas dos clientes, oferecendo serviços eficientes, convenientes e personalizados, a fim de conquistar sua satisfação e fidelidade.

A prestação de um serviço, de acordo com Lovelock, Wirtz e Hemzo (2011), envolve a ação de servir, ajudar ou beneficiar, buscando o bem-estar e a vantagem do cliente. No entanto, as promessas feitas na entrega dos serviços muitas vezes não são completamente cumpridas, devido a falhas humanas, de equipamentos ou de processos. A qualidade de serviços está relacionada à percepção de excelência por parte dos clientes em relação a um produto ou serviço oferecido por uma empresa. Isso envolve a conformidade com as expectativas e requisitos dos clientes, assim como a capacidade de atender às suas necessidades. É fundamental que as empresas compreendam as expectativas dos clientes e se esforcem para superá-las, a fim de proporcionar uma experiência positiva e alcançar a satisfação dos clientes.

As expectativas dos consumidores, segundo Lovelock, Wirtz e Hemzo (2011), são influenciadas por informações de mercado, recomendações “boca a boca” e experiências anteriores, resultando em certos padrões de serviço que eles esperam antes do consumo. Durante o encontro com o serviço, os clientes avaliam o desempenho em relação ao nível previsto, formando assim suas avaliações de satisfação. Essas avaliações podem resultar em não confirmação positiva (se o serviço superou as expectativas), não confirmação negativa (se ficou aquém das expectativas) ou simples confirmação (se atendeu às expectativas). Portanto, a satisfação do cliente é determinada pela comparação entre o que eles esperavam e o que percebem ter recebido do provedor de serviço. É importante ressaltar que as expectativas são subjetivas e variam entre os clientes, o que pode tornar o que é satisfatório para um cliente não necessariamente satisfatório para outro.

As métricas de qualidade desempenham um papel fundamental na avaliação e monitoramento da qualidade de serviços. Elas abrangem uma variedade de indicadores, incluindo satisfação do cliente, tempo de resposta, taxa de resolução de problemas e taxa de reclamações, entre outros. Essas métricas permitem que as empresas avaliem seu desempenho e identifiquem áreas que necessitam de melhorias. Além disso, a qualidade de serviço pode ser avaliada por meio de métricas e indicadores que consideram diversos aspectos da experiência do cliente, como tempo de espera, cortesia do atendimento, confiabilidade e eficácia na resolução de problemas. Ao utilizar essas métricas, as empresas podem obter conhecimentos valiosos para aprimorar a qualidade de seus serviços, proporcionando uma experiência satisfatória e positiva aos clientes.

Métricas de qualidade são essenciais para gerenciar a qualidade de serviços. Sem elas, é difícil detectar lacunas no serviço e, conseqüentemente, planejar ações corretivas e medir a eficácia das melhorias implementadas. Parasuraman, Zeithaml e Berry (1988) realizaram uma extensa pesquisa sobre qualidade em serviços e identificaram cinco critérios usados pelos consumidores para avaliá-la. Eles constataram uma alta correlação entre essas variáveis, o que os levou a consolidá-las em cinco dimensões: tangibilidade (aparência física), confiabilidade (desempenho preciso e digno de confiança), responsividade (rapidez e prestimosidade), segurança (credibilidade, segurança, competência e cortesia) e empatia (acesso fácil, boa comunicação e entendimento do cliente). Utilizar essas métricas é fundamental para monitorar o desempenho da empresa e identificar áreas que precisam de melhorias para atender às expectativas dos clientes e proporcionar uma experiência satisfatória.

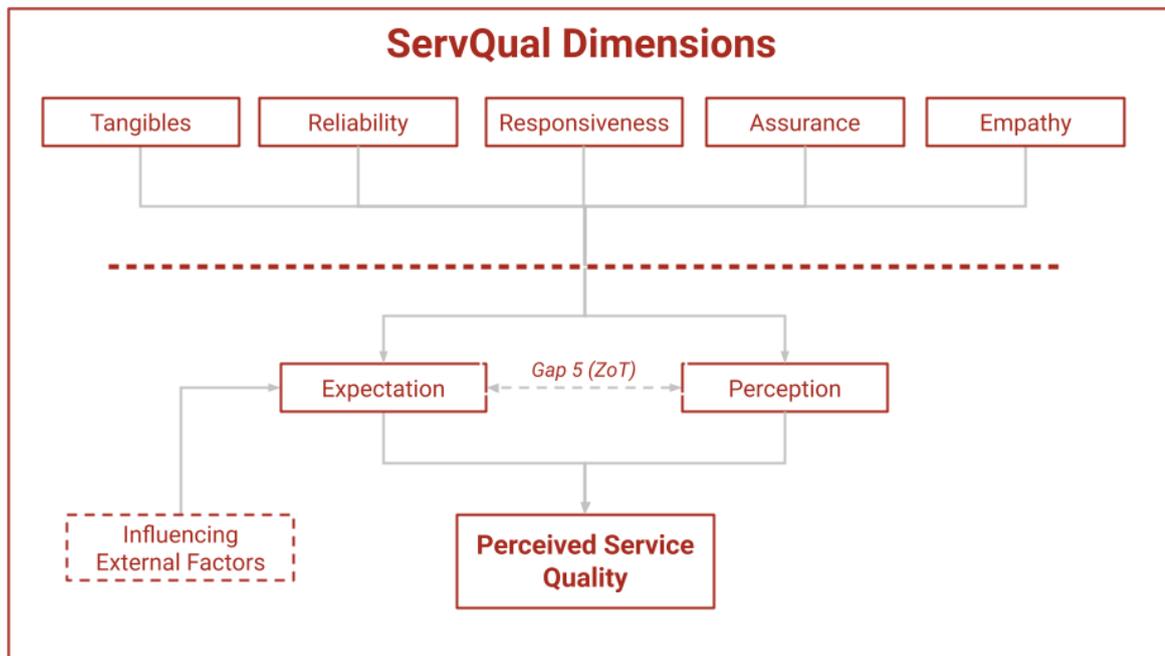


Figura 2 – Diagrama ServQual Gap 5 (Zona de Tolerância)

Fonte: Parasuraman, Zeithaml e Berry (1988)

O modelo ServQual, proposto por Parasuraman, Zeithaml e Berry (1988) e referenciado por Kotler e Keller (2016), bem como Zeithaml, Bitner e Gremler (2017), é uma técnica amplamente reconhecida para avaliar a qualidade de um serviço com base nas expectativas e percepções dos usuários. Ela utiliza a técnica de desconfirmação para medir a diferença entre as expectativas do usuário e sua percepção real. Essa métrica avalia a percepção do usuário em relação à qualidade do serviço prestado, o que está associado à satisfação. Ela mede a tolerância, analisando a diferença entre as expectativas dos usuários e a performance observada, o que influencia a percepção da qualidade. Em outras

palavras, se o serviço estiver dentro dos limites aceitáveis, ele é considerado como um serviço de qualidade.

O modelo considera a Zona de Tolerância (ZoT) como uma faixa de níveis de serviço satisfatórios para os usuários. Durante a experiência do serviço, as expectativas do usuário atuam como padrões de referência, enquanto as percepções são avaliações subjetivas, baseadas em experiências reais, que acabam por formar a ZoT. O modelo ServQual tem sido amplamente adotado na literatura como uma metodologia para avaliar a qualidade dos serviços, sendo um dos instrumentos mais utilizados nesse contexto.

A Figura 3 representa de forma resumida o modelo ServQual, que se concentra na análise das discrepâncias entre as expectativas e percepções dos usuários, identificando lacunas importantes na avaliação da qualidade dos serviços. O “*gap 5*”, conhecido como *gap* do usuário, é especialmente relevante, pois mede a diferença entre as expectativas e percepções dos usuários em relação ao serviço recebido. Além disso, a Zona de Tolerância é uma medida abrangente, que indica a diferença entre o nível de serviço considerado adequado e o nível desejado pelos usuários, auxiliando na avaliação do atendimento (ou do não atendimento) das expectativas. Em última análise, a aplicação desse modelo permite avaliar a qualidade percebida pelos usuários em relação a um serviço específico, desempenhando um papel fundamental no sucesso de qualquer produto ou serviço.

O modelo ServQual foi desenvolvido para medir a qualidade do serviço com base nas percepções dos usuários, através de dimensões que formam uma estrutura do conceito de qualidade do serviço. São elas:

- ❑ **Confiabilidade (*Reliability*):** refere-se à capacidade de fornecer o serviço de forma confiável, precisa e consistente. Os clientes esperam que o serviço seja realizado de forma confiável e sem erros, dentro do prazo prometido e conforme acordado;
- ❑ **Responsividade (*Responsiveness*):** diz respeito à disposição e prontidão dos agentes envolvidos em prestar assistência e atender às necessidades dos clientes. Os clientes valorizam uma resposta rápida e eficiente a solicitações, bem como a disposição em ajudar e resolver problemas;
- ❑ **Competência (*Assurance*):** refere-se à habilidade e conhecimento dos agentes envolvidos em fornecer o serviço. Os clientes esperam que os agentes envolvidos sejam competentes e capazes de realizar suas tarefas de forma adequada, com a expertise e habilidades necessárias;
- ❑ **Empatia (*Empathy*):** envolve o cuidado, a atenção e a consideração demonstrados pelos agentes envolvidos em relação aos clientes. Os clientes valorizam ser tratados com cortesia, respeito e empatia, além de receber atenção personalizada às suas necessidades individuais;

- ❑ Tangibilidade (*Tangibles*): diz respeito aos aspectos físicos e visíveis do serviço, como instalações, equipamentos, materiais e comunicação visual. Os clientes avaliam a aparência física do ambiente de serviço, bem como os materiais tangíveis fornecidos, para formar percepções sobre a qualidade do serviço.

Lovelock, Wirtz e Hemzo (2011) discutem a importância das métricas de qualidade de serviços, que podem ser classificadas em métricas intangíveis (*soft measures*) e métricas tangíveis (*hard measures*). As métricas intangíveis são obtidas por meio de conversas e avaliações de clientes, funcionários e outros *stakeholders*. Um exemplo sofisticado de métrica intangível é o ServQual (LOVELOCK; WIRTZ; HEMZO, 2011). Por outro lado, as métricas tangíveis estão relacionadas a características e atividades mensuráveis diretamente, como o número de ligações telefônicas interrompidas durante a espera do cliente, a precisão no preenchimento de pedidos, o tempo necessário para concluir uma tarefa, o tempo de espera em filas, a pontualidade de trens e o extravio de bagagens. Estabelecer padrões e metas é essencial, e os profissionais devem garantir que as métricas operacionais de qualidade forneçam insights para a implementação de melhorias. É fundamental que o cliente seja a principal fonte de referência para avaliar o desempenho do serviço.

Para avaliar a satisfação do cliente em relação à qualidade do serviço, Parasuraman, Zeithaml e Berry (1988) desenvolveram o instrumento de pesquisa conhecido como ServQual, que se baseia na comparação entre as percepções dos clientes e suas expectativas. O ServQual é amplamente utilizado em diversos setores de serviços e consiste em uma escala básica com 22 itens de percepção, acompanhados por uma série de itens de expectativa que abrangem as cinco dimensões da qualidade de serviço. No entanto, apesar de sua ampla adoção pelas empresas de serviços, o ServQual possui algumas limitações. Muitos pesquisadores optam por omitir, aprimorar ou modificar a lista de afirmações a fim de capturar alguns aspectos da qualidade de serviço mais específicos de um contexto. Lovelock, Wirtz e Hemzo (2011) sugerem que o ServQual mede principalmente dois fatores: a qualidade intrínseca do serviço (similar ao conceito de qualidade funcional proposto por Grönroos) e a qualidade extrínseca do serviço (que se refere aos aspectos tangíveis da entrega do serviço e se assemelha ao conceito de qualidade técnica, também proposto por Grönroos). Dessa forma é possível concluir que existe uma relação entre as métricas técnicas, a percepção do usuário e a resultante dessa relação, que é a qualidade.

Lovelock, Wirtz e Hemzo (2011) utilizam a Figura 3 para ilustrar a dinâmica dos fatores que impactam a Zona de Tolerância, de Parasuraman, Berry e Zeithaml (1991).

- ❑ O serviço desejado representa o tipo de serviço que os clientes esperam receber e consideram ideal, sendo uma combinação do que acreditam que pode e deve ser entregue, levando em consideração suas necessidades pessoais. O serviço desejado é influenciado por promessas explícitas e implícitas feitas pelos fornecedores, comentários “boca a boca” e experiências anteriores. É aquilo que os clientes acreditam

que poderão receber, caso tudo ocorra de forma adequada. No entanto, a maioria dos clientes é realista e compreende que nem sempre as empresas podem fornecer o nível desejado de serviço. Portanto, eles também estabelecem um patamar mínimo para suas expectativas, chamado de serviço adequado;

- ❑ O serviço adequado representa o nível mínimo de serviço que os clientes aceitarão sem ficar insatisfeitos. Por outro lado, o serviço previsto é o nível de serviço que os clientes esperam receber, e também pode ser influenciado por promessas dos fornecedores, comentários “boca a boca” e experiências anteriores. O nível de serviço previsto tem um impacto direto na forma como os clientes definem o que é considerado “serviço adequado” em uma determinada situação. Se os clientes esperam um bom serviço, o nível adequado será mais elevado do que quando esperam um serviço de qualidade inferior;
- ❑ As previsões de serviço feitas pelos clientes podem ser contextuais. Por exemplo, com base em experiências anteriores, clientes que visitam um museu em um dia de verão podem esperar encontrar um maior número de pessoas se o clima estiver ruim, em comparação a um dia ensolarado. Portanto, uma espera de dez minutos para comprar ingressos em um dia frio e chuvoso não será considerada abaixo do nível adequado de serviço, de acordo com a percepção dos clientes;
- ❑ O nível de serviço previsto oferecido por fornecedores alternativos é outro fator que influencia as expectativas dos clientes. Os clientes compreendem que pode ser desafiador para as empresas manter uma entrega consistente de serviço em todos os pontos de contato, atravessando diversos canais, filiais e envolvendo milhares de funcionários. A proporção dessa variação que os clientes estão dispostos a tolerar é conhecida como zona de tolerância, que abrange o intervalo entre o nível mínimo adequado e o nível máximo desejado pelo cliente.

Quando o desempenho fica abaixo do nível de serviço adequado, isso resulta em frustração e insatisfação para os clientes. Por outro lado, quando o desempenho excede o nível de serviço desejado, os clientes ficam surpreendidos e encantados. Podemos também visualizar a zona de tolerância como a faixa na qual os clientes não prestam atenção explícita ao desempenho do serviço. No entanto, quando o serviço se desvia dessa faixa, as reações dos clientes podem ser positivas ou negativas.

A zona de tolerância explicada por Lovelock, Wirtz e Hemzo (2011) refere-se à faixa que compreende as expectativas mínimas e máximas do cliente em relação a um serviço. Dentro dessa faixa, os clientes podem se sentir satisfeitos ou indiferentes em relação ao serviço recebido. No entanto, caso o serviço fique aquém das expectativas mínimas ou ultrapasse as expectativas máximas, os clientes tendem a ficar insatisfeitos. Lovelock,

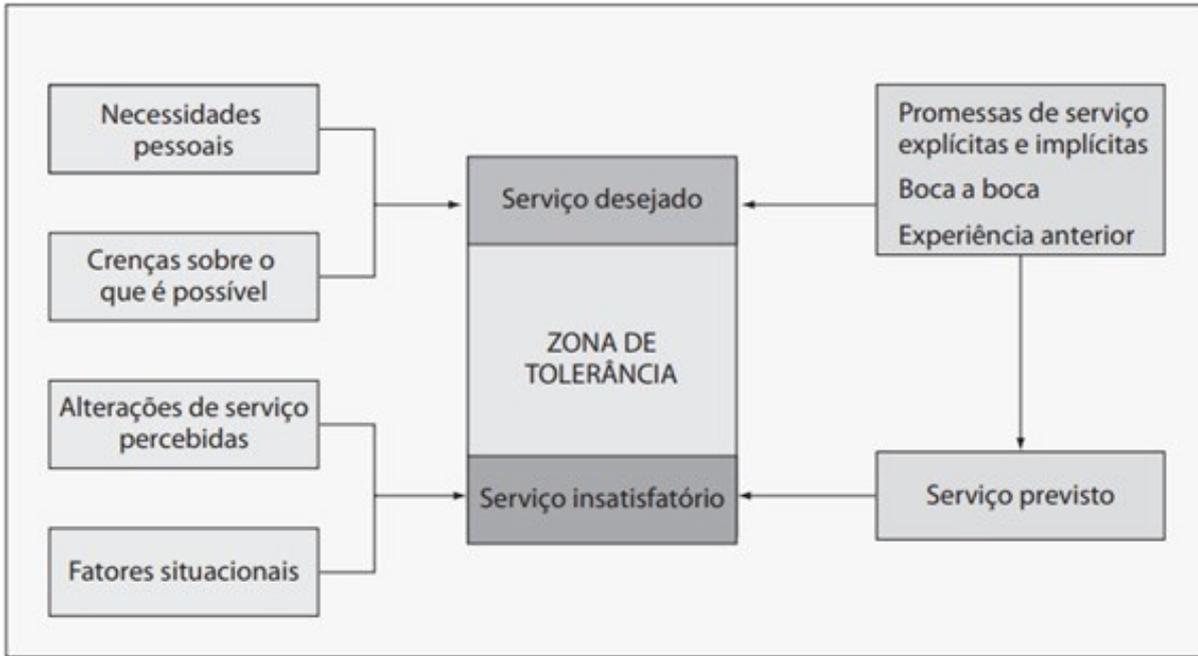


Figura 3 – Fatores relacionados ao serviço desejado e insatisfatório (Zona de Tolerância)

Fonte: Lovelock, Wirtz e Hemzo (2011)

Wirtz e Hemzo (2011) destacam que a zona de tolerância pode variar de cliente para cliente, influenciada por fatores como concorrência, preço ou importância atribuída a características específicas do serviço, os quais podem afetar o nível considerado adequado de serviço. Por outro lado, os níveis de serviço costumam evoluir lentamente em resposta às experiências acumuladas pelos clientes.

Os clientes tendem a ficar razoavelmente satisfeitos quando percebem que o desempenho está dentro da zona de tolerância, ou seja, acima do nível mínimo adequado. À medida que a percepção de desempenho se aproxima dos níveis desejados ou os supera, os clientes ficam muito satisfeitos, tornando-se mais propensos a repetir a compra e a se tornarem leais ao provedor, além de compartilhar comentários positivos “boca a boca”. No entanto, se a experiência de serviço não atender às expectativas, ficando abaixo da zona de tolerância, os clientes podem manifestar insatisfação, sofrer em silêncio (mas considerando alternativas e expressando comentários negativos “boca a boca”) ou buscar outro fornecedor. Em mercados altamente competitivos, muitos clientes de serviços esperam que seus provedores antecipem suas necessidades implícitas e as satisfaçam. Essa capacidade de antecipação é vista como um componente do serviço e afeta a percepção da qualidade do serviço.

Por outro lado, em mercados não competitivos ou situações em que os consumidores não têm escolha livre (devido a custos de troca proibitivos ou restrições de tempo ou localização, por exemplo), existem riscos em definir a satisfação do cliente em relação

às expectativas anteriores. Por exemplo, se as expectativas do cliente forem baixas e o serviço fornecido atender apenas ao nível mínimo esperado, ele dificilmente considerará ter recebido um serviço de boa qualidade.

A satisfação do cliente desempenha um papel fundamental na qualidade dos serviços. Refere-se ao grau de resposta ao atendimento do cliente, ou seja, é o julgamento de que um produto ou serviço proporciona um nível prazeroso de contentamento. Clientes satisfeitos tendem a demonstrar maior fidelidade à empresa, recomendar seus serviços e continuar realizando negócios com ela. A Figura 4 exemplifica a relação de fatores que impactam a satisfação do cliente.

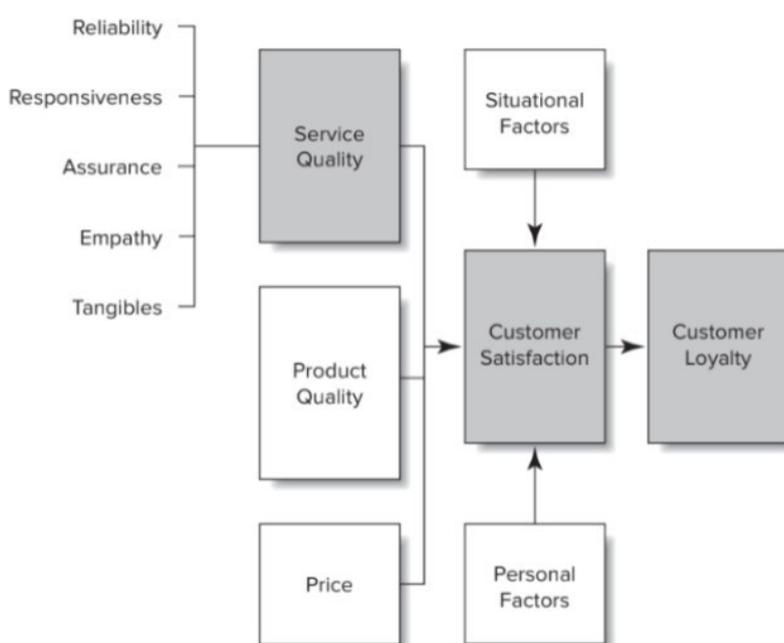


Figura 4 – Percepção do consumidor sobre a qualidade e a satisfação do cliente.

Fonte: Zeithaml, Bitner e Gremler (2017)

A satisfação é um indicador de extrema importância, frequentemente mensurado por gestores de serviços. Zeithaml, Bitner e Gremler (2017) e Lovelock, Wirtz e Hemzo (2011), bem como Oliver (2010) relatam pesquisas que fornecem evidências convincentes da relação existente entre o nível de satisfação dos clientes em relação a uma empresa e seu desempenho geral. É importante ressaltar que muitas pesquisas de satisfação levam em consideração produtos com atributos de busca e experiência elevados. A satisfação pode ser entendida como uma avaliação de atitude, que ocorre após uma experiência de consumo. Entretanto muitos estudos baseiam-se na teoria de que a confirmação ou não confirmação das expectativas prévias desempenha um papel fundamental na determinação da satisfação (LOVELOCK; WIRTZ; HEMZO, 2011).

A avaliação da opinião do consumidor sobre a qualidade dos serviços é comumente realizada por meio de pesquisas de opinião, conforme apontado por Lovelock, Wirtz e Hemzo (2011). No entanto, é importante reconhecer que surgem desafios significativos quando os consumidores são solicitados a avaliar a qualidade de serviços que possuem características altamente credíveis, como casos jurídicos ou tratamentos médicos complexos. Esses serviços são frequentemente percebidos como difíceis de ser avaliados de maneira precisa após a sua entrega. Isso se deve, em parte, à natureza complexa e subjetiva desses serviços, além da influência de fatores externos que podem afetar a percepção do consumidor.

A recuperação de serviços envolve as ações tomadas por uma empresa para resolver problemas e atender às reclamações dos clientes. Uma recuperação eficaz é capaz de restaurar a confiança do cliente, melhorar a satisfação e fortalecer o relacionamento com ele. Lovelock, Wirtz e Hemzo (2011) destacam a recuperação de serviço como uma ação que engloba os esforços sistemáticos da empresa para corrigir um problema imediatamente após uma falha e manter a boa vontade do cliente. Os esforços de recuperação de serviço desempenham um papel crucial na conquista (ou restauração) da satisfação e fidelidade do consumidor. Durante a prestação do serviço, podem ocorrer eventos que afetam negativamente os relacionamentos com os clientes, representando um teste para o comprometimento da empresa com a satisfação e qualidade do serviço. Isso vai além das promessas de publicidade e se reflete em sua reação quando as coisas dão errado para o cliente. Quando as reclamações são resolvidas de maneira satisfatória, os clientes tendem a se manter leais, e aqueles que experimentaram uma falha de serviço solucionada adequadamente em geral demonstram-se mais satisfeitos do que aqueles que nunca tiveram problemas.

Lovelock, Wirtz e Hemzo (2011) citam pesquisas que revelam o “paradoxo da recuperação de serviço”, que se aplica apenas à primeira falha de serviço recuperada de maneira totalmente satisfatória para o cliente. No entanto, quando ocorre uma segunda falha, o paradoxo não se repete. Parece que os clientes podem perdoar a empresa uma vez, mas ficam desapontados se os erros persistirem. A frustração pode ser maior devido à ineficiência do que à falta de ação. Além disso, o estudo demonstrou que as expectativas dos clientes aumentam após uma recuperação muito bem-sucedida, tornando a excelência na recuperação o padrão que eles esperam para o tratamento de problemas futuros. Vale ressaltar que a satisfação do cliente com a recuperação de serviço também depende da gravidade e da eficácia da correção da falha. Nem sempre é possível oferecer uma compensação adequada que restaure completamente a satisfação. Nesse contexto, Lovelock, Wirtz e Hemzo (2011) consideram a recuperação de serviço como a transformação de uma falha de serviço em uma oportunidade que a empresa desejaria nunca ter tido.

O modelo de Kano (1984), também conhecido como análise de Kano, é uma abordagem teórica desenvolvida pelo professor japonês Noriaki Kano na década de 1980. Ele propôs um modelo para entender e classificar as expectativas dos clientes em relação aos

atributos de um produto ou serviço. A análise de Kano é uma abordagem teórica utilizada no gerenciamento da qualidade e do marketing a fim de entender e classificar as expectativas dos clientes em relação aos atributos de um produto ou serviço. Esse modelo fornece um guia valioso sobre como os atributos de atingimento das funcionalidades afetam a satisfação do cliente e pode auxiliar as organizações a priorizar seus esforços de desenvolvimento e melhoria com foco na experiência do usuário.

O modelo de Kano (1984) categoriza os atributos em cinco tipos principais:

- ❑ **Atributos básicos:** são aqueles que são esperados pelos clientes e não necessariamente os satisfazem quando estão presentes, mas causam insatisfação quando ausentes. São considerados requisitos mínimos e, quando atendidos, não geram diferenciação competitiva. Um exemplo seria a capacidade de um carro ter rodas e volante;
- ❑ **Atributos de desempenho:** são aqueles que estão diretamente relacionados ao desempenho do produto ou serviço. Quanto melhor esses atributos forem, maior será a satisfação do cliente. A presença desses atributos gera satisfação e sua ausência gera insatisfação. Por exemplo, a velocidade de processamento de um computador ou a qualidade do som de um sistema de áudio;
- ❑ **Atributos de encantamento:** são aqueles que excedem as expectativas dos clientes, surpreendendo-os positivamente. Sua presença cria satisfação extrema e sua ausência não gera insatisfação. Esses atributos geralmente não são mencionados pelos clientes, pois estes não estão necessariamente cientes de sua existência. Eles podem ser recursos inovadores ou extras que agregam valor ao produto ou serviço;
- ❑ **Atributos de indiferença:** são aqueles que não têm impacto significativo na satisfação do cliente. Sua presença ou ausência não influencia a percepção do cliente. Por exemplo, a cor de um produto pode ser um atributo de indiferença, pois a maioria dos clientes não atribui importância a ela;
- ❑ **Atributos reversos:** são aqueles que, quando presentes, podem causar insatisfação aos clientes. A ausência desses atributos não gera satisfação nem insatisfação, mas sua presença pode resultar em uma experiência negativa. Um exemplo seria uma interface de usuário complexa e difícil de usar.

O modelo de Kano descrito na Figura 5, auxilia as organizações a entender as diferentes expectativas dos clientes e a priorizar funcionalidades cujo os atributos mais impactam a satisfação do cliente. Ao identificar o nível de atingimento e satisfação percebido pelos seus consumidores e ao classificar os atributos básicos, de desempenho, de encantamento, de indiferença e reversos, as empresas podem direcionar seus esforços para criar valor para os clientes e diferenciar-se da concorrência.

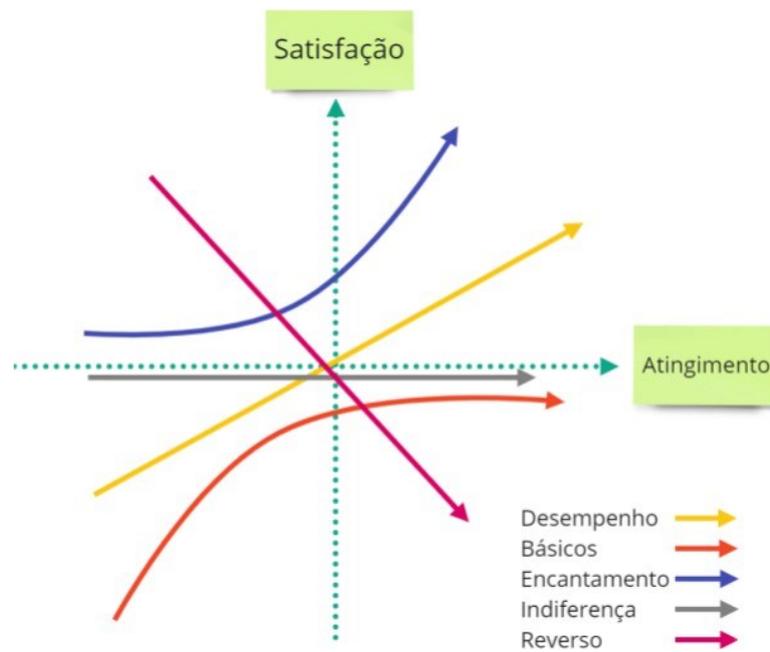


Figura 5 – Percepção do consumidor sobre as funcionalidades do produto ou serviço.

Fonte: Kano (1984)

Em resumo, a satisfação do usuário é amplamente utilizada como uma medida essencial da qualidade percebida, mas é influenciada por diversos fatores além da própria qualidade. Com base em modelos da literatura que abordam a percepção do usuário e o modelo de desconformação, conhecido como zona de tolerância, foi desenvolvido um modelo teórico (descrito na Figura 6) para compreender o conceito de qualidade da experiência. Um dos modelos mais conhecidos para avaliar a qualidade é o modelo ServQual, ilustrado na Figura 2, que propõe cinco dimensões de qualidade: confiabilidade, responsividade, competência, empatia e tangibilidade. Esse modelo de qualidade de serviços utiliza essas dimensões para medir a lacuna entre as expectativas dos clientes e suas percepções reais do serviço recebido, contribuindo para a avaliação da qualidade percebida. No contexto do modelo teórico descrito na Figura 6, considera-se a experiência e a percepção do impacto das falhas como única dimensão de qualidade. Neste modelo estão representadas as forças mais significativas que afetam diretamente a satisfação do usuário, com base em modelos disponíveis na literatura (ilustrados nas Figuras 4 e 3) e na base de conhecimento relatada.

No modelo descrito na Figura 6, a zona de tolerância está relacionada à percepção dos clientes em relação ao serviço desejado, serviço aceitável e serviço indesejado. O serviço desejado representa o nível ideal de serviço que os clientes esperam receber, com base em suas necessidades, preferências, experiências anteriores e promessas de serviços. A zona de tolerância é o ponto de referência para avaliar a qualidade do serviço. O serviço indesejado refere-se a um nível de serviço abaixo das expectativas dos clientes, resultando

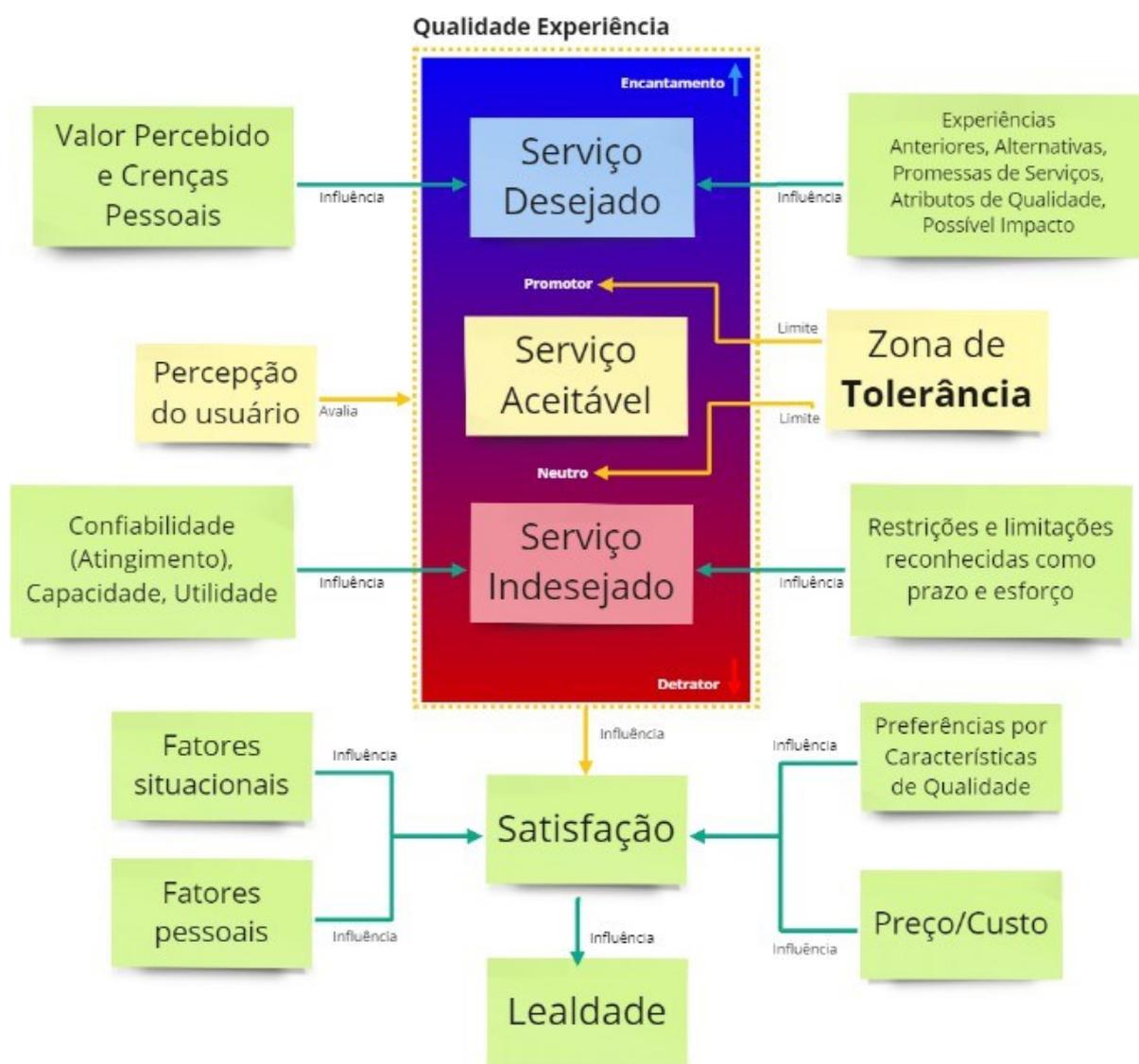


Figura 6 – Resumo do modelo teórico da relação entre satisfação e a qualidade da experiência.

em insatisfação e percepção de baixa qualidade. Por fim, o serviço aceitável abrange a faixa entre o serviço desejado e o serviço indesejado, representando a zona de tolerância. Dentro dessa zona, as variações na entrega do serviço são consideradas aceitáveis e não afetam significativamente a satisfação do cliente. No entanto, se a qualidade do serviço se desviar além dessa zona, os clientes podem se sentir insatisfeitos (detratores) ou podem se sentir encantados. Portanto, a zona de tolerância é o mecanismo que avalia a qualidade do serviço através das percepções dos clientes.

O atingimento, as restrições e as limitações estão intimamente relacionadas com o serviço prestado. O atingimento refere-se à capacidade da organização de alcançar e superar as expectativas dos clientes, proporcionando um serviço de alta qualidade que atenda às

suas necessidades e desejos. Isso envolve a entrega eficiente e eficaz do serviço, cumprindo os padrões e promessas estabelecidos. Por outro lado, as restrições e limitações podem surgir de diversos fatores, como recursos financeiros, disponibilidade de pessoal, restrições regulatórias, entre outros. Essas restrições podem afetar a capacidade da organização de fornecer o serviço desejado ou alcançar um alto nível de qualidade. O equilíbrio entre o atingimento das expectativas dos clientes e a avaliação das restrições e limitações por parte dos usuários é essencial para não gerar impacto na experiência garantindo a satisfação através da qualidade do serviço prestado.

A satisfação do usuário é influenciada por uma variedade de fatores, incluindo fatores situacionais, fatores pessoais, características de qualidade e preço. Fatores situacionais são as circunstâncias específicas em que ocorre a interação entre um usuário e um serviço. Estes incluem o ambiente físico, o contexto social, o tempo e a conveniência. Esses fatores influenciam a percepção do usuário sobre a qualidade do serviço e sua satisfação geral, e proporcionam uma experiência positiva aos usuários. Fatores pessoais são características individuais que influenciam a percepção e resposta de um usuário em relação a um serviço. Eles incluem expectativas, necessidades, valores, personalidade e experiências anteriores. Esses fatores são únicos para cada usuário e devem ser considerados pelas empresas ao fornecer serviços, visando atender às necessidades individuais e promover a satisfação do usuário.

O preço está relacionado ao valor percebido pelo usuário em troca do serviço recebido. Quando o preço é considerado justo e proporciona um bom custo-benefício, isso tende a aumentar a satisfação do cliente. No entanto, se o preço for considerado alto demais em relação ao valor percebido, pode resultar em menor satisfação. As características de qualidade são os atributos que definem a excelência de um serviço. Essas características são essenciais para a percepção da qualidade do serviço e a satisfação do cliente. Altos níveis de qualidade levam à lealdade do usuário. A lealdade é a disposição contínua de escolher uma empresa específica em vez da concorrência. É baseada em experiências positivas, satisfação contínua, confiança na marca e relacionamento duradouro. Usuários leais utilizam os serviços prestados repetidamente, recomendam e contribuem para o sucesso do negócio.

2.2.5 Qualidade da Experiência (QoE)

A qualidade de serviços é um tema abordado em diversos contextos (incluindo as telecomunicações), e que já tem sido objeto de estudo já há muitos anos. Segundo o ITUT (2008), a qualidade é a combinação de características de uma entidade que afetam sua capacidade de atender às necessidades declaradas e implícitas, desde que essas características possam ser observáveis e mensuráveis. Além disso, o ITUT (2008) define a Qualidade de Serviço (QoS) como o conjunto de características de um serviço de telecomunicações que influenciam sua capacidade de satisfazer as necessidades declaradas e implícitas do usuário. O termo “característica” é percebido como uma propriedade que distingue os indivíduos em uma população, podendo ser quantitativa (variáveis) ou qualitativa (atributos).

Com o rápido desenvolvimento de sistemas e serviços de Tecnologia da Informação e Comunicação (TIC), tornou-se crucial avaliar sua qualidade. Embora seja comumente acreditado que alto desempenho leva a alta aceitação dos serviços, há exemplos que contradizem essa relação, onde sistemas de baixo desempenho são percebidos como um sucesso pelos usuários, enquanto sistemas de alto desempenho podem ser considerados um fracasso. Para abordar essa questão, o ITUT (2008) introduziu o conceito de qualidade de serviço experienciada e percebida pelo usuário (QoSE), que expressa o nível de qualidade que os usuários acreditam ter experimentado. A QoSE é baseada na opinião e avaliação dos usuários e possui dois componentes principais: quantitativo e qualitativo. O componente quantitativo é influenciado pelos efeitos do sistema, como infraestrutura de rede, enquanto o componente qualitativo é influenciado pelas expectativas do usuário, condições ambientais, fatores psicológicos, contexto de aplicação, entre outros.

No contexto atual, a opinião dos usuários desempenha um papel crucial na definição da qualidade esperada de um serviço. Möller e Raake (2014) explicam que a qualidade, do ponto de vista de uma pessoa, é o resultado do julgamento sobre a harmonização percebida de uma entidade em relação ao seu arranjo desejado. Esse conceito foi denominado “Quality of Experience” (QoE) e se popularizou, especialmente no contexto de sistemas e serviços de transmissão de mídia. A QoE complementa a Qualidade de Serviço (QoS), que se concentra em métricas objetivas relacionadas à entrega de dados, como largura de banda, atraso e taxa de perda de pacotes. A principal diferença entre QoE e QoS reside na perspectiva adotada. Enquanto a QoS avalia parâmetros técnicos e objetivos, a QoE leva em consideração a percepção subjetiva do usuário e seu nível de satisfação com o produto ou serviço. A QoE considera fatores emocionais, expectativas individuais, contexto de uso e preferências pessoais, os quais podem variar de usuário para usuário.

De acordo com Möller e Raake (2014), a experiência é definida como o fluxo individual de percepções, incluindo sentimentos, sensações e conceitos, que ocorrem em uma situação específica de referência. A Qualidade de Serviço (QoS), por sua vez, abrange todas as características de um serviço que influenciam sua capacidade de satisfazer as necessidades

declaradas e implícitas do usuário. Já a Qualidade da Experiência (QoE) refere-se ao grau de prazer ou desconforto de uma pessoa em relação a um aplicativo, serviço ou sistema, resultante da avaliação da pessoa em relação ao cumprimento de suas expectativas e necessidades em termos de utilidade (pragmática e hedônica) e/ou satisfação, considerando o contexto de uso, a personalidade e o estado atual do indivíduo.

O projeto centrado no usuário é uma abordagem que prioriza as necessidades e requisitos dos usuários, considerando fatores humanos, ergonomia e usabilidade. Isso resulta em maior eficácia, eficiência e satisfação do usuário, além de melhorar o bem-estar geral. A usabilidade desempenha um papel fundamental na Qualidade da Experiência (QoE), influenciando fatores que ajudam no sucesso na realização de tarefas. Uma boa usabilidade contribui para uma experiência agradável, e ajuda a amenizar possíveis problemas que podem prejudicar a QoE. A relação entre desempenho, qualidade, usabilidade e aceitação ainda não é completamente compreendida. O estudo de Interfaces homem-máquina (HCI) se concentra em projetar experiências do usuário, indo além da usabilidade tradicional para abordar a *User Experience* (UX). No entanto, é importante distinguir entre UX e usabilidade, pois são conceitos diferentes e não devem ser usados de forma equivocada.

A relação entre Qualidade da Experiência (QoE) e satisfação do usuário é estreita, pois a percepção de qualidade influencia diretamente a satisfação (Möller e Raake, 2014). Quando a QoE é alta, ou seja, quando o usuário tem uma experiência positiva, a satisfação geral é maior. Por outro lado, uma QoE baixa pode resultar em insatisfação, frustração e até mesmo na interrupção do uso do produto ou serviço. A opinião média dos usuários (Mean Opinion Score - MOS) é uma métrica comumente utilizada para medir a QoE, sendo obtida a partir da resposta dos usuários a perguntas como “Como você avalia a sua experiência durante o uso do aplicativo?” ou relacionadas a características específicas do produto ou serviço, como a qualidade do áudio ou do vídeo, conforme apontado por Möller e Raake (2014). É importante destacar que a QoS e a QoE estão interligadas, uma vez que problemas como latência, lentidão e perda de sincronia em um vídeo online afetam diretamente a QoE do usuário. Em geral, a QoE não é expressa em termos de satisfação, mas sim como um indicativo de prazer ou aborrecimento envolvido em uma relação direta.

A opinião dos usuários desempenha um papel crucial na avaliação da Qualidade da Experiência (QoE) e está intrinsecamente ligada a uma ampla gama de emoções, estados afetivos, fisiológicos e comportamentais. Segundo Möller e Raake (2014), a QoE pode ser influenciada por diversos fatores, incluindo características do usuário, do sistema, do serviço, da aplicação e do contexto do momento. No que diz respeito aos fatores humanos, são consideradas características do usuário que podem ser variáveis ou invariáveis, como histórico demográfico, socioeconômico, constituição física, saúde mental e estado emocional. Os fatores de influência do sistema estão relacionados às propriedades e características técnicas que afetam a qualidade da aplicação ou serviço. Já os fatores de

influência de contexto abrangem propriedades situacionais que descrevem o ambiente em que o usuário está inserido. Essas variáveis contextuais são frequentemente utilizadas para complementar as inferências feitas durante a avaliação da QoE. É essencial considerar esses fatores ao analisar e compreender a QoE, pois desempenham um papel significativo na percepção e na experiência do usuário.

Em resumo, a Qualidade da Experiência (QoE) concentra-se na percepção subjetiva do usuário em relação à qualidade de um produto ou serviço. É fundamental compreender e aprimorar a QoE para oferecer produtos e serviços que atendam às expectativas dos usuários e proporcionem experiências positivas e gratificantes. Segundo Möller e Raake (2014), as qualidades hedônicas abrangem os aspectos não funcionais do sistema, enquanto as qualidades pragmáticas referem-se aos aspectos relacionados à conclusão de tarefas em um sistema. As qualidades pragmáticas de um sistema são apenas suficientes para satisfazer as necessidades, mas não são uma fonte de experiência positiva. Por outro lado, as qualidades hedônicas estão associadas à capacidade do sistema de evocar prazer e promover o bem-estar psicológico do usuário, ou seja, são responsáveis por criar uma experiência positiva. O objetivo de “fazer as pessoas felizes” está intrinsecamente ligado ao objetivo de “projetar produtos melhores” e, do ponto de vista da indústria e dos negócios, isso implica em um possível aumento do sucesso no mercado.

Trabalhos Relacionados

3.1 Confiabilidade percebida pelo usuário

O trabalho de Assunção (2019) aborda a confiabilidade de software, integrando fatores quantitativos e qualitativos, com foco na percepção dos usuários (QoE). Tradicionalmente, as métricas de Quality of Service (QoS), como o Mean Time Between Failures (MTBF), não refletem com precisão como os usuários percebem falhas. Para isso, o estudo utilizou questionários com cenários realistas de uso, identificando dimensões emocionais como estresse, pontualidade e paciência como cruciais para a percepção de confiabilidade. Com base nisso, foi proposto um modelo analítico centrado na QoE, concluindo que avaliações apenas com QoS são insuficientes para captar a experiência dos usuários.

Yabe et al. (2018) exploraram os efeitos de diferentes tipos de falhas na percepção dos usuários sobre a confiabilidade de um Sistema Operacional (SO). Os experimentos foram realizados utilizando cenários realistas de falhas de software do SO Microsoft Windows 7, que são comumente percebidas pelos usuários. O estudo propôs uma variação de cenários considerando o contexto (profissional ou lazer), o conhecimento em tecnologia, o tipo de falha (*kernel* ou aplicação) e as consequências geradas pela falha, como a perda de parte de um documento em edição. A pesquisa foi conduzida em duas fases. Na primeira fase, questionários *online* foram aplicados a 121 participantes, enquanto na segunda fase, questionários presenciais foram realizados com 120 participantes. Cada participante teve a oportunidade de vivenciar um cenário de falha por meio de um simulador. Os pesquisadores concluíram que existem fatores que influenciam a percepção do usuário em relação à qualidade da experiência, mas descobriram que o tipo de falha e o nível de conhecimento em tecnologia não impactam diretamente essa percepção. Com base em suas observações, os pesquisadores identificaram fatores mais relevantes para a percepção da qualidade da experiência (QoE), tais como a consequência da falha, a relação com prazos apertados, o formato da mensagem de falha, a possibilidade de recuperação da falha, a expectativa em relação à solução, o tipo do evento de falha, a frequência do evento de falha e a familiaridade com o evento de falha. Esses fatores desempenham um papel importante na

forma como os usuários percebem a confiabilidade do sistema.

Kuipers et al. (2010) compilaram um conjunto de técnicas para medir a qualidade da experiência do usuário (QoE) com base em medidas de qualidade de serviço (QoS). O objetivo da pesquisa foi avaliar os desafios enfrentados na produção e distribuição de conteúdo de vídeo, áudio e jogos *online*. Os autores destacaram a dificuldade de capturar medidas subjetivas, que envolvem fatores emocionais dos usuários (QoE), e relacioná-las a modelos que utilizam métricas de QoS. De maneira objetiva, os autores conseguiram estabelecer uma correlação entre medidas de QoS e a opinião subjetiva dos usuários, usando um grupo menor de participantes, a fim de prever a QoE com base em um padrão de QoS. O estudo avaliou três contextos de uso: comunicações de áudio, jogos *online* e transmissões de vídeo. Para cada contexto, foram utilizadas métricas diferentes para prever a QoE. Por exemplo, no contexto de áudio, o atraso na rede é um fator importante, enquanto nos jogos *online*, além da latência, a variação da latência (*jitter*) tem uma relação direta com a experiência percebida. Já nas transmissões de vídeo, as características do conteúdo, a sincronização entre áudio e vídeo, e, no caso de transmissões ao vivo, o tempo de propagação são relevantes. A descrição do *framework* de previsão da QoE considera que, para prever corretamente a qualidade de vídeo e áudio, é importante garantir a qualidade da rede. No entanto, muitas vezes, essa qualidade não está sob o controle do provedor de serviços. Ainda assim, essas métricas podem ser monitoradas e combinadas com as percepções humanas do ambiente e as expectativas do usuário. O artigo conclui que conhecer essas métricas é fundamental para realizar um bom *design* de produtos, considerando a experiência do usuário.

Wang e Trivedi (2009) definem a confiabilidade do serviço percebida pelo usuário como a probabilidade de que todas as requisições sejam satisfeitas com sucesso durante a sessão do usuário. Segundo eles, o sucesso da sessão do usuário depende não apenas do estado do sistema, mas também do comportamento do usuário, como o número de solicitações durante a sessão, tempos de resposta do usuário, duração da sessão, entre outros fatores. Sua proposta consiste em utilizar métricas de uso para construir o grafo de comportamento do usuário (*user behavior graph*) como base para um modelo baseado em cadeias de Markov, a fim de calcular a confiabilidade e criar um mecanismo de alocação de serviços em um *cluster* compartilhado, garantindo um nível de serviço adequado. Os autores argumentam que a indisponibilidade de uma parte do *cluster* não resultará necessariamente em uma falha, pois nem todas as partes do sistema são acionadas ao executar uma função específica do sistema, quando comparadas com as métricas em um cenário sem falhas. Eles concluem que a confiabilidade baseada no comportamento do usuário pode ajudar a priorizar as funções mais utilizadas em vez de considerar apenas a indisponibilidade parcial do *cluster* como um todo.

Jalote et al. (2004) propõem um modelo de cálculo de confiabilidade de software que incorpora a percepção do usuário, aplicando pesos diferentes para tipos específicos de

falhas. Esses tipos de falhas são escolhidos de forma a representar as várias percepções de um grupo de usuários e levam em consideração informações complementares sobre o contexto em que a falha ocorre. Essa abordagem de classificação de falhas e atribuição de pesos ao seu impacto foi uma maneira de aproximar a análise tradicional da confiabilidade à perspectiva do usuário. Os autores concluíram que a abordagem tradicional de confiabilidade não atende às necessidades de produtos de massa, nos quais existem grandes variações nos perfis operacionais e grupos de usuários. A pesquisa demonstrou que ao considerar a percepção dos usuários por meio da classificação de falhas, foi possível priorizar iniciativas de melhoria do software que ajudaram a aprimorar a experiência do usuário em relação à confiabilidade do sistema. Essa abordagem mais centrada no usuário proporcionou uma melhor compreensão das necessidades e expectativas destes, levando a melhorias significativas no software.

Cheung (1980) defende o uso de técnicas distintas para medir a confiabilidade, levando em conta o objetivo desejado. Ele argumenta que, no caso de software, é possível oferecer uma boa confiabilidade aos usuários, desde que as partes menos confiáveis do sistema não sejam acionadas. Em sua proposta de um modelo baseado em cadeias de Markov, ele considera o sistema como um conjunto de componentes interativos, em que as ações do usuário influenciam a ativação desses componentes. Isso permite avaliar a confiabilidade ao aprimorar os processos mais frequentemente utilizados. Cheung conclui que, embora não seja possível garantir a acuidade absoluta, é importante orientar as validações levando em consideração a confiabilidade.

A dependabilidade (*dependability*) é uma propriedade essencial dos sistemas computacionais, que se refere à capacidade de fornecer um serviço em que os usuários possam confiar de forma justificada. No entanto, a avaliação da confiabilidade nem sempre reflete a percepção do usuário, como foi relatado por Yabe et al. (2018). Diversos estudos, como o de Kuipers et al. (2010), demonstraram a existência de uma relação entre métricas de Qualidade de Serviço (QoS) e Qualidade de Experiência (QoE). Jalote et al. (2004) concluíram que essa relação pode ser estabelecida através de um fator de correlação, embora encontrar esse fator seja um desafio em si.

Cheung (1980) defende que observar a confiabilidade de um sistema de forma abrangente pode ser custoso, e sugere que apenas as funções relevantes do código, aquelas mais frequentemente utilizadas, sejam consideradas. Por sua vez, Wang e Trivedi (2009) propõem que a confiabilidade seja calculada com base em métricas de comportamento de uso do usuário durante sua sessão de utilização. No entanto, esses modelos não levam em conta fatores emocionais que podem influenciar a percepção de confiabilidade.

O trabalho de Kusters et al. (1997) investiga o conceito de confiabilidade em sistemas de software embarcado, enfatizando a importância de considerar a perspectiva do usuário. Os autores argumentam que as definições clássicas de confiabilidade são insuficientes para expressar a confiabilidade percebida pelos usuários porque ignoram a visão do mesmo

sobre o que constitui um sistema confiável. Os autores argumentam que diferentes usuários possuem diferentes expectativas de confiabilidade baseadas, por exemplo, em suas funções, objetivos e como interagem com o sistema. Para lidar com a dificuldade em se estabelecer, bem como esclarecer os requisitos de confiabilidade, o artigo propõe um novo modelo denominado *Multi-Party Chain* (MPC).

O MPC surge como uma ferramenta para o desenvolvimento de software, com o objetivo de facilitar a comunicação e garantir que a confiabilidade final atenda a todos os envolvidos. O MPC inicia identificando todos os stakeholders (desenvolvedores, usuários, clientes etc.) e coletando seus requisitos específicos de confiabilidade. Esses requisitos são então traduzidos para uma linguagem comum, analisados para identificar conflitos e, a partir deles, definidas métricas para avaliar a confiabilidade ao longo do desenvolvimento. O MPC permite ainda o monitoramento dos requisitos, garantindo que a confiabilidade se mantenha como prioridade. Benefícios como comunicação clara e objetiva, melhora no entendimento dos requisitos, e o desenho centrado no usuário, proporcionam produtos mais confiáveis, uma vez que é estabelecido um processo de monitoramento e controle para o desenvolvimento de software e que atenda aos objetivos estabelecidos.

O MPC foi aplicado em um estudo de caso na Schlumberger RPS, empresa internacional de tecnologia que atua na área de operações petrolíferas, e demonstra a aplicação e a efetividade do modelo na identificação dos motivos subjacentes aos requisitos dos interessados por meio de testes de campo. A Schlumberger utiliza sistemas embarcados em suas bombas de combustíveis, exigindo alta confiabilidade e disponibilidade 24/7. O estudo aplicou o modelo MPC para analisar e identificar os requisitos de confiabilidade subjacentes e revelou que a demanda por confiabilidade de seus usuários originou diversos requisitos, incluindo:

- ❑ Disponibilidade: o sistema precisa estar em operação contínua para evitar perdas de receita;
- ❑ Maturidade: o sistema precisa ser estável e livre de bugs antes da implementação em larga escala;
- ❑ Ausência de Falhas Conhecidas: o sistema não deve apresentar falhas graves ou recorrentes;
- ❑ Suporte em Caso de Falha: o cliente precisa ter acesso a suporte rápido e eficaz em caso de falhas;
- ❑ Recuperabilidade: o sistema precisa ser capaz de se recuperar de falhas sem perda significativa de dados ou tempo de operação;
- ❑ Atender a Requisitos Implícitos: o sistema precisa atender a expectativas não explicitamente definidas pelos usuários.

O estudo demonstrou que o modelo MPC é uma ferramenta eficaz para identificar os diversos requisitos de confiabilidade dos usuários e entender os motivos por trás de suas demandas. Essa compreensão permite que as empresas desenvolvam soluções mais adequadas às necessidades dos usuários, reduzindo custos e aumentando a satisfação. O estudo conclui que a confiabilidade em sistemas embarcados deve ser definida a partir de uma perspectiva centrada no usuário, considerando os diversos requisitos e expectativas dos usuários finais. O modelo MPC é uma ferramenta que auxiliar nesse processo, facilitando a comunicação entre as partes interessadas e garantindo que o produto atenda às necessidades de todos. O autor conclui que é importante que pesquisas futuras se concentrem em aprimorar o modelo MPC com novos estudos de caso, através de pesquisas empíricas e desenvolvimento de ferramentas a fim de apoiar a abordagem centrada no usuário na definição de requisitos de confiabilidade.

O trabalho de Kusters et al. (1997) enfatiza a importância de considerar a confiabilidade em sistemas pela perspectiva do usuário, e apresenta uma proposta de modelo para o levantamento de requisitos dos sistema que visa atender as necessidade do usuário. O trabalho de Yabe et al. (2018), que investigou os fatores que afetam a percepção de confiabilidade do usuário durante sua experiência de uso, buscou encontrar fundamentos que pudessem ser utilizados para classificar a importância de diferentes tipos de falhas, seguindo um modelo semelhante ao proposto por Jalote et al. (2004). Embora todos esses estudos tenham contribuído para a compreensão da importância da confiabilidade baseada na interação dos usuários, apenas o estudo de Yabe et al. (2018) se aprofundou nos fatores que afetam a experiência do usuário. Este trabalho pretende explorar outros fatores que influenciam a experiência do usuário, especialmente sob a perspectiva da tolerância dos usuários.

Tabela 1 – Resumo dos trabalhos relacionados sobre confiabilidade percebida

Referência (Autor, Ano)	Objetivo Principal	Metodologia	Métricas Utilizadas	Cenários de Aplicação	Resultados/Relevância	Limitações	Atributos que Impactam a Confiabilidade
(CHEUNG, 1980)	Avaliar a confiabilidade de software com base em componentes mais frequentemente utilizados	Modelagem com cadeias de Markov e comportamento de uso	Confiabilidade de componentes, Ação do usuário	Sistemas computacionais interativos	Proposta de um modelo que avalia a confiabilidade com base nos componentes mais utilizados, reduzindo o impacto das partes menos confiáveis do sistema.	Avaliação restrita ao comportamento de uso mais frequente; não considera o impacto emocional ou de falhas raras.	Confiabilidade de componentes, frequência de utilização, ativação de componentes relevantes pelo comportamento do usuário
(JALOTE et al., 2004)	Incorporar a percepção do usuário na avaliação da confiabilidade de software	Modelagem com pesos diferentes para tipos de falhas	QoE, Tipo de falha, Impacto do contexto	Produtos de software de massa	Modelo baseado em pesos atribuídos a falhas específicas para melhorar a experiência do usuário em termos de confiabilidade.	Abordagem limitada a produtos de software de massa, sem generalização para outros sistemas.	Impacto da falha, contexto da falha, percepção de diferentes grupos de usuários
(KUSTERS et al., 1997)	Introduzir um modelo centrado no usuário para levantar requisitos de confiabilidade	Estudo de caso com o modelo Multi-Party Chain (MPC)	Confiabilidade percebida, requisitos de stakeholders	Sistemas embarcados (indústria petrolífera)	O modelo MPC facilita a comunicação e o levantamento de requisitos, melhorando a confiabilidade com base nas necessidades de diferentes stakeholders e no monitoramento contínuo.	Aplicação limitada a sistemas embarcados industriais; pode não se aplicar a outros tipos de software.	Disponibilidade, maturidade do sistema, ausência de falhas conhecidas, recuperação, suporte em caso de falha, requisitos implícitos dos usuários
(WANG; TRIVEDI, 2009)	Definir a confiabilidade percebida do serviço com base no comportamento do usuário	Modelagem com base em métricas de uso e cadeias de Markov	Confiabilidade percebida, comportamento do usuário	Serviços em nuvem compartilhados	Proposta de um modelo baseado em comportamento do usuário para calcular a confiabilidade e melhorar a alocação de recursos no cluster.	Não aborda diretamente fatores emocionais ou subjetivos que podem impactar a percepção do usuário.	Número de solicitações, tempos de resposta, comportamento do usuário, utilização de partes do sistema
(KUIPERS et al., 2010)	Relacionar medidas de QoS com a percepção de QoE em diferentes contextos	Avaliação experimental, correlação entre QoS e QoE	QoE, QoS, Latência, Jitter, Sincronização	Áudio, vídeo e jogos online	Estabelecimento de uma correlação entre QoS e QoE em diferentes contextos de uso, prevendo a QoE com base em parâmetros de QoS.	Grupo de participantes reduzido; limitações na generalização para outros contextos de uso.	Latência, variação de latência (jitter), qualidade de rede, sincronização áudio-vídeo, características do conteúdo
(YABE et al., 2018)	Explorar o impacto de falhas do SO na percepção de confiabilidade dos usuários	Questionários e simulação de falhas	QoE, Tipo de falha, Expectativa	Microsoft Windows 7 (Sistema Operacional)	Fatores como consequência da falha, prazos apertados e recuperação são mais relevantes que o tipo de falha ou nível de conhecimento do usuário.	Não considera falhas em sistemas críticos.	Consequência da falha, recuperação, formato da mensagem de falha, frequência e familiaridade com falhas
(ASSUNÇÃO, 2019)	Integrar fatores quantitativos e qualitativos para avaliar a percepção de QoE	Questionários com cenários realistas	QoE, QoS, MTBF	Software geral	Modelo analítico centrado na QoE. Dimensões emocionais como estresse e paciência são cruciais para a percepção de confiabilidade.	Avaliação limitada a falhas simuladas em cenários específicos.	Estresse, pontualidade, paciência, frequência de falhas, impacto emocional

3.2 Estudos relacionados a Qualidade de Serviços

Prakash et al. (2019) realizaram uma revisão da literatura sobre qualidade de serviços (SQ), com o objetivo de examinar as questões envolvidas em sua conceituação e operacionalização, identificando mais de 60 modelos de SQ. Os autores destacam que o mercado está cada vez mais orientado para o cliente e enfatizam que a chave para obter uma vantagem competitiva sustentável está na prestação de serviços de alta qualidade, resultando em clientes satisfeitos. A pesquisa evidencia que a SQ possui uma forte relação com o desempenho dos negócios, redução de custos, satisfação e fidelidade do cliente, bem como com a lucratividade.

O estudo foi conduzido por meio da revisão de artigos publicados em periódicos e revistas acadêmicas, utilizando palavras-chave frequentemente encontradas na literatura. Ao todo, foram analisados 828 artigos. Em resumo, os autores observaram que a maioria dos estudos sobre SQ concentra-se em três aspectos principais: o “aquilo que” é entregue, que diz respeito à qualidade física ou tangível do serviço; o “como” é entregue, que se refere aos aspectos funcionais ou interativos da qualidade do serviço (fluxo bidirecional do serviço); e o “resultado” do processo de avaliação, que envolve a comparação entre as expectativas do cliente e suas percepções em relação ao serviço efetivamente recebido, ou seja, o paradigma da desconfirmação.

Prakash et al. (2019) reconhecem que medir a qualidade de serviço (SQ) é um desafio em muitos estudos, devido à sua natureza intangível, à complexidade do envolvimento humano, à variabilidade na prestação de serviços, à dificuldade de aplicar padrões e limites ao processo e à identificação de pontos de verificação e controle adequados. Embora as expectativas dos clientes em relação à SQ sejam difíceis de compreender, é amplamente aceito que as percepções de SQ são significativamente influenciadas pela qualidade técnica, que afeta o julgamento do cliente em relação à excelência geral do serviço.

O autor também destaca uma série de modelos focados em serviços digitais, como o e-SQ, que inclui dimensões como bom design, fácil acesso, suporte, velocidade e manutenção. Os serviços digitais são geralmente avaliados com base na qualidade percebida do processamento do site, qualidade do resultado e qualidade da recuperação, os quais influenciam a satisfação do usuário e suas intenções comportamentais. Em conclusão, o autor ressalta que a literatura sobre SQ está repleta de redefinições conceituais, adaptações contextuais e medidas, e que vários contextos discutidos ajudam os profissionais a identificar e adaptar as dimensões de SQ mais relevantes para seus negócios.

Parasuraman, Zeithaml e Malhotra (2005) desenvolveram a escala E-S-QUAL para medir a qualidade do serviço prestado por sites de vendas *online*. A pesquisa resultou em quatro dimensões principais: eficiência, cumprimento, disponibilidade do sistema e privacidade, juntamente com três dimensões adicionais: responsividade, compensação e assistência em casos não rotineiros. Os autores destacam que a Internet continua sendo um canal crucial para a venda de diversos produtos e serviços, e quando os consumidores

enfrentam problemas como transações não concluídas, atrasos ou falhas na entrega de produtos, falta de resposta aos e-mails e dificuldade em acessar informações desejadas, a reputação dos canais *online* é prejudicada. Tanto no âmbito acadêmico quanto empresarial, é evidente que a falta de qualidade adequada no serviço prestado pelas empresas na Internet leva à rejeição pelos clientes. A pesquisa se concentrou exclusivamente em sites de compras *online*, em que os consumidores percebem e avaliam a qualidade do atendimento dado ao cliente nesse ambiente. Para isso, foi aplicado o procedimento meio-fim (*means-end framework*) conduzido por uma empresa terceirizada, que fez entrevistas com clientes de sites populares de vendas *online*.

Parasuraman, Zeithaml e Malhotra (2005) propuseram que a e-SQ se baseasse na literatura tradicional de qualidade de serviço, onde a qualidade do serviço é percebida por meio da comparação entre o desempenho real do serviço e o que deveria ser entregue. Eles citam a pesquisa de Gefen (2002), que adaptou a conceituação do ServQual para o contexto eletrônico, condensando as cinco dimensões tradicionais de qualidade de serviço em três, para serviços *online*: tangíveis; uma dimensão combinada de responsividade; confiabilidade e segurança; e empatia. Os autores também mencionam o eTailQ de Wolfinbarger e Gilly (2003), que confirmou que as classificações de confiabilidade/atendimento (*fulfillment*) eram o preditor mais forte de satisfação e qualidade do cliente, além de ser o segundo preditor mais forte das intenções de recompra em um site. A pesquisa destaca que alguns aspectos hedônicos da experiência, como diversão ou prazer, não puderam ser considerados, mas acredita-se que os benefícios buscados em um site sejam importantes para a qualidade do serviço. Além disso, a relação entre fatores hedônicos e orientados a objetivos também é relevante para a qualidade do serviço.

No E-S-QUAL, Parasuraman, Zeithaml e Malhotra (2005) definem o atendimento (*fulfillment*) como a medida em que as promessas do site em relação à entrega do pedido e disponibilidade do item são cumpridas. Eles ressaltam que essa é uma das facetas mais críticas e igualmente importantes da qualidade do serviço do site, observando algumas sobreposições com estudos anteriores de Wolfinbarger e Gilly (2003) e Gefen (2002). Os autores ainda mencionam que há um número limitado de artigos acadêmicos que abordam diretamente como os clientes avaliam a e-SQ, bem como seus antecedentes e consequências em relação aos aspectos relevantes da SQ tradicional (serviços prestados por pessoas), sugerindo a necessidade de pesquisas adicionais no contexto eletrônico.

Fassnacht e Koese (2006) realizaram um estudo no qual desenvolveram e testaram um modelo hierárquico para a Qualidade dos Serviços Eletrônicos (QES). O processo de pesquisa envolveu uma revisão abrangente da literatura, combinada com um estudo qualitativo, no qual os entrevistados foram convidados a compartilhar suas experiências por meio de respostas a perguntas abertas sobre os serviços eletrônicos em questão. A Figura 7 ilustra o modelo hierárquico proposto por Fassnacht e Koese (2006). A pesquisa concentrou-se em serviços nos quais os clientes pagam diretamente pelo serviço recebido,

embora não tenha abordado todas as diversas finalidades que a Internet pode ter na entrega de serviços. O objetivo dos autores era desenvolver e testar empiricamente uma conceituação que proporcionasse uma compreensão mais abrangente da QES em comparação a estudos anteriores, adotando uma visão mais ampla das dimensões abordadas no tema.

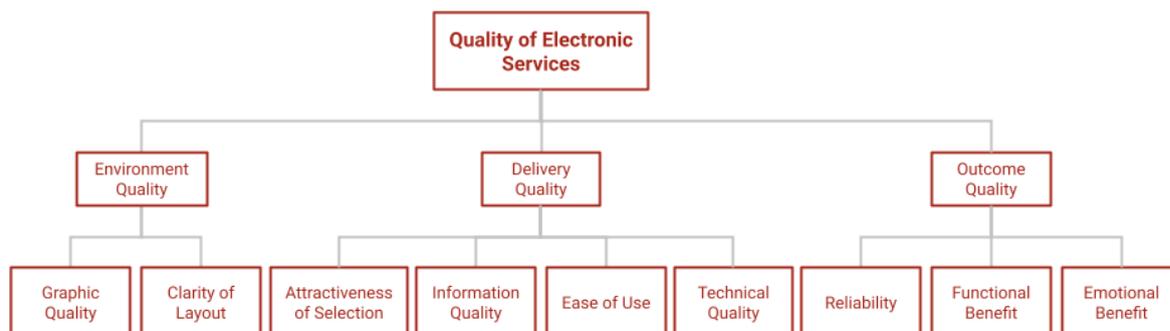


Figura 7 – Modelo conceitual para qualidade dos serviços eletrônicos

Fonte: Fassnacht e Koese (2006)

Os autores destacam duas maneiras importantes nas quais a Qualidade dos Serviços Eletrônicos (QES) difere da qualidade de serviço tradicional. Primeiramente, ocorre um autoatendimento durante a experiência de serviço, uma vez que os clientes interagem apenas com uma interface técnica, resultando em um encontro direto entre pessoa e tecnologia. Além disso, os clientes são imersos em um ambiente de serviço que é decorado por recursos de design específicos da interface gráfica do usuário. Os autores também reconhecem que os consumidores de serviços na Internet estão menos dispostos a tolerar a má qualidade do serviço.

Os autores argumentam que, embora existam várias escalas disponíveis para medir a QES, a maioria delas está limitada a algumas dimensões para capturar a essência do conceito. Essas dimensões variam de acordo com os diferentes contextos nos quais são aplicadas. Por exemplo, perguntas diferentes são aplicadas para avaliar a qualidade em compras online em comparação com jornais online. No entanto, os benefícios buscados pelos clientes podem ser aplicáveis em diferentes contextos, o que ressalta a necessidade de um modelo conceitual abrangente.

No estudo, os autores discutem a possibilidade de obter medidas internas de qualidade, como o tempo de resposta do servidor ou dados de sequência de cliques, relacionados à parte técnica do serviço prestado. No entanto, eles observam que os *insights* obtidos por meio dessas medidas tendem a ser limitados, uma vez que não refletem adequadamente as percepções dos clientes. Para os autores, o resultado do serviço, que ocorre após a entrega

do mesmo e é visto a partir da perspectiva do cliente, é um componente importante da QES. Eles destacam que a confiabilidade é a única faceta do resultado que geralmente é considerada nos estudos existentes. Além disso, os autores lamentam a falta de avaliação dos benefícios hedônicos e a falta de atenção dada à percepção dos usuários no contexto de serviços eletrônicos, principalmente quando ocorre uma falha e é necessário um processo de recuperação do serviço com a ajuda de agentes de atendimento ou comunicação.

Ao avaliarem a importância da dimensão do resultado, especialmente a subdimensão do benefício funcional, os autores ressaltam que a tecnologia é um meio e não um fim em si mesma. Suas análises também revelaram uma considerável sobreposição entre as subdimensões propostas, como facilidade de uso, qualidade de aparência/*layout*, qualidade da informação, privacidade/segurança, confiabilidade, velocidade/capacidade de resposta e conteúdo. Os autores acreditam que a conceitualização adequada e as medições são cruciais para uma gestão eficaz da qualidade do serviço.

A Figura 7 representa o modelo hierárquico para Qualidade dos Serviços Eletrônicos proposto e suas definições:

- ❑ Qualidade do ambiente (*environment quality*): está relacionada à aparência da interface do usuário;
- ❑ Qualidade de entrega (*delivery quality*): diz respeito à interação cliente-máquina durante a utilização do serviço, ou seja, inclui aspectos relevantes para o cliente quando ele está buscando informações, selecionando entre as opções disponíveis ou realizando transações;
- ❑ Qualidade do resultado (*outcome quality*): é a percepção do cliente após a entrega do serviço;
- ❑ Qualidade gráfica (*graphic quality*): captura o quão bem os vários elementos da interface do usuário (por exemplo, texto, ícones, imagens digitais ou planos de fundo) são representados visualmente;
- ❑ Clareza de layout (*clarity of layout*): é definido como o grau em que a estrutura de design da interface do usuário ajuda os usuários a se encontrar na experiência;
- ❑ Atratividade da Seleção (*attractiveness of selection*): é definida como a medida em que a variedade de ofertas disponíveis atrai o cliente. Ela espelha a subdimensão de “conteúdo”;
- ❑ Qualidade da informação (*information quality*): cobre até que ponto informações completas, precisas e oportunas são fornecidas ao cliente durante o processo de interação com a interface do usuário (por exemplo: descrições de produtos, informações de pagamento ou perguntas frequentes);

- ❑ Facilidade de Uso (*ease of use*): é definido como o grau em que a funcionalidade da interface do usuário facilita a recuperação do serviço eletrônico pelo cliente;
- ❑ Qualidade Técnica (*technical quality*): captura a qualidade da transferência e processamento de dados durante a entrega do serviço eletrônico. Ela engloba vários elementos técnicos do processo de entrega, que em nossa opinião pertencem um ao outro, mas são encontrados em duas dimensões de qualidade diferentes: velocidade/capacidade de resposta e privacidade/segurança;
- ❑ Confiabilidade (*reliability*): é definida também como cumprimento (*fulfillment*), que por sua vez é definida como a medida em que o provedor mantém sua promessa de serviço. É importante observar que esta subdimensão não se refere ao funcionamento confiável da infraestrutura técnica do provedor durante a prestação do serviço, pois esse aspecto já está coberto pela qualidade técnica;
- ❑ Benefício funcional (*functional benefit*): é definido como a medida em que o serviço atende ao seu propósito real;
- ❑ Benefício Emocional (*emotional benefit*): é o grau em que o uso do serviço desperta sentimentos positivos.

Diversos estudos têm avançado no campo da qualidade de serviço, principalmente na compreensão de que um software é desenvolvido para fornecer um serviço aos clientes. Prakash et al. (2019) destacaram que a qualidade de serviço é um conceito complexo, levando a modelos com sobreposições ou omissões, a depender do contexto em que é aplicado. No entanto, eles reconhecem que é difícil encontrar métricas que avaliem os aspectos intangíveis devido à sua natureza subjetiva e ao alto envolvimento de elementos humanos.

O E-S-QUAL proposto por Parasuraman, Zeithaml e Malhotra (2005) concentrou-se principalmente no contexto de sites de vendas online, o que levou a um foco maior em métricas de qualidade técnica do que em métricas intangíveis. O estudo também redefiniu a confiabilidade como “atendimento/cumprimento” (*fulfillment*). A classificação das principais métricas em categorias, como Prakash et al. (2019), com “o que”, “como” e “resultado”; Gefen (2002) com as dimensões “tangíveis” e “empatia”, e Fassnacht e Koese (2006), com “ambiente”, “entrega” e “resultado”, demonstra a importância de separar as perspectivas em que as métricas são representativas e indica que a qualidade é a representação dessa hierarquização.

A confiabilidade/atendimento (*fulfillment*), conforme evidenciado por Wolfinbarger e Gilly (2003) e Prakash et al. (2019), é considerada o fator mais influente na satisfação e na qualidade percebida pelos clientes. Fassnacht e Koese (2006), por sua vez, acreditam na criação de um modelo de propósito geral, mas realizaram um avanço significativo ao distinguir a qualidade de entrega da qualidade de resultado, considerando confiabilidade,

aspectos funcionais e emocionais como componentes essenciais no estudo da confiabilidade percebida pelos usuários.

Vários estudos têm como base o modelo ServQual, um dos primeiros a estabelecer uma metodologia para medir métricas intangíveis por meio da comparação entre expectativas e percepções do serviço efetivamente recebido, ou seja, seguindo o paradigma da desconfirmação.

Tabela 2 – Resumo dos estudos relacionados sobre SQ

Referência (Autor, Ano)	Objetivo Principal	Metodologia	Métricas Utilizadas	Cenários de Aplicação	Resultados/Relevância	Limitações
(GEFEN, 2002)	Adaptação do modelo SERVQUAL ao contexto eletrônico	Pesquisa quantitativa com consumidores	Tangíveis, responsividade, confiabilidade, segurança, empatia	Serviços online em geral	SERVQUAL adaptado para ambiente eletrônico, com fortes predições de qualidade	Foco em serviços online, sem capturar fatores emocionais
(WOLFINBARGER; GILLY, 2003)	Avaliação da confiabilidade e atendimento em websites de e-commerce	Pesquisa empírica com usuários de e-commerce	Confiabilidade, atendimento, segurança	Sites de vendas online	Confiabilidade é o preditor mais forte de satisfação e intenções de recompra	Aplicação limitada ao contexto de e-commerce
(PARASURAMAN; ZEITHAML; MALHOTRA, 2005)	Desenvolvimento da escala E-S-QUAL para medir a qualidade de serviços online	Questionários e entrevistas com consumidores	Eficiência, cumprimento, disponibilidade do sistema, privacidade	Sites de vendas online	E-S-QUAL ajuda a prever satisfação e intenções de recompra	Foco limitado em métricas hedônicas
(FASSNACHT; KOESE, 2006)	Desenvolvimento e teste de um modelo hierárquico para Qualidade dos Serviços Eletrônicos (QES)	Revisão de literatura e estudo qualitativo	QES, confiabilidade, benefício emocional	Diversos serviços eletrônicos pagos	Proposta de um modelo hierárquico de QES, diferenciando entrega e resultado	Falta de avaliação de aspectos hedônicos e percepção durante recuperação de falhas
(PRAKASH et al., 2019)	Revisão da literatura sobre qualidade de serviços (SQ)	Revisão de artigos acadêmicos	Modelos de SQ: SERVQUAL, e-SQ	Diversos contextos de serviços digitais	Identificou mais de 60 modelos de SQ; foco no desempenho, satisfação e fidelidade do cliente	Desafios na medição devido à natureza intangível da SQ

Material e Métodos

4.1 Introdução

Este capítulo descreve o desenho do estudo exploratório sobre a tolerância dos usuários a diferentes níveis de confiabilidade de software, detalhando as ferramentas e procedimentos utilizados para a coleta e análise de dados.

O estudo em questão avalia a crescente dependência de softwares em diversos setores da sociedade, com o intuito de aprofundar a compreensão sobre como os usuários avaliam a confiabilidade desses sistemas e reagem a falhas. Compreender esses aspectos é fundamental para identificar os fatores que tornam algumas falhas mais toleráveis do que outras, proporcionando informações valiosas para o desenvolvimento de softwares que ofereçam uma melhor Qualidade da Experiência (QoE) e, conseqüentemente, aumentem a satisfação dos usuários com os serviços prestados.

A seção 4.2 descreve os recursos utilizados na coleta de dados, incluindo a plataforma online empregada para o questionário, a estrutura e o conteúdo do questionário, bem como as questões éticas envolvidas e as características da amostra. O principal objetivo dessa etapa é identificar os fatores chave que influenciam a tolerância dos usuários a falhas de software. Para isso, definiu-se o conceito de “tolerância” neste contexto e estabeleceu-se uma metodologia para quantificar e observar o nível de tolerância dos usuários.

A metodologia proposta avalia a tolerância dos usuários a falhas de software por meio de um experimento mental com múltiplos cenários. Nesses cenários hipotéticos, os usuários são convidados a imaginar como lidariam com diferentes falhas e a responder a perguntas sobre suas percepções. Para cada cenário apresentado, são coletadas três métricas principais: aceitabilidade, nível de satisfação e QoE.

Adicionalmente, para avaliar o conceito de tolerância a falhas, foi aplicada uma escala Likert elaborada com base nos fatores relacionados a essa tolerância. O objetivo da utilização desta escala é permitir que os usuários expressem seu nível de concordância ou discordância em relação a diferentes afirmações, no que tange à sua percepção de cada fator ao lidar com falhas.

A seção 4.3 detalha a abordagem da pesquisa e as técnicas de análise de dados. Essa seção especifica os procedimentos de coleta de dados e descreve o procedimento da análise realizada. Para todas as análises foi feito o uso do software Jamovi¹. Para a análise dos dados fez-se uso da técnica de Análise de Variância (ANOVA) não paramétrica. Ademais, explicar-se-á como os resultados serão apresentados e interpretados.

Ao comparar os dados coletados entre os diferentes cenários e grupos, a pesquisa visa identificar os fatores que mais influenciam a tolerância a falhas, determinar a relação entre a tolerância e a satisfação do usuário, e analisar a relação entre a tolerância e outras variáveis de interesse na pesquisa.

A pesquisa também avaliará a experiência do usuário em cada cenário de falha. Através dessa avaliação, será possível analisar como os fatores de tolerância impactam a QoE e identificar maneiras de minimizar os efeitos negativos das falhas, aprimorando a qualidade da experiência do usuário.

Finalmente, o estudo investigará a relação entre os fatores de tolerância e a QoE, explorando a correlação entre a aceitabilidade das falhas e a satisfação do usuário. A partir dessa investigação, pretende-se determinar a viabilidade de definir um valor de referência para o nível de tolerância, indicando um limite abaixo do qual a QoE não é significativamente afetada.

O objetivo deste capítulo é descrever o desenvolvimento da metodologia proposta e assegurar a credibilidade da pesquisa, fornecendo informações detalhadas que permitam a replicação do estudo. Este relato minucioso garante que os leitores compreendam como a pesquisa foi conduzida e possam avaliar a qualidade dos resultados obtidos.

No próximo capítulo (Capítulo 5), serão apresentados os resultados dos experimentos realizados, oferecendo uma análise detalhada das descobertas obtidas ao longo desta pesquisa.

4.2 Material

Esta seção descreve o material de pesquisa desenvolvido e como ele se integra à metodologia apresentada na seção 1.5, permitindo uma melhor compreensão dos fatores que influenciam a tolerância dos usuários a falhas de software. O principal material de pesquisa é um questionário, descrito na subseção 4.2.4, com múltiplos cenários detalhados na subseção 4.2.3. Esse questionário apresenta aos participantes diferentes situações hipotéticas de falhas de software, meticulosamente elaboradas para explorar os diversos fatores de tolerância descritos na Figura 1.

Ao utilizar esse questionário, a metodologia ganha profundidade na análise da percepção dos usuários em relação a falhas, variando um fator de tolerância em cada cenário. A coleta de dados sobre como os participantes reagem aos diferentes cenários possibilitará a

¹ Disponível em: <https://www.jamovi.org/>

identificação dos fatores que exercem mais influência na tolerância, utilizando uma escala definida na subseção 4.2.2. Além disso, o estudo avaliará o impacto da tolerância na QoE. O questionário incluirá perguntas específicas sobre a aceitabilidade e a satisfação dos participantes com o software em cada cenário, permitindo relacionar a tolerância com a QoE.

A subseção 4.2.1 descreve as principais ferramentas utilizadas e as restrições às quais a pesquisa foi submetida para atender aos requisitos estabelecidos pelo Comitê de Ética em Pesquisa (CEP) e Comissão Nacional de Ética em Pesquisa (CONEP). Essas ferramentas e restrições garantem a conformidade com os padrões éticos e metodológicos necessários para a credibilidade e replicabilidade da pesquisa.

4.2.1 Ambiente de execução

A construção do ambiente de execução foi orientada pelas diretrizes detalhadas sobre a condução de experimentos que envolvem a coleta de dados das percepções humanas, conforme estabelecido pelo Comitê de Ética em Pesquisa (CEP). O projeto de pesquisa foi desenvolvido e submetido para avaliação na Plataforma Brasil. A análise ética do CEP, relativa ao protocolo de pesquisa, resultou em um parecer favorável, conforme Anexo B.0.2, garantindo a conformidade com os princípios éticos no processo de coleta e uso dos dados.

O experimento conduzido nesta pesquisa foi realizado por meio de questionários online utilizando a plataforma Google Forms². A escolha do Google Forms foi motivada por sua praticidade, acessibilidade e recursos abrangentes, que possibilitaram a criação de um questionário estruturado e eficiente, atendendo aos requisitos mínimos de segurança e confidencialidade. O Anexo B.0.3 contém a extração do formulário criado na plataforma.

Os participantes foram convidados eletronicamente através do convite disponível no Anexo B.0.1, utilizando listas de e-mails e redes sociais. Eles receberam informações sobre o propósito da pesquisa, juntamente com instruções para acessar o formulário. Para garantir a confidencialidade dos dados pessoais dos participantes, a pesquisa foi realizada de forma anônima, em conformidade com as regulamentações e requisitos de proteção de dados, oferecendo a possibilidade de não responder a alguma questão ou abandonar o formulário a qualquer momento. As informações relevantes para o participante estão descritas no Termo de Consentimento Livre e Esclarecido (TCLE), onde apenas após a concordância com os termos apresentados, o participante tem acesso às questões do formulário.

Apesar do rigor metodológico adotado na pesquisa, alguns vieses precisam ser considerados ao interpretar os resultados. O viés de seleção, resultante do recrutamento de participantes por listas públicas e redes sociais, pode ter gerado uma amostra não

² Disponível em: <https://workspace.google.com/intl/pt-BR/lp/forms/>

representativa, favorecendo usuários mais experientes em tecnologia. O viés de resposta também pode ter ocorrido, com participantes respondendo de forma socialmente desejável ou interpretando incorretamente algumas perguntas. Esses vieses limitam a generalização dos resultados, sugerindo a necessidade de estudos complementares com uma amostra mais diversificada.

Um teste piloto foi realizado com outros pesquisadores para avaliar a clareza das questões do questionário e a compreensão por parte dos participantes. Essa etapa foi importante para garantir que o formulário estivesse bem formulado e que as instruções fossem facilmente compreendidas, visando à remoção de possíveis vieses ou limitações da plataforma de formulários. Os dados coletados no teste piloto foram descartados.

Com base em pesquisas anteriores (YABE et al., 2018), não foram identificadas diferenças significativas entre métodos de pesquisa presencial e online. Dessa forma, esta pesquisa optou por coletar dados apenas por meio de questionários online, permitindo um maior alcance e diversidade de participantes.

4.2.2 Conceito de tolerância

Este trabalho tem como objetivo investigar fatores que afetam a tolerância do usuário a falhas de software. A revisão da literatura identificou os principais elementos relacionados à Tolerância (Figura 1) e, para avaliar a influência desses fatores na percepção do usuário, serão utilizados cenários de falha realistas, e um questionário coletará opiniões e percepções a respeito do efeito percebido das falhas.

A coleta da percepção do usuário utiliza uma escala triádica de Likert, elaborada com a finalidade de investigar fatores que afetam a tolerância do usuário a falhas de software. Essa técnica, introduzida por Likert (1932), é amplamente utilizada em pesquisas sociais, por possibilitar a avaliação equilibrada de atitudes e percepções, além de oferecer uma opção neutra para expressar incerteza.

Para atender ao objetivo de determinar um conceito de tolerância, foram utilizados como insumo para a elaboração da escala os fatores representados na Figura 1, que seguem uma abordagem de decomposição por meio de afirmações positivas e negativas, conforme descrito na Tabela 3.

A escala triádica de Likert foi aplicada por meio de um formulário de seleção que adiciona um ponto para a afirmativa positiva, subtrai um ponto para a afirmativa negativa e considera o valor neutro como zero quando o participante não seleciona nenhuma alternativa. Caso o participante selecione tanto uma afirmativa positiva quanto uma negativa, o resultado é considerado neutro. Essa metodologia permite uma análise detalhada e equilibrada das percepções dos usuários, capturando nuances nas respostas que outras escalas podem não identificar.

O atributo Resultado / Atributos de Qualidade avalia a qualidade geral do software em relação ao seu desempenho e funcionalidade, medindo se o software atende aos padrões

Tabela 3 – Afirmativas positivas e negativas das dimensões de Tolerância

Fator	Positiva	Negativa
Resultado / Atributos de Qualidade	A qualidade do software atende às minhas necessidades em relação ao meu objetivo.	A qualidade do software é insuficiente para atender às minhas necessidades com relação ao meu objetivo.
Resultado / Atingimento / Confiabilidade	Tenho confiança de que conseguirei alcançar meus objetivos utilizando esse software.	É improvável que eu consiga alcançar meus objetivos com esse software.
Fatores Emocionais / Utilidade e Necessidade	As funcionalidades desse software me proporcionam uma sensação de realização.	As funcionalidades desse software dificultam que eu atinja meus objetivos.
Fatores Emocionais / Crenças / Capacidade	Acredito que consigo operar esse software.	Esse software é muito difícil e tenho dificuldade em operá-lo.
Fatores Emocionais / Promessas	As funcionalidades do software estão alinhadas com as promessas feitas para esse tipo de produto.	Esse software deixa a desejar em relação a algumas das funcionalidades prometidas para esse tipo de produto.
Fatores de Economicidade / Valor Percebido	Esse software me proporciona benefícios claros.	Os benefícios desse software são inferiores às minhas expectativas.
Fatores de Economicidade / Esforço Necessário	Considero que esse software reduz meu esforço quando escolho utilizá-lo.	A utilização desse software demanda muito esforço.
Fatores Mercadológicos / Alternativas do Mercado	Esse modelo de software atende perfeitamente às minhas necessidades.	Acredito que há opções mais apropriadas no mercado que estou disposto(a) a considerar.
Fatores Mercadológicos / Experiências Passadas	Minha experiência prévia com esse tipo de software me permitiu utilizá-lo com eficácia.	Tive experiências frustrantes com esse tipo de software no passado.
Contexto / Impacto	Acredito que as falhas desse software são superáveis e causam um impacto mínimo em mim.	Acredito que as falhas desse software me afetam significativamente.
Contexto / Prazo	Esse software me ajuda a concluir meus objetivos no menor tempo possível.	Quando uso esse software, tenho mais trabalho e preciso de tempo extra para concluir meus objetivos com sucesso.

esperados pelos usuários para cumprir seus objetivos. Um exemplo positivo seria: “O editor de texto oferece todas as ferramentas de formatação de que preciso para preparar documentos profissionais”. Um exemplo negativo seria: “As características do editor de texto não são boas suficiente para que eu faça um bom trabalho”.

O atributo Resultado / Atingimento / Confiabilidade mede a confiança dos usuários na capacidade do software de ajudá-los a atingir seus objetivos, avaliando a percepção de confiabilidade e eficácia do software. Um exemplo positivo seria: “Confio que o editor de texto não terá problemas quando eu estiver revisando documentos importantes”. Um exemplo negativo seria: “É improvável que eu consiga terminar meu relatório a tempo, pois o editor de texto é muito instável”.

O atributo Emocional / Utilidade e Necessidade avalia a utilidade do software e sua capacidade de satisfazer as necessidades dos usuários, proporcionando uma sensação de realização. Um exemplo positivo seria: “Usar as ferramentas de formatação e design do editor de texto me faz sentir produtivo e eficiente”. Um exemplo negativo seria: “A falta

de recursos avançados no editor de texto dificulta a criação de documentos complexos”.

O atributo Emocional / Crenças / Capacidade mede a autoconfiança dos usuários em sua capacidade de utilizar o software de forma eficaz, avaliando a percepção de facilidade de uso. Um exemplo positivo seria: “Eu me sinto confortável usando todas as funções do editor de texto, mesmo as mais avançadas”. Um exemplo negativo seria: “Eu sempre tenho dificuldades para encontrar as funções que preciso no editor de texto”.

O atributo Emocional / Promessas avalia se o software cumpre as promessas feitas durante a promoção do produto, medindo a congruência entre expectativas e realidade. Um exemplo positivo seria: “O editor de texto oferece todas as funcionalidades de que eu preciso, como prometido nas especificações do produto”. Um exemplo negativo seria: “O editor de texto não possui todas as ferramentas de edição avançada que foram anunciadas”.

O atributo Economicidade / Valor Percebido avalia a percepção dos usuários sobre os benefícios claros proporcionados pelo software, justificando seu uso em relação ao custo. Um exemplo positivo seria: “O editor de texto me ajuda a economizar tempo e melhorar a qualidade dos meus documentos”. Um exemplo negativo seria: “Os recursos oferecidos pelo editor de texto não justificam o preço pago”.

O atributo Economicidade / Esforço Necessário mede o esforço necessário para utilizar o software, avaliando se ele facilita ou dificulta as tarefas dos usuários. Um exemplo positivo seria: “O editor de texto automatiza muitas tarefas, o que me economiza bastante tempo”. Um exemplo negativo seria: “Cada tarefa simples no editor de texto leva muito tempo e esforço para ser concluída”.

O atributo Mercadológico / Alternativas do Mercado avalia a percepção dos usuários sobre a adequação do software em comparação com outras opções disponíveis no mercado. Um exemplo positivo seria: “Eu prefiro usar este editor de texto porque ele tem as melhores funcionalidades em comparação com outros no mercado”. Um exemplo negativo seria: “Estou considerando mudar para outro editor de texto, pois acredito que existem opções melhores”.

O atributo Mercadológico / Experiências Passadas mede a influência das experiências anteriores dos usuários com software similares na percepção do software atual. Um exemplo positivo seria: “Por já ter usado outros editores de texto, achei fácil me adaptar a este”. Um exemplo negativo seria: “Minhas experiências anteriores com editores de texto foram frustrantes, o que me deixa receoso ao usar este”.

O atributo Contexto / Impacto avalia a percepção dos usuários sobre o impacto das falhas do software em suas atividades. Um exemplo positivo seria: “Mesmo quando o editor de texto falha, consigo facilmente recuperar meu trabalho”. Um exemplo negativo seria: “Quando o editor de texto falha, perco muito do meu trabalho e tempo, o que é frustrante”.

O atributo Contexto / Prazo mede a eficiência do software em ajudar os usuários a concluir suas tarefas no menor tempo possível. Um exemplo positivo seria: “Com o

editor de texto, consigo terminar meus documentos mais rapidamente do que antes”. Um exemplo negativo seria: “Usar este editor de texto me faz perder tempo, pois preciso de mais trabalho para finalizar meus documentos”.

A Tabela 3 apresenta uma lista de afirmativas positivas e negativas que foram formuladas a fim de avaliar os atributos de tolerância do usuário a falhas de software. Cada atributo é utilizado para a identificação e formulação de fatores que ajudam a compreender a percepção dos usuários sobre diferentes aspectos do software em situação de falha. A aplicação da escala triádica de Likert permite uma análise fatorial exploratória das percepções dos usuários, capturando nuances nas respostas, o que ajuda a identificar áreas de melhoria e a desvendar aspectos que podem tornar o software mais confiável e com melhor qualidade de experiência para os usuários.

A escala Likert de tolerância foi projetada para capturar de forma quantitativa a percepção dos usuários sobre diversos fatores associados à tolerância a falhas de software. Ao utilizar essa escala, os participantes podem expressar seu grau de concordância ou discordância em relação a uma série de afirmações que medem aspectos humanos, técnicos e econômicos, permitindo a análise sistemática das variáveis que influenciam a tolerância. Essa abordagem é particularmente eficaz para responder à pergunta de pesquisa, que busca identificar os principais fatores que afetam a tolerância e como esses fatores se associam entre si.

Por exemplo, a escala Likert pode incluir afirmativas que avaliem a familiaridade do usuário com o software, como “Eu me sinto confortável utilizando todas as funções do software”, o que permite mensurar como a experiência e o conhecimento técnico impactam a disposição dos usuários em aceitar falhas. Da mesma forma, afirmações relacionadas à importância da tarefa, como “As falhas do software me impedem de concluir tarefas críticas”, podem ajudar a avaliar como o contexto de uso e a gravidade das falhas afetam a tolerância.

Essas afirmativas, uma vez respondidas pelos participantes, são organizadas em fatores como confiabilidade percebida, esforço necessário e valor econômico percebido, que são avaliados por meio de análises estatísticas, como a Análise Fatorial Exploratória (AFE). A AFE permite identificar correlações entre os fatores medidos pela escala Likert, respondendo à hipótese de que categorias como fatores humanos, fatores técnicos, e fatores econômicos influenciam diretamente o nível de tolerância a falhas de software.

Assim, ao usar a escala Likert, a pesquisa não só quantifica a tolerância a falhas de software de forma objetiva e sistemática, mas também revela associações entre esses fatores. Isso permite uma compreensão mais profunda de como diferentes aspectos da experiência do usuário — como o esforço para contornar falhas, a confiança na estabilidade do software, e a percepção de valor econômico — influenciam a tolerância, alinhando-se diretamente com a hipótese de que múltiplas categorias de fatores estão envolvidas na forma como os usuários lidam com falhas.

A validade de constructo, que avalia se os instrumentos utilizados medem corretamente os conceitos teóricos do estudo, enfrenta desafios significativos neste contexto, especialmente na definição e medição da tolerância a falhas. Tolerância a falhas é um conceito multifacetado, influenciado por fatores humanos, técnicos e emocionais, e capturá-lo em sua totalidade por meio de uma escala de Likert pode ser complexo. Embora a pesquisa tenha utilizado uma escala Likert triádica, validada por pré-teste com afirmações positivas e negativas para melhorar a precisão, essa abordagem pode não captar as nuances mais sutis das percepções dos usuários. Por exemplo, a tolerância a falhas pode variar com base em características pessoais, como resiliência e experiência prévia com tecnologias semelhantes, que não são diretamente abordadas pela escala. Além disso, a escala Likert, por ser baseada em respostas quantitativas, pode deixar de fora nuances emocionais ou contextuais que influenciam as respostas. Dessa forma, ainda que a escala tenha sido cuidadosamente desenhada para mitigar esses desafios, ela não elimina completamente o risco de não capturar toda a complexidade envolvida na tolerância a falhas, e variáveis externas não consideradas podem influenciar os resultados, comprometendo parcialmente a validade do constructo.

4.2.3 Cenários de pesquisa

Esta pesquisa utiliza cenários cuidadosamente elaborados para observar diferentes níveis de tolerância e capturar a percepção dos participantes sobre o efeito das falhas em variados contextos. Os participantes serão solicitados a avaliar os cenários propostos e expressar suas preferências, percepções e atitudes por meio de escalas de avaliação disponíveis no questionário. Essa abordagem permite coletar dados sobre como os participantes reagem a diferentes situações de falha e como suas escolhas podem variar com base nos fatores apresentados nos cenários.

Um experimento mental é uma técnica amplamente utilizada na pesquisa psicológica e comportamental, conduzida a fim de explorar como as pessoas reagem e tomam decisões em diferentes cenários hipotéticos. Essa abordagem permite aos pesquisadores manipular variáveis de interesse e observar as respostas dos participantes sem a necessidade de realizar experimentos físicos ou situações reais. No contexto desta pesquisa, um experimento mental será aplicado por meio da criação de cenários variados nos quais os participantes são convidados a imaginar situações específicas relacionadas à ocorrência de falhas de software. Essa abordagem proporciona dados que permitirão avaliar o quanto os fatores influenciam a tolerância dos usuários em uma análise multifatorial, e ajudará a compreender o peso e a relação de cada fator na qualidade da experiência do usuário.

Cada cenário é descrito de forma imparcial, detalhando o passo a passo para que os participantes possam se imaginar neles, contribuindo para um experimento mental. A metodologia do experimento mental, em conjunto com a técnica de múltiplos cenários, é empregada para explorar as percepções e respostas dos participantes de uma maneira mais otimizada. Essa abordagem flexível e controlada é essencial para uma análise aprofundada dos fatores que influenciam a tolerância dos usuários às falhas de software, sem a necessidade de execução prática, possibilitando uma investigação objetiva dos processos mentais e comportamentais envolvidos. A preocupação com o viés é constante ao descrever os cenários, garantindo que eles sejam apresentados de maneira neutra, sem qualquer inclinação que possa influenciar as respostas dos participantes.

De maneira geral, o participante é convidado a se colocar na perspectiva de um estudante universitário com um trabalho acadêmico importante e com prazo próximo, necessitando utilizar um editor de texto para realizar modificações na formatação e impressão de um artigo. Nessa situação, o participante é orientado a utilizar um editor de texto, com os recursos necessários, disponível em seu computador, para realizar a tarefa da forma como está familiarizado. Em seguida, é convidado a experimentar este cenário mentalmente e, diante dessa situação imaginada, responder ao questionário a fim de capturar suas percepções.

A dinâmica de elaboração dos cenários se inicia com a apresentação ao participante do cenário principal, que representa a situação em que não ocorre nenhum impacto e é considerado como grupo controle. A tarefa consiste em abrir o arquivo no editor, realizar

a alteração e, em seguida, utilizar a função de conversão para PDF para imprimir o artigo no formato desejado. Os demais cenários são contextualizados com falhas, onde, por exemplo, o aplicativo é abruptamente encerrado, resultando na perda das alterações, retrabalho, atraso na entrega e impactos em seus planos pessoais.

A Tabela 4 detalha cada um dos cenários e os fatores relacionados com o conceito de tolerância associados à classificação a seguir:

- Cenário 1: o usuário consegue editar o documento normalmente;
- Cenário 2: ocorre uma falha ao editar o texto e o aplicativo reinicializa;
- Cenário 3: ressalta as promessas de uso do aplicativo;
- Cenário 4: ressalta a perda do prazo para envio;
- Cenário 5: ressalta as opções concorrentes do mercado;
- Cenário 5.1: ressalta as opções concorrentes do mercado em relação ao comprometimento financeiro;
- Cenário 5.2: ressalta as opções concorrentes do mercado em relação ao impacto causado com a perda do prazo;
- Cenário 6: ressalta a atualização do aplicativo e o esforço em melhorar as falhas;
- Cenário 6.1: ressalta a atualização do aplicativo e o esforço em melhorar as falhas em relação ao comprometimento financeiro;
- Cenário 7: ressalta a avaliação do usuário frente à reincidência do mesmo problema;
- Cenário 7.1: ressalta a avaliação do usuário frente à reincidência do mesmo problema em relação ao comprometimento financeiro;
- Cenário 7.2: ressalta a avaliação do usuário frente à reincidência do mesmo problema em relação ao esforço de aprendizado;
- Cenário 7.3: ressalta a avaliação do usuário frente à reincidência do mesmo problema em relação à falta de alternativas.

Esses cenários foram elaborados para explorar diversos aspectos relacionados à tolerância do usuário a falhas de software. No Cenário 1, o usuário encontra uma experiência sem falhas, permitindo-nos estabelecer uma base para compreender o uso típico do software. No Cenário 2, introduzimos um elemento de falha, forçando a reinicialização do aplicativo. Isso nos permite avaliar como a tolerância do usuário é afetada quando ocorre uma interrupção inesperada. No Cenário 3, concentramo-nos nas promessas feitas pelo fabricante do aplicativo aos usuários, explorando as dimensões emocionais de sua experiência.

O Cenário 4 destaca a importância do contexto, enfocando a perda de prazo, um fator que pode intensificar a intolerância do usuário. No Cenário 5, examinamos os aspectos mercadológicos, considerando as opções de mercado, o que pode influenciar a tolerância do usuário às falhas. O Cenário 6 se concentra na economicidade, realçando os esforços necessários para realizar a mudança frente à ação do aplicativo para melhorar e corrigir as falhas. Por fim, no Cenário 7, investigamos como a repetição de problemas afeta a tolerância do usuário, explorando a dimensão de recuperação, bem como a familiaridade, frequência e falta de alternativas. Juntos, esses cenários fornecem uma visão dos fatores que influenciam a tolerância do usuário, permitindo-nos entender, comparando entre eles, como eles se relacionam com a QoE, a Satisfação e, por fim, avaliar a Tolerância.

Os cenários com desfecho explícito de falha, presentes nos Cenários 1 a 5, permitem uma comparação direta sobre como as falhas afetam a tolerância dos usuários. Nos demais cenários, o foco é identificar outros aspectos além da falha, como promessas de melhoria, alternativas de mercado e reincidência de problemas. A comparação entre cenários quanto à falha só pode ser realizada nos cenários com desfecho explícito de falha, proporcionando uma análise mais precisa e específica desses eventos.

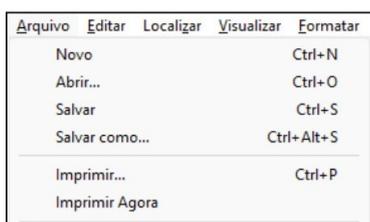
Esses cenários foram estrategicamente desenhados para aplicar diferentes tipos de fatores ao contexto em que a falha é experimentada, permitindo observar, de maneira mensurável, como cada fator afeta a tolerância dos usuários. A variação entre os cenários permitirá testar a hipótese de que o nível de tolerância a falhas de software pode ser quantificado de forma objetiva e sistemática, proporcionando uma comparação entre situações de falha explícita e outros fatores, como promessas de melhoria e alternativas de mercado.

A revisão bibliográfica indicou os principais componentes que moldam a Tolerância, permitindo realizar a avaliação da Tolerância frente a falhas de software. O foco central é compreender os fatores que influenciam a Tolerância dos usuários às falhas de software (discutidos na Figura 1) e como essas falhas impactam na QoE. A proposta é apresentar para o mesmo participante da pesquisa todos os cenários de forma sequencial e incremental; dessa forma, a cada cenário, espera-se a soma de um novo fator à sua experiência.

Tabela 4 – Cenários de pesquisa

Cenário 1 - Edição sem falhas

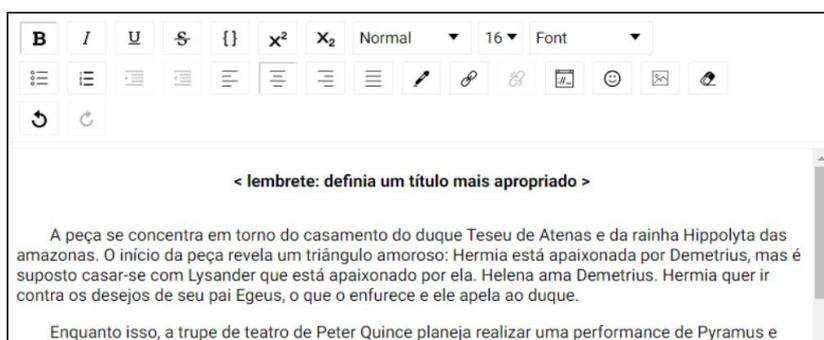
- ❑ Imagine que você é um estudante universitário e está finalizando seu trabalho acadêmico de conclusão de curso.
- ❑ A data de entrega está se aproximando e você precisa garantir que o trabalho esteja bem formatado. Para realizar as tarefas de edição e formatação, você utiliza o editor de texto que você tem mais familiaridade, instalado no seu computador. Imagine que você está abrindo o arquivo com o seu trabalho.



- ❑ Então, você localiza o seu arquivo e aciona a função “Abrir” do software de edição de texto.



- ❑ Com seu documento aberto, você facilmente identifica o ponto que precisa ser alterado e realiza a alteração necessária.



- ❑ Após finalizar a edição, você utiliza a função impressão do arquivo para finalizar a conversão para PDF e salva o arquivo.

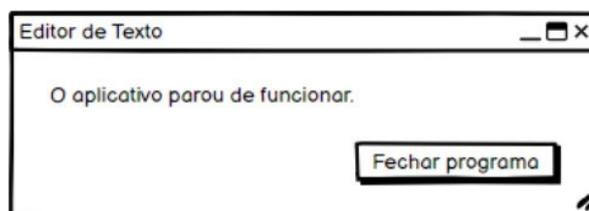


- ❑ Com o trabalho terminado e o arquivo editado e convertido para PDF, você o envia ao seu professor e **conclui sua tarefa com sucesso.**

Continua na próxima página.

Tabela 4 – Continuação da página anterior.**Cenário 2 - Fator Resultado**

- ❑ Imagine que você é um estudante universitário e está finalizando seu trabalho acadêmico de conclusão de curso. A data de entrega está se aproximando. Para garantir que o trabalho esteja bem formatado, você precisa realizar as tarefas de edição e formatação. Para isso, você utiliza o editor de texto instalado no seu computador. Ao abrir o arquivo com o trabalho, você identifica o ponto que precisa ser alterado e realiza a edição necessária.
- ❑ Após realizar alguns ajustes no texto, você se depara com uma situação inesperada. O software para de responder, “trava” e uma mensagem do sistema operacional é exibida informando que “O aplicativo parou de funcionar”.



- ❑ O software é encerrado subitamente e, após reiniciá-lo, você percebe que perdeu todas as alterações realizadas. **Apesar da falha, você consegue refazer as alterações perdidas e enviar o arquivo convertido em PDF ao seu professor dentro do prazo de entrega.**

Continua na próxima página.

Tabela 4 – Continuação da página anterior.

Cenário 3 - Fator Emocional

- ❑ Imagine que você é um estudante universitário e está finalizando seu trabalho acadêmico de conclusão de curso. A data de entrega está se aproximando. Para garantir que o trabalho esteja bem formatado, você precisa realizar as tarefas de edição e formatação. Para isso, você utiliza o editor de texto instalado no seu computador. Ao abrir o arquivo com o trabalho, você identifica o ponto que precisa ser alterado e realiza a edição necessária.
- ❑ Entretanto, após realizar alguns ajustes no texto o software foi encerrado subitamente, causando a perda de todas as alterações realizadas. Apesar de ter conseguido concluir a entrega do trabalho, a utilização do software acabou lhe causando perda de tempo e a necessidade de refazer o trabalho de edição.
- ❑ Neste cenário você se lembra das **promessas de funcionalidades feitas pelo fabricante do software que te fizeram adquirir esse produto (editor de texto)**, dentre elas, **acelerar a sua produtividade**, além da funcionalidade já esperada de edição de textos. No entanto, você percebe que o trabalho repetido causado pela falha acabou na verdade lhe **causando um esforço extra, tornando-o menos produtivo**.

Cenário 4 - Fator Contexto

- ❑ Imagine que você é um estudante universitário e está finalizando seu trabalho acadêmico de conclusão de curso. A data de entrega está se aproximando. Para garantir que o trabalho esteja bem formatado, você precisa realizar as tarefas de edição e formatação. Para isso, você utiliza o editor de texto instalado no seu computador. Ao abrir o arquivo do trabalho, você identifica o ponto que precisa ser alterado e realiza a edição necessária. Entretanto, após realizar alguns ajustes no texto o software foi encerrado subitamente, causando a perda de todas as alterações realizadas, e consequente perda de tempo e a necessidade de refazer o trabalho de edição. Isso quebrou suas expectativas em relação à promessa do produto de acelerar a sua produtividade.
- ❑ Infelizmente, essa situação fez com que **você perdesse o prazo para a entrega do seu trabalho**, resultando na necessidade de repetir a disciplina, o que acontecerá apenas no próximo ano, **impactando seus planos objetivos pessoais**.

Continua na próxima página.

Tabela 4 – Continuação da página anterior.**Cenário 5 - Fator Mercadológico**

- Imagine que você é um estudante universitário e está finalizando seu trabalho acadêmico de conclusão de curso. A data de entrega está se aproximando. Para garantir que o trabalho esteja bem formatado, você precisa realizar as tarefas de edição e formatação. Para isso, você utiliza o editor de texto instalado no seu computador. Ao abrir o arquivo do trabalho, você identifica o ponto que precisa ser alterado e realiza a edição necessária. Entretanto, após realizar alguns ajustes no texto o software foi encerrado subitamente, causando a perda de todas as alterações realizadas, e conseqüente perda de tempo e a necessidade de refazer o trabalho de edição. Isso quebrou suas expectativas em relação à promessa do produto de acelerar a sua produtividade.
- Você decide então **buscar alternativas no mercado para substituir o seu software editor de texto atual**, para que o problema não ocorra novamente, afinal, você **não quer perder o prazo de entrega** do seu trabalho.
- Você encontra opções similares no mercado. No entanto, o custo da licença da melhor alternativa encontrada é **aproximadamente 10 vezes maior que a sua opção atual, o que representa um custo significativo**. É sabido que essa decisão o afetará, pois, além do custo financeiro, você precisará se capacitar para usar o novo software, **processo que demandará um esforço pessoal extra**.

Cenário 5.1 - Comprometimento financeiro

- Agora que você se imaginou nessa situação, responda às perguntas a seguir. Qual sua decisão?
 - Comprar uma das opções de editores de texto do mercado, mesmo ciente dos impactos que terá (financeiro, esforço de aprendizado do novo software);
 - Não compraria no momento e daria uma segunda chance ao editor de texto atual.

Continua na próxima página.

Tabela 4 – Continuação da página anterior.**Cenário 5.2 - Perda do Prazo**

- Mesmo sabendo que a troca do software **provavelmente levará à perda do prazo de publicação**, devido às adaptações necessárias, o que impacta seus planos e objetivos pessoais, qual a sua decisão?
 - Comprar uma das opções de editores de texto do mercado, mesmo ciente dos impactos que terá (financeiro, esforço de aprendizado do novo software);
 - Não compraria no momento e daria uma segunda chance ao editor de texto atual.

Cenário 6 - Fator Economicidade

- Imagine que você é um estudante universitário e está finalizando seu trabalho acadêmico de conclusão de curso. A data de entrega está se aproximando. Para garantir que o trabalho esteja bem formatado, você precisa realizar as tarefas de edição e formatação. Para isso, você utiliza o editor de texto instalado no seu computador. Ao abrir o arquivo do trabalho, você identifica o ponto que precisa ser alterado e realiza a edição necessária. Entretanto, após realizar alguns ajustes no texto o software foi encerrado subitamente, causando a perda de todas as alterações realizadas, e consequente perda de tempo e a necessidade de refazer o trabalho de edição. Isso quebrou suas expectativas em relação à promessa do produto de acelerar a sua produtividade.
- Você está cogitando buscar alternativas no mercado para substituir o seu software editor de texto atual, apesar do custo da licença e o esforço para capacitação. **Porém, você não se decidiu ainda.**
- Ao tentar utilizar novamente seu software editor de texto atual, **você recebe uma mensagem informando que o erro que você encontrou anteriormente foi corrigido** em uma atualização que foi aplicada automaticamente.
- Diante desse cenário de promessa de correção do problema no software atual, por meio da atualização recente, você avalia as opções de **manter o uso do software atual ou adquirir e usar um novo software**, sabendo que a adoção do novo software significaria um **esforço extra que levaria à perda do prazo** para publicar seu artigo, o que vai prejudicar e atrasar seus planos pessoais.

Continua na próxima página.

Tabela 4 – Continuação da página anterior.**Cenário 6.1 - Comprometimento financeiro**

- Agora que você se imaginou nessa situação, responda às perguntas a seguir. Qual sua decisão?
 - Comprar uma das opções de editores de texto do mercado, mesmo me comprometendo;
 - Não compraria no momento e daria uma segunda chance ao editor de texto atual.

Cenário 7 - Recuperação e Aceitabilidade

- Imagine que você é um estudante universitário com um trabalho acadêmico importante prestes a ser entregue e você utiliza um editor de texto para fazer as alterações necessárias na formatação. Porém, o software é encerrado abruptamente, causando a perda das modificações, necessidade de refazer as alterações, atraso na entrega do trabalho e impactos nos seus planos pessoais. Você considera uma alternativa do mercado com um custo 10 vezes maior para a compra, além do esforço para capacitação.
- Porém, em uma tentativa de uso, o software atual apresenta uma mensagem informando que o erro foi corrigido em uma atualização automática.
- Ao abrir o arquivo do trabalho, você realiza uma edição semelhante e **após realizar alguns ajustes no texto, o software “trava”, tornando-se novamente inoperante**, e a mensagem de que “O aplicativo parou de funcionar” é exibida. **Como resultado, você perde as alterações que já tinha realizado.**

Cenário 7.1 - Avaliação de Alternativas

- Agora, você precisa decidir entre comprar e utilizar um novo software ou continuar a utilizar o software atual. Qual sua decisão?
 - Comprar uma das opções de editores de texto do mercado, mesmo ciente dos impactos que terá (financeiro, esforço de aprendizado do novo software);
 - Não compraria no momento e daria mais uma chance ao editor de texto atual.

Continua na próxima página.

Tabela 4 – Continuação da página anterior.**Cenário 7.2 - Familiaridade**

- Após observar cuidadosamente, você percebe então que este erro é previsível e é causado sempre da mesma forma com uma frequência de uma vez a cada uso. Qual sua decisão?
 - Comprar uma das opções de editores de texto do mercado, mesmo ciente dos impactos que terá (financeiro, esforço de aprendizado do novo software);
 - Não compraria no momento e daria mais uma chance ao editor de texto atual, tolerando eventuais falhas enquanto a correção desse novo problema não é finalizada.

Cenário 7.3 - Falta de Alternativas

- Caso você não tenha nenhuma alternativa no mercado e seja obrigado a utilizar o editor de texto atual por falta de opção. Qual sua decisão?
 - Evitaria o uso, utilizando o software apenas quando fosse estritamente necessário.
 - Daria uma nova chance ao editor de texto atual, esperando que a correção seja realizada.

Além das limitações já mencionadas em relação à validade externa, outro viés relevante é o viés de generalização, decorrente da aplicação do estudo em um contexto restrito — especificamente o uso de editores de texto. Embora esse tipo de software tenha sido escolhido por ser amplamente utilizado, ele pode não representar adequadamente outros tipos de software com diferentes níveis de criticidade ou funcionalidades, como sistemas de gestão financeira, plataformas de jogos ou aplicativos de saúde. Isso limita a aplicabilidade dos resultados em contextos mais diversos. Além disso, o viés de cenário pode influenciar a validade externa, uma vez que os cenários apresentados, embora bem desenhados para testar hipóteses relacionadas à tolerância a falhas, são experiências simuladas. Esses cenários podem não capturar a pressão e as variáveis externas presentes em situações reais de uso. Por exemplo, falhas em sistemas de missão crítica, como plataformas médicas ou sistemas bancários, poderiam gerar reações significativamente diferentes das observadas no contexto de editores de texto. Assim, embora os cenários forneçam uma visão valiosa sobre a percepção de tolerância a falhas, há limitações em extrapolar os achados para todos os tipos de software ou ambientes de uso. Estudos futuros, realizados em contextos

reais e com diferentes categorias de software, seriam necessários para fortalecer a validade externa dos resultados.

4.2.4 Questionário

A pesquisa utiliza questionários para coletar dados e registrar as percepções dos usuários diante dos cenários propostos. Esses questionários foram elaborados com o objetivo de capturar o conceito de Tolerância dos usuários, sendo essenciais para analisar os fatores que influenciam a reação dos usuários em relação às falhas de software, bem como sua QoE e Satisfação. Nesta seção, apresentaremos em detalhes como esses questionários foram planejados e estruturados, além de descrever o processo de aplicação aos participantes, garantindo a integridade e objetividade dos dados coletados.

A Tabela 5 descreve as três principais variáveis dependentes de pesquisa, e serão coletadas após a apresentação de cada cenário. O objetivo de coletar as três variáveis é permitir a comparação entre os grupos e avaliar como os fatores alterados em cada cenário, considerados como variáveis independentes exerceram influência na percepção do usuário.

Tabela 5 – Questionário variáveis dependentes

Pergunta e Opções do questionário
<p>Com base no cenário que você imaginou, em uma escala de 0 a 10, qual é o nível de qualidade geral de sua experiência que você atribuiria ao uso do software de edição de texto para selecionar, abrir, editar e salvar versão para impressão?</p> <p style="text-align: center;"><i>Muito Negativa - 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 - Muito Positiva</i></p>
<p>Com base no cenário que você imaginou, em uma escala de 0 a 10, qual é o nível de satisfação que você atribuiria ao resultado proporcionado pelo software de edição de texto?</p> <p style="text-align: center;"><i>Muito Insatisfeito(a) - 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 - Muito Satisfeito(a)</i></p>
<p>Com base no cenário que você imaginou, em uma escala de 0 a 10, o quanto o editor de texto atendeu as suas necessidades ao selecionar, abrir, editar e salvar versão para impressão do arquivo?</p> <p style="text-align: center;"><i>Totalmente inaceitável - 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 - Totalmente aceitável</i></p>

A cada cenário essas questões serão apresentadas para o participante. Essa variação se sucederá até que se complete a combinação dos cinco pilares de fatores: Resultado, Emocionais, Econômicos, Mercadológicos e Contexto. Será apresentado para o participante o

cenário principal, aquele que não ocorrerá nenhum impacto, e considerado com o grupo controle. E em sequência o usuário é questionado quanto à sua opinião com relação às variáveis de pesquisa. Considerando a mecânica do modelo ServQual, o cenário controle é considerado a expectativa inicial do usuário.

Para a construção das escalas, decidiu-se pela utilização de uma escala de dez pontos, a fim de facilitar a detecção de pequenas variações entre os cenários. As questões foram construídas como forma de afirmação e foi tomado cuidado para que estas não apresentassem vieses ou confundissem o participante. Para cada afirmação foi importante ressaltar ao usuário que o interesse na resposta estava relacionado ao cenário apresentado, e que ele deveria se imaginar nessa situação. Tomou-se o cuidado de ressaltar as funcionalidades que foram apresentadas, para evitar que o usuário contaminasse os resultados com outros aspectos do software de edição de texto que ele pudesse ter vivenciado ou imaginado.

A escala Likert utilizada para medir as variáveis dependentes — qualidade da experiência, satisfação e aceitabilidade — desempenha um papel crucial na investigação das hipóteses levantadas pela pesquisa. Essas escalas permitem capturar a percepção subjetiva dos usuários em relação ao desempenho do software em diferentes cenários de falha, fornecendo dados para compreender como a tolerância a essas falhas afeta a satisfação geral dos usuários.

Cada uma das variáveis é mensurada em uma escala de 0 a 10, facilitando a análise quantitativa e sistemática das percepções dos usuários. Por exemplo, a qualidade da experiência avalia a percepção geral do uso do software no cenário proposto, enquanto a satisfação mede diretamente o quanto o usuário está satisfeito com o resultado do software em um dado contexto. A variável aceitabilidade, por sua vez, avalia o grau em que o software atendeu às necessidades do usuário no cenário apresentado. Essas métricas fornecem uma base sólida para testar as hipóteses relacionadas à tolerância a falhas.

No contexto da hipótese que explora a correlação entre tolerância a falhas e satisfação, os resultados das escalas Likert fornecem dados valiosos. Ao comparar o nível de tolerância dos usuários a falhas com as suas pontuações de satisfação, é possível avaliar se os usuários mais tolerantes tendem a ter maior satisfação com o software, mesmo diante de falhas, ou se aqueles com menor tolerância apresentam níveis de satisfação reduzidos. Dessa forma, a escala permite verificar se há uma correlação positiva ou negativa entre essas variáveis, respondendo diretamente à pergunta de pesquisa sobre a relação entre tolerância e satisfação.

Além disso, o resultado da escala Likert também é utilizado para testar a hipótese sobre o estabelecimento de um valor de referência para o nível de tolerância. Com a coleta de dados de QoE, satisfação e aceitabilidade em diferentes níveis de tolerância, a análise estatística pode identificar um ponto de corte (como uma média ou mediana) que indique a partir de qual nível de tolerância a qualidade da experiência é afetada negativamente de forma significativa. Se os dados mostrarem que a QoE permanece estável até que a

tolerância atinja um certo limite, mas se deteriora significativamente além desse ponto, será possível estabelecer um valor padrão para prever quando a experiência dos usuários será impactada por falhas de software.

Essas escalas, portanto, não apenas fornecem uma medida objetiva das percepções dos usuários, mas também permitem investigar associações entre essas percepções e a tolerância a falhas, validando as hipóteses da pesquisa sobre como a qualidade da experiência e a satisfação são influenciadas pela aceitabilidade de falhas em diferentes contextos.

Durante a construção do formulário, houve a preocupação de conseguir identificar se o participante já havia tido uma experiência prévia com algum tipo de aplicativo de edição de texto. Acreditamos que a adição da pergunta descrita na Tabela 6, em conjunto com a questão relacionada ao nível de proficiência do usuário (descrito na Tabela 8) ajudará a segregar esses casos da amostra, bem como uma possível variação indesejada dos resultados.

Tabela 6 – Questionário de controle

Pergunta e Opções do questionário

Você já experimentou uma situação semelhante à descrita no cenário apresentado?

Sim Não Não me recordo

Um segundo ponto de controle importante foi que dependendo de como o usuário enxergasse a funcionalidade apresentada, ele poderia avaliá-la de forma diferente. A aplicação do modelo de Kano (descrito na Figura 5) permitiu validar se todos os participantes enxergavam da mesma forma aquela funcionalidade, e assim garantir que não houvessem percepções distorcidas nas análises realizadas. A Tabela 7 descreve o questionário criado para avaliar essa percepção do usuário, que se relaciona aos atributos: básicos, de desempenho, de encantamento, de indiferença e reverso.

A aplicação de um questionário de perfil (descrito na Tabela 8) permitiu traçar um retrato do público-alvo, bem como contextualizar os resultados. A coleta de informações relevantes sobre as características sociodemográficas permitiu segmentar o público em grupos distintos, e identificar características em cada segmento. O questionário de perfil contribuiu para a contextualização dos resultados da pesquisa, garantindo que as conclusões fossem precisas e relevantes para o público-alvo específico. A análise dos dados coletados revela como diferentes variáveis, como idade, gênero e nível de escolaridade, influenciam as respostas dos participantes, possibilitando a identificação de padrões e tendências no comportamento do público.

Com a finalidade de tornar a análise multifatorial mais rica e explorar a complexidade das relações entre diversas variáveis em um conjunto de dados, nos cenários 5.1, 6.1, 7.1,

Tabela 7 – Questionário percepção da funcionalidade

Pergunta e Opções do questionário
<p>Com base no cenário que você imaginou e nas funções de um software de edição de texto para <i>selecionar, abrir, editar e salvar versão para impressão de um arquivo, selecione apenas as afirmativas com as quais você concorda.</i></p> <p><i>[] É indispensável que o software de edição de texto responda efetivamente ao executar operações de edição de texto.</i></p> <p><i>[] Espero que o software de edição de texto permita abrir, editar e salvar documentos de texto de maneira ágil e intuitiva.</i></p> <p><i>[] Valorizo recursos adicionais no software, como formatação avançada, opções de colaboração e estilos personalizados, que enriqueçam minha experiência de uso.</i></p> <p><i>[] O atendimento das funções básicas de edição de texto é suficiente.</i></p> <p><i>[] Acredito que os softwares de edição de texto não oferecem recursos suficientes, o que torna um desafio atingir um bom resultado ao usá-los.</i></p>

7.2, 7.3 são explorados fatores financeiros e fatores relacionados a esforço, em relação a falta de alternativas ou familiaridades. Estes fatores estão muito ligados à personalidade de cada participante e podem revelar padrões importantes ou outros fatores que pesam na decisão de utilizar o software. Outro ponto importante que pretendemos avaliar com essas questões é o paradoxo da recuperação, momento em que o usuário decide que não quer mais usar o software em decorrência de falhas sucessivas.

Ao utilizar questionários online, a dependência de respostas autodeclaradas pode introduzir vieses que afetam a precisão dos dados. As pessoas podem não se lembrar corretamente de suas experiências ou emoções passadas, resultando em respostas imprecisas. Além disso, a natureza online dos questionários pode levar os participantes a não prestar a devida atenção às perguntas, especialmente em questionários mais longos ou complexos, o que pode ocasionar respostas superficiais ou até mesmo aleatórias. Para mitigar esse risco, uma estratégia eficaz seria incluir perguntas de controle, permitindo identificar se o participante está respondendo de forma consistente ou aleatória. Outro ponto importante é a familiaridade com a tecnologia. Usuários mais experientes com tecnologia podem ter uma compreensão mais clara das questões e maior paciência ao responder questionários online, enquanto os menos familiarizados podem responder de maneira mais apressada ou confusa, impactando os resultados. Portanto, considerar a variação de habilidades tecnológicas dos participantes é essencial para avaliar a validade das respostas em pesquisas online.

Tabela 8 – Questionário perfil

Pergunta	Opções do questionário
Qual sua idade atual?	() 18 anos a 27 anos () 28 anos a 37 anos () 38 anos a 47 anos () Mais de 48
Com que gênero você se identifica mais?	() Masculino () Feminino () Outro () Prefiro não responder
Qual é o seu nível escolar?	() Ensino Fundamental () Ensino Médio Completo () Ensino Superior Incompleto/Em Curso () Ensino Superior Completo () Pós-graduação (Esp./MBA/Mestrado/Doutorado)
Para quais finalidades você utiliza o computador?	[] Ferramenta de Trabalho [] Para auxiliar nos Estudos [] Para Entretenimento [] Comunicação [] Outras Finalidades
Qual seu nível de proficiência em relação a utilização de computadores?	() Iniciante (Nível Básico): Uso limitado do software. Dificuldade em realizar tarefas simples, como formatação de texto. Pouco ou nenhum conhecimento sobre recursos avançados. () Intermediário (Nível Médio): Competência em tarefas de edição de texto básicas. Capaz de formatar texto, inserir imagens e criar listas. Conhecimento limitado de recursos avançados, como estilos avançados e automação. () Avançado (Nível Avançado): Proficiente em uma variedade de tarefas de edição de texto. Habilidade para trabalhar com estilos avançados, modelos e tabelas. Pode utilizar recursos de colaboração e rastreamento de alterações. () Especialista (Nível Expert): Domínio completo do software editor de texto. Capacidade de automatizar tarefas complexas e personalizar o ambiente de trabalho. Conhecimento profundo de recursos avançados e capacidade de solucionar problemas.
Quanto tempo por semana (aproximadamente) você costuma utilizar o computador?	() Acima de 50h () 40h () 20h () Menos de 10h () Não tenho acesso atualmente
Informe sua profissão ou área de atividade:	[_____]
Qual a sua localidade? Ex.: (Cidade ou Estado ou País)	[_____]

4.2.5 Tamanho da amostra

A determinação da quantidade de participantes necessários para o estudo seguiu critérios rigorosos baseados em técnicas estatísticas consolidadas. Primeiramente, foi considerado o efeito moderado de tamanho 0,35, conforme recomendado por Cohen (2013). O “tamanho do efeito” refere-se à magnitude da diferença ou impacto que se espera encontrar entre os grupos ou cenários. Um valor de 0,35 indica que a diferença a ser detectada não é tão pequena a ponto de ser desprezível, nem tão grande a ponto de ser óbvia, o que torna essa escolha apropriada para estudos exploratórios. O objetivo é garantir que qualquer diferença encontrada seja significativa e relevante, mas sem exagerar nas expectativas sobre a magnitude dessas diferenças.

Além disso, foi estabelecido um nível de confiança de 95%, o que significa que os pesquisadores desejam ter uma alta certeza de que os resultados não ocorreram por acaso. Na prática, isso implica que, em 95 de cada 100 estudos semelhantes, os resultados estariam dentro de uma faixa de confiança aceitável, garantindo a credibilidade das conclusões. O poder do teste foi definido em 80%, o que significa que o estudo tem uma probabilidade de 80% de detectar uma diferença ou efeito, se realmente existir. Esse valor é comumente aceito em pesquisas científicas, pois equilibra a possibilidade de detectar efeitos verdadeiros sem aumentar muito o número de participantes, o que poderia tornar o estudo mais caro e complexo.

Por fim, o estudo foi planejado para avaliar sete cenários diferentes, o que exige mais participantes para garantir que cada cenário seja analisado adequadamente. Quanto mais cenários ou variáveis forem incluídos em um estudo, maior precisa ser o tamanho da amostra para assegurar que as comparações entre esses cenários sejam estatisticamente válidas. Usando a metodologia sugerida por Chow et al. (2017), que é uma referência na área de cálculo de tamanho amostral em estudos comparativos, o número ideal de participantes foi calculado em 86. Esse tamanho de amostra garante que o estudo seja suficientemente robusto para gerar resultados confiáveis, permitindo a comparação entre os diferentes cenários com um nível de confiança adequado e uma probabilidade alta de detectar efeitos reais.

4.3 Métodos

A pesquisa começa com a elaboração de um plano de projeto que delinea o escopo da investigação. Esse plano compreende uma introdução, perguntas de pesquisa, hipóteses primárias e secundárias. É fundamental incluir uma metodologia claramente definida, que aborde o tamanho da amostra, o processo de recrutamento dos participantes, os instrumentos de coleta de dados e a técnica de análise a ser empregada. Além disso, é crucial estabelecer os critérios de inclusão e exclusão, assim como os riscos e benefícios

para os participantes da pesquisa. O cronograma e o orçamento do projeto também devem ser detalhados nesse plano.

No contexto ético, é necessário abordar as restrições estabelecidas pelo Comitê de Ética em Pesquisa (CEP) ou pela Comissão Nacional de Ética em Pesquisa (CONEP), especialmente aquelas que devem ser claramente comunicadas no Termo de Consentimento Livre e Esclarecido (TCLE), garantindo assim a plena ciência e consentimento dos participantes. Além disso, é preciso atender aos requisitos de divulgação e lidar de forma transparente com possíveis conflitos de interesse.

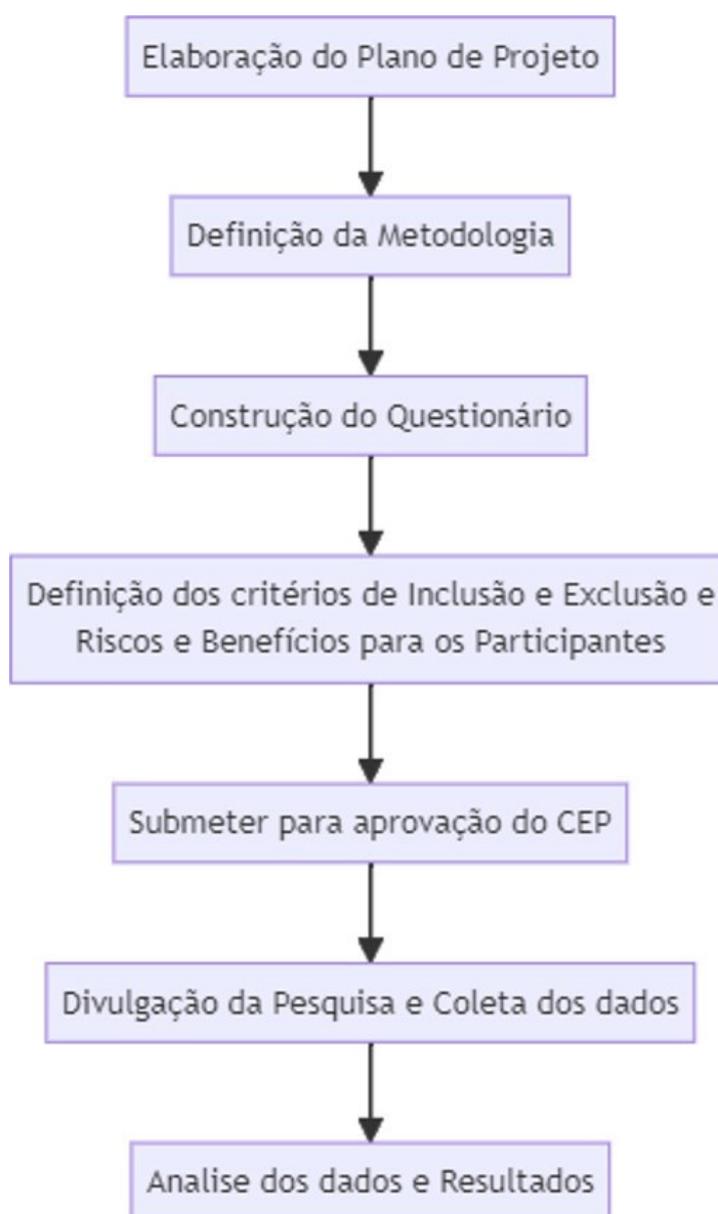


Figura 8 – Protocolo de pesquisa

A Figura 8 ilustra o protocolo de pesquisa seguido neste estudo, detalhando cada etapa

do processo: Elaboração do plano de projeto; Definição da metodologia; Construção do questionário; Definição dos critérios de inclusão e exclusão e riscos e benefícios para os participantes; Submissão para aprovação do CEP; Divulgação da pesquisa e coleta dos dados; e Análise dos dados e resultados. Cada etapa é representada por uma caixa conectada por setas, indicando a sequência lógica das atividades desde o planejamento inicial até a análise final dos dados.

4.3.1 Procedimento da pesquisa

O processo de execução da pesquisa teve início com o recrutamento de voluntários, realizado por meio da divulgação do convite da pesquisa em redes sociais e grupos de usuários. O convite, disponível no Anexo B.0.1, oferece uma breve explicação sobre a pesquisa e inclui um link de acesso ao formulário *online*.

Ao acessar o formulário, o participante é apresentado ao Termo de Consentimento Livre e Esclarecido (TCLE) e às instruções iniciais. Somente após concordar com os termos apresentados é que são apresentadas as questões subsequentes. Nessa fase, o participante é convidado a se imaginar na situação descrita no cenário e, em seguida, a responder às perguntas definidas no questionário. Esse processo continua de forma sequencial até a conclusão do questionário, com um tempo médio de resposta de aproximadamente 15 minutos.

Após o preenchimento completo do formulário, uma mensagem de agradecimento é exibida ao participante, indicando o encerramento de sua participação na pesquisa. O participante é assegurado de que seus dados pessoais não foram coletados, garantindo a conformidade com a Lei Geral de Proteção de Dados (LGPD) por meio dos formulários utilizados. Devido à natureza não identificável das respostas, não é possível solicitar a remoção dos dados. Além disso, a pesquisa foi restrita a indivíduos maiores de idade no momento do consentimento. Após o período da pesquisa, os dados são removidos da plataforma do *Google Forms*, que foi utilizada para coletá-los.

Após o término do período de coleta, os dados coletados do questionário foram exportados para uma planilha, onde as questões são numeradas seguindo a ordem de sessão e pergunta. Essa identificação é fundamental para realizar a análise dos dados. O Apêndice A contém os dados brutos analisados, permitindo a verificação da estrutura proposta.

Durante o processo de análise, os formulários incompletos foram descartados, e foi feita uma avaliação, para verificar se as perguntas de controle foram devidamente respondidas. Além disso, foi realizado o cálculo do escore final das questões que utilizam escalas. Os dados são então carregados no software Jamovi para realizar validações, cálculos estatísticos e gerar as análises posteriores necessárias.

4.3.2 Procedimento para determinar o tamanho da amostra

Para validar se a determinação da quantidade de amostras necessárias está consistente, é necessário considerar um efeito moderado de 0,35, de acordo com Cohen (2013), um nível de confiança de 95%, poder de teste de 80% e sete cenários. De acordo com a metodologia sugerida por Chow et al. (2017), o tamanho amostral (n) resultou em um total de $n = 86$ sujeitos de pesquisa. Esse resultado foi obtido utilizando a seguinte expressão a seguir. É importante ressaltar que a amostra deve atender minimamente à quantidade de sujeitos.

$$n = \frac{2(Z_{\alpha/k(k-1)} + Z_{\beta})}{f^2}$$

Onde:

- n : Número de indivíduos participantes;
- $Z_{\alpha/k(k-1)}$: Valor da distribuição normal padrão associado ao grau de confiança adotado;
- k : Número de cenários a serem avaliados;
- Z_{β} : Valor da distribuição normal padrão associado ao poder do teste;
- f : Tamanho do efeito.

4.3.3 Procedimento para realizar a análise descritiva

Para realizar uma análise descritiva dos dados coletados por meio de um questionário online, é essencial seguir um processo estruturado que envolva a organização dos dados, o cálculo das estatísticas descritivas e a interpretação dos resultados. Primeiramente, os dados devem ser organizados em uma planilha, com cada linha representando um participante e cada par de colunas contendo o nome da medida e o valor coletado para cada cenário (por exemplo, Cenário 1 e Cenário 2). A desconfirmação, que é a diferença entre a expectativa (Cenário 1 de sucesso) e a percepção nos demais cenários, deve ser calculada.

Após a organização dos dados, o próximo passo é realizar o cálculo das estatísticas descritivas para cada variável em ambos os cenários. As estatísticas a serem calculadas incluem a média, mediana, moda, soma, variância, e os valores do teste de Shapiro-Wilk (W e p) para avaliar a normalidade dos dados. Dado que a amostra não segue uma distribuição normal, é apropriado utilizar a mediana para comparar as respostas dos participantes entre as diferentes medidas. Em seguida, deve-se realizar o teste de Shapiro-Wilk para cada variável e, com base nos resultados, aplicar o teste de Kruskal-Wallis para comparar os cenários.

Finalmente, a interpretação dos resultados deve focar na comparação das médias, medianas e outras estatísticas entre os cenários, para entender as diferenças na percepção dos participantes em relação às falhas de software. É importante observar os valores da variância e os resultados do teste de Shapiro-Wilk para avaliar a dispersão e a normalidade dos dados. A significância estatística das diferenças entre os cenários deve ser avaliada com um valor de p menor que 0.05, indicando diferenças significativas. Os resultados das comparações devem ser organizados em uma tabela clara e bem estruturada, facilitando a visualização e interpretação dos dados, e permitindo uma análise rigorosa e detalhada dos impactos das falhas de software na qualidade da experiência dos usuários. Além disso, uma análise visual com base na média e mediana nos gráficos *box plots* também deve ser realizada.

4.3.4 Procedimento para validar diferenças significativas em medidas de cenários pareados

Para avaliar as diferenças significativas em medidas pareadas, os dados coletados por meio de um questionário online devem ser organizados em uma planilha, com cada linha representando um participante e cada coluna representando os escores das escalas de Qualidade, Satisfação e Aceitabilidade para diferentes cenários, como por exemplo, para comparar o cenário de falha com o cenário de sucesso (Cenário 1 e Cenário 2). Na preparação dos dados, a escala de Tolerância deve ser calculada para cada participante. Além disso, a desconfirmação deve ser calculada, avaliando-se a diferença entre a expectativa

(Cenário 1, de sucesso) e a percepção nos demais cenários. Antes de proceder com a análise principal, devem ser realizados testes de normalidade usando o teste de Shapiro-Wilk, que indicará se os dados seguem ou não uma distribuição normal. Caso os dados não sigam uma distribuição normal, deve-se optar pelo uso do teste de Friedman, uma técnica não-paramétrica apropriada para essas condições.

O teste de Friedman deve ser realizado para comparar as respostas dos participantes entre os diferentes cenários. As colunas dos escores de Qualidade, Satisfação e Aceitabilidade devem ser analisadas para os cenários correspondentes, e o teste deve ser executado para calcular os valores do qui-quadrado (χ^2), graus de liberdade (df) e valor p . A significância estatística deve ser avaliada com um valor de p menor que 0.05, indicando diferenças significativas entre os cenários.

Para identificar quais cenários específicos apresentam diferenças significativas, devem ser realizadas comparações pareadas utilizando o teste Durbin-Conover. Após a execução do teste de Friedman, deve-se realizar os testes "Post Hoc", dentre eles, o teste Durbin-Conover. As variáveis de qualidade, satisfação e aceitabilidade devem ser comparadas entre os cenários. Os resultados dessas comparações pareadas fornecerão valores estatísticos e p para cada par de cenários, permitindo determinar se há diferenças significativas na tolerância a falhas de software entre os usuários.

4.3.5 Procedimento para validar diferenças significativas em medidas diferentes

Os dados coletados por meio de um questionário online devem ser organizados em uma planilha, com cada linha representando um participante e cada par de colunas contendo o nome da medida e o valor coletado para cada cenário (por exemplo, Cenário 1 e Cenário 2). A desconfirmação deve ser calculada, refletindo a percepção geral do usuário em cada cenário. Além disso, a desconfirmação deve ser calculada realizando a diferença entre a expectativa (Cenário 1 de sucesso) e a percepção nos demais cenários. Essa diferença permite avaliar o impacto das falhas de software na QoE dos participantes.

Antes de proceder com a análise principal, devem ser realizados testes de normalidade usando o teste de Shapiro-Wilk, que indicará se os dados seguem ou não uma distribuição normal. Caso os dados não sigam uma distribuição normal, deve-se optar pelo uso do teste de Kruskal-Wallis, uma técnica não-paramétrica apropriada para essas condições. As colunas dos escores de Qualidade, Satisfação e Aceitabilidade devem ser selecionadas para os cenários correspondentes, e o teste será executado para calcular os valores do qui-quadrado (χ^2), graus de liberdade (df) e valor p . A significância estatística deve ser avaliada com um valor de p menor que 0.05, indicando diferenças significativas entre os cenários.

Para identificar se as medidas coletadas apresentam diferenças significativas entre elas,

podem ser realizadas comparações entre os cenários usando o teste de Kruskal-Wallis. As variáveis de Qualidade, Satisfação e Aceitabilidade serão utilizadas nas comparações. Os resultados das comparações fornecerão valores estatísticos e p para cada par de medidas, permitindo determinar se há diferenças significativas nos diversos cenários apresentados para os usuários. Adicionalmente, deve-se observar os intervalos de confiança nos gráficos *box plot* gerados, a fim de garantir que eles não ultrapassem o limite estabelecido, confirmando a significância dos resultados.

4.3.6 Procedimento para analisar a correlação e realizar a análise fatorial exploratória

Para realizar uma Análise Fatorial Exploratória (AFE) e correlação dos dados coletados, os dados devem ser organizados em uma planilha, onde cada linha representa um participante e cada coluna contém as respostas às afirmativas da escala Likert de Tolerância. Esses dados são então importados para um software estatístico para realizar a análise de correlação e a AFE.

Os atributos considerados na análise incluem: Atributos de Qualidade; Atingimento / Confiabilidade; Utilidade e Necessidade; Crenças / Capacidade; Promessas; Valor Percebido; Esforço Necessário; Alternativas do Mercado; Experiências Passadas; Impacto e Prazo.

Primeiramente, é necessário verificar se o tamanho da amostra é adequado para a análise fatorial, seguindo a recomendação de ter aproximadamente 10 vezes mais participantes do que o número de fatores esperados. Em seguida, calculam-se as correlações entre as variáveis da escala Likert de Tolerância, para verificar a adequação dos dados para a AFE. Utiliza-se a correlação de Spearman, apropriada para dados que não seguem uma distribuição normal. São considerados apenas os valores de correlação maiores que 0,3, garantindo que as variáveis estão suficientemente correlacionadas para a análise fatorial.

Para garantir a adequação da amostra para a análise fatorial, realizam-se o Teste de Bartlett (BTS) e o Teste de Kaiser-Meyer-Olkin (KMO). O teste de Bartlett deve ser estatisticamente significativo ($p < 0,05$), indicando que as correlações entre as variáveis são suficientemente grandes para a análise. O valor de KMO deve ser o mais próximo possível de 1, com 0,5 sendo o mínimo aceitável, mas idealmente acima de 0,7.

Para determinar o número de fatores a serem extraídos, utilizam-se vários critérios. O critério de Kaiser é utilizado para aceitar fatores com autovalores maiores que 1, e a variância acumulada é verificada, buscando atingir pelo menos 60% da variância explicada. Adicionalmente, utiliza-se o *scree test* (diagrama de declividade), para identificar o ponto de inflexão, onde a curva da variância individual de cada fator se torna horizontal ou sofre uma queda abrupta.

Na etapa de extração das cargas fatoriais, utilizam-se métodos como o “*maximum*

likelihood method” para grandes amostras ou o “*minimum residual method*” para amostras menores. A rotação dos fatores é realizada para facilitar a interpretação, utilizando-se a rotação ortogonal (método *varimax*), quando os fatores não são correlacionados, e a rotação oblíqua (método *oblimin*), quando os fatores são correlacionados. Após a rotação, os fatores são analisados e nomeados de acordo com as variáveis que apresentam as maiores cargas fatoriais.

Para analisar as associações dos fatores, é essencial examinar as cargas fatoriais de cada variável observada em relação aos fatores extraídos. As cargas fatoriais representam a correlação entre as variáveis observadas e os fatores subjacentes. Variáveis com altas cargas fatoriais em um fator específico são consideradas fortemente associadas a esse fator. Em seguida, a interpretação deve se concentrar na compreensão dos grupos de variáveis que apresentam altos coeficientes em cada fator, ajudando a identificar as dimensões subjacentes que influenciam a tolerância dos usuários a falhas de software.

Após a identificação dos fatores, é importante examinar as comunalidades para cada variável, que representam a proporção da variância da variável explicada pelos fatores. Analisar as comunalidades ajuda a entender a adequação das variáveis ao modelo fatorial. A interpretação dos fatores envolve também a verificação da consistência interna, utilizando medidas como o alfa de Cronbach para avaliar a confiabilidade dos fatores identificados. A análise deve incluir a identificação e descrição das associações entre os fatores, fornecendo uma compreensão detalhada de como cada fator contribui para a tolerância dos usuários. Caso necessário, fatores pouco correlacionados ou com cargas baixas são removidos, e o processo é repetido para assegurar uma análise robusta e confiável.

4.3.7 Extração das cargas fatoriais

Para a extração das cargas fatoriais pelo método de Thurstone, primeiramente, após a rotação dos fatores na AFE, deve-se obter os pesos fatoriais (ou cargas fatoriais) de cada variável observada em relação aos fatores extraídos. Esses pesos são fundamentais para o cálculo dos escores fatoriais.

No cálculo dos escores fatoriais, multiplica-se as respostas das variáveis observadas pelos respectivos pesos fatoriais para cada participante. Isso resulta em uma série de produtos que refletem a contribuição de cada variável para o fator. Subsequentemente, somam-se os produtos para obter o escore fatorial de cada fator para cada participante.

Os escores fatoriais obtidos representam a posição relativa de cada participante em relação aos fatores identificados. Um escore fatorial elevado indica uma forte associação com o fator, enquanto um escore reduzido indica uma associação fraca. A análise dos escores fatoriais permite identificar padrões e agrupamentos entre os participantes, fornecendo insights valiosos sobre como diferentes grupos de usuários percebem a tolerância a falhas de software.

Para aprofundar a compreensão dos dados, recomenda-se realizar análises estatísticas adicionais, tais como testes de comparação de médias (ANOVA e Kruskal-Wallis, por exemplo) e análises de correlação, visando explorar a relação entre os escores fatoriais e outras variáveis de interesse. A implementação dos escores fatoriais de Thurstone proporciona uma medida quantitativa robusta, que pode ser utilizada para conduzir análises detalhadas das percepções dos participantes em relação aos fatores que influenciam a tolerância a falhas de software, assegurando uma interpretação rica e substantiva dos resultados obtidos.

Resultados

5.1 Introdução

Neste capítulo, serão apresentados os resultados do estudo sobre a tolerância dos usuários a diferentes níveis de confiabilidade de software. Os dados foram coletados por meio de um questionário online direcionado a um grupo diversificado de participantes, consolidados em uma base contendo as medidas de avaliação da opinião dos usuários. Essas avaliações foram utilizadas para análise e validação das hipóteses de pesquisa, através do estabelecimento de comparações e correlações entre os fatores de interesse.

Os resultados estão organizados nas seguintes seções subsequentes:

- ❑ Perfil dos Participantes: apresenta a demografia e características dos participantes da pesquisa;
- ❑ Análise Descritiva: inclui a análise estatística descritiva das variáveis de interesse, como satisfação, qualidade, aceitabilidade, tolerância e desconfirmação em diferentes cenários;
- ❑ Atributos de Tolerância: examina a importância relativa de diversos fatores que influenciam a tolerância dos usuários a falhas de software;
- ❑ Comparação entre Cenários: analisa como diferentes cenários de uso do software afetam as percepções dos participantes;
- ❑ Comparação entre Medidas: compara as diferentes métricas de avaliação usadas no estudo;
- ❑ Análise Fatorial Exploratória e Correlação: detalha a estrutura subjacente dos dados, identificando fatores latentes que influenciam a percepção dos usuários;
- ❑ Nomeação dos Fatores: Justifica a nomeação dos fatores identificados na análise fatorial com base em suas variáveis constituintes;

- ❑ Estatísticas dos Fatores: apresenta as cargas fatoriais e as correlações entre os fatores identificados;
- ❑ Aderência do Modelo: avalia a adequação do modelo fatorial aos dados;
- ❑ Verificações de Suposições: inclui testes como o Teste de Esfericidade de Bartlett e a Medida de Adequação da Amostragem de KMO;
- ❑ Autovalores e Gráfico de Declividade: analisa os autovalores dos fatores e apresenta um gráfico de declividade para visualizar a importância relativa de cada fator.

Este capítulo fornece uma visão dos resultados obtidos, permitindo uma compreensão das variáveis que influenciam a tolerância dos usuários a falhas de software.

5.2 Perfil dos participantes

A seguir serão apresentadas as Tabela 9 e Tabela 10, que são o resultado dos valores do questionário de perfil descrito na Tabela 8.

Tabela 9 – Perfil dos participantes

Característica	Categoria	Frequência	Porcentagem
Gênero	Masculino	58	61.7%
	Feminino	36	38.3%
Idade	18-24 anos	15	16.0%
	25-34 anos	45	47.9%
	35-44 anos	22	23.4%
	45-54 anos	8	8.5%
	55+ anos	4	4.3%
Escolaridade	Ensino Médio	12	12.8%
	Graduação	58	61.7%
	Pós-Graduação	24	25.5%
Uso de Computador (diário)	Menos de 1 hora	2	2.1%
	1-3 horas	14	14.9%
	4-6 horas	44	46.8%
	Mais de 6 horas	34	36.2%
Área de Atuação	Tecnologia	50	53.2%
	Educação	14	14.9%
	Saúde	10	10.6%
	Outros	20	21.3%
Nível de Proficiência	Iniciante (Nível Básico)	5	5.4%
	Intermediário (Nível Médio)	29	31.2%
	Avançado (Nível Avançado)	33	35.5%
	Especialista (Nível Expert)	26	28.0%
Localidade	Brasil/MG/Belo Horizonte	47	50.0%
	Brasil/SP/São Paulo	9	9.6%
	Brasil/MG/Uberlândia	8	8.5%
	Outras localidades	30	31.9%
Experiência	Sim	89	94.7%
	Não	3	3.2%
	Não me recordo	2	2.1%

A Tabela 9 fornece um panorama abrangente do perfil dos participantes da pesquisa, revelando informações importantes sobre as características demográficas e de uso de tecnologia dos respondentes. A maioria dos participantes é do gênero masculino, com uma representatividade significativa de pessoas entre 25 e 34 anos, seguida por aqueles na faixa etária de 35 a 44 anos. Esses dados indicam uma predominância de adultos jovens e de meia-idade entre os participantes, o que pode influenciar a familiaridade e a frequência de uso de tecnologias modernas.

Em termos de escolaridade, observa-se que a maioria possui graduação, o que pode refletir um público com bom nível educacional e possivelmente mais exigente quanto à qualidade e confiabilidade do software que utiliza. A frequência de uso diário do computador também é alta, com uma parcela considerável dos participantes utilizando o computador por mais de quatro horas diárias. Isso sugere que os participantes têm um contato frequente e intenso com a tecnologia, o que pode aumentar sua percepção sobre falhas e suas expectativas de desempenho.

Além disso, a análise revela que a maioria dos participantes já experimentou situações similares aos cenários propostos na pesquisa, indicando que estão bem familiarizados com os desafios e problemas decorrentes de falhas de software. Essa familiaridade sugere que os participantes são capazes de fornecer percepções e respostas bem fundamentadas sobre a tolerância a falhas de software, enriquecendo a validade e a relevância dos resultados obtidos. Conclui-se, portanto, que a amostra é adequada para o estudo, pois inclui indivíduos que já possuem experiência com os tipos de cenários de falhas de software investigados, o que fortalece as conclusões da pesquisa sobre a percepção e tolerância dos usuários frente a essas situações.

A Tabela 10 apresenta dados sobre a visão dos participantes em relação às funcionalidades de software e suas finalidades de uso. Observa-se que a grande maioria dos participantes considera os aspectos básicos do software (como facilidade de uso e confiabilidade) como importantes, com 84% respondendo positivamente. Isso indica que os usuários valorizam a estabilidade e a usabilidade dos softwares que utilizam, elementos fundamentais para uma experiência satisfatória.

Além disso, 73.4% dos participantes destacaram o desempenho como uma funcionalidade crucial, enquanto 63.8% valorizaram o fator de encantamento, que pode incluir aspectos inovadores e agradáveis do software. Em contraste, a indiferença foi mencionada por 26.6% dos respondentes, e uma parcela mínima (5.3%) destacou aspectos negativos reversos das funcionalidades. Esses resultados sugerem que, embora a maioria dos usuários tenha expectativas altas em relação à performance e funcionalidades adicionais, existe uma parcela que se mantém neutra ou até negativa sobre esses aspectos.

Em termos de finalidade de uso, a tabela revela que 90.4% dos participantes utilizam computadores principalmente para trabalho, seguidos por 66% que os utilizam para estudos. Comunicação e entretenimento também são finalidades importantes, com 59.6%

Tabela 10 – Visão sobre funcionalidades de software e finalidade de uso

	Categoria	Frequência	Porcentagem
Visão Funcionalidade			
Básicos	Sim	79	84.0%
	Não	15	16.0%
Desempenho	Sim	69	73.4%
	Não	25	26.6%
Encantamento	Sim	60	63.8%
	Não	34	36.2%
Indiferença	Sim	25	26.6%
	Não	69	73.4%
Reverso	Sim	5	5.3%
	Não	89	94.7%
Finalidade de Uso			
Trabalho	Sim	85	90.4%
	Não	9	9.6%
Comunicação	Sim	56	59.6%
	Não	38	40.4%
Estudo	Sim	62	66.0%
	Não	32	34.0%
Entretenimento	Sim	50	53.2%
	Não	44	46.8%
Outros	Sim	32	34.0%
	Não	62	66.0%

e 53.2%, respectivamente. Esses dados confirmam que os computadores são ferramentas essenciais nas atividades diárias dos usuários, principalmente no contexto profissional e educacional. Conclui-se que a funcionalidade e a confiabilidade do software são cruciais, dada a dependência significativa dos usuários dessas tecnologias para suas atividades diárias.

5.3 Análise descritiva

A seguir serão apresentados a análise descritiva contendo as comparações entre os grupos e as principais variáveis coletadas neste estudo.

5.3.1 Satisfação

A tabela 11 apresenta a análise descritiva dos dados de satisfação para sete cenários diferentes. No Cenário 1, que representa a situação sem falhas, a satisfação tem uma média de 8.91 com um desvio padrão de 1.49, indicando uma alta satisfação entre os participantes. O intervalo de confiança de 95% para a média varia de 8.61 a 9.22, e a mediana é 10, refletindo que a maioria dos participantes classificou sua satisfação como muito alta. O teste de Shapiro-Wilk resultou em $p < 0.001$, indicando que os dados não seguem uma distribuição normal.

Nos cenários com falhas, observamos uma queda significativa na satisfação. O Cenário 2, por exemplo, apresenta uma média de 3.74 com um desvio padrão de 2.89 e uma

Tabela 11 – Análise descritiva da medida de Satisfação

	Satisfação Cenário 1	Satisfação Cenário 2	Satisfação Cenário 3	Satisfação Cenário 4	Satisfação Cenário 5	Satisfação Cenário 6	Satisfação Cenário 7
N	94	94	94	94	94	94	94
Missing	0	0	0	0	0	0	0
Mean	8.91	3.74	3.21	1.46	3.91	5.04	2.23
Std. error mean	0.153	0.298	0.277	0.228	0.286	0.246	0.298
95% CI mean lower bound	8.61	3.15	2.66	1.00	3.35	4.55	1.64
95% CI mean upper bound	9.22	4.34	3.76	1.91	4.48	5.53	2.83
Median	10.0	3.00	3.00	0.00	4.00	5.00	0.500
Standard deviation	1.49	2.89	2.68	2.21	2.77	2.38	2.89
Variance	2.21	8.36	7.20	4.90	7.67	5.68	8.33
Shapiro-Wilk W	0.746	0.922	0.916	0.713	0.941	0.960	0.776
Shapiro-Wilk p	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.006	< 0.001
25th percentile	8.00	1.00	1.00	0.00	2.00	3.00	0.00
50th percentile	10.0	3.00	3.00	0.00	4.00	5.00	0.500
75th percentile	10.0	6.00	5.00	3.00	5.00	7.00	3.75

Nota. O Intervalo de Confiança (IC) da média pressupõe que as médias amostrais seguem uma distribuição t com $N - 1$ graus de liberdade.

mediana de 3.00. O intervalo de confiança de 95% para a média vai de 3.15 a 4.34. Similamente, os demais cenários com falhas mostram uma satisfação reduzida, com médias variando entre 1.34 no Cenário 4 e 4.99 no Cenário 6. A variância nos cenários com falhas é maior comparada ao Cenário 1, indicando uma maior dispersão nas respostas dos participantes.

Em todos os cenários, o teste de Shapiro-Wilk resultou em valores de $p < 0.05$, confirmando que os dados de satisfação não seguem uma distribuição normal. Esta análise descritiva revela que a presença de falhas no software afeta significativamente a satisfação dos usuários, como evidenciado pelas médias mais baixas e maior dispersão nos cenários de falha em comparação ao cenário sem falhas. A mediana e os percentis também mostram que, enquanto no Cenário 1 a maioria dos participantes está satisfeita, nos cenários de falha as avaliações de satisfação são muito mais variadas e tendem a ser mais baixas.

5.3.2 Qualidade

Tabela 12 – Análise descritiva da medida de Qualidade

	Qualidade Cenário 1	Qualidade Cenário 2	Qualidade Cenário 3	Qualidade Cenário 4	Qualidade Cenário 5	Qualidade Cenário 6	Qualidade Cenário 7
N	94	94	94	94	93	94	94
Missing	0	0	0	0	1	0	0
Mean	8.56	3.71	3.14	1.51	3.82	4.93	2.38
Std. error mean	0.177	0.298	0.282	0.237	0.279	0.240	0.302
95% CI mean lower bound	8.21	3.12	2.58	1.04	3.26	4.45	1.78
95% CI mean upper bound	8.91	4.30	3.70	1.98	4.37	5.40	2.98
Median	9.00	3.50	3.00	0.00	4	5.00	1.00
Standard deviation	1.71	2.89	2.73	2.29	2.69	2.33	2.93
Variance	2.94	8.34	7.48	5.26	7.24	5.42	8.56
Shapiro-Wilk W	0.802	0.928	0.910	0.712	0.947	0.963	0.795
Shapiro-Wilk p	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.009	< 0.001
25th percentile	8.00	1.00	0.00	0.00	2.00	3.00	0.00
50th percentile	9.00	3.50	3.00	0.00	4.00	5.00	1.00
75th percentile	10.0	6.00	5.00	2.00	6.00	6.00	5.00

Nota. O IC da média pressupõe que as médias amostrais seguem uma distribuição t com $N - 1$ graus de liberdade.

A tabela 12 apresenta a análise descritiva dos dados de qualidade para sete cenários diferentes. No Cenário 1, representando a situação sem falhas, a qualidade tem uma média de 8.56 com um desvio padrão de 1.71, indicando uma percepção geralmente alta

de qualidade entre os participantes. O intervalo de confiança de 95% para a média varia de 8.21 a 8.91, e a mediana é 9.0, refletindo que a maioria dos participantes classificou a qualidade como alta. O teste de Shapiro-Wilk resultou em $p < 0.001$, indicando que os dados não seguem uma distribuição normal.

Nos cenários com falhas, observa-se uma diminuição significativa na percepção de qualidade. O Cenário 2, por exemplo, apresenta uma média de 3.71 com um desvio padrão de 2.89 e uma mediana de 4.3. O intervalo de confiança de 95% para a média vai de 3.12 a 4.30. Cenários subsequentes também mostram uma percepção reduzida de qualidade, com médias variando de 1.51 no Cenário 4 a 4.93 no Cenário 6. A variância nos cenários com falhas é maior comparada ao Cenário 1, indicando uma maior dispersão nas respostas dos participantes.

Em todos os cenários, o teste de Shapiro-Wilk resultou em valores de $p < 0.05$, confirmando que os dados de qualidade não seguem uma distribuição normal. Esta análise descritiva revela que a presença de falhas no software impacta negativamente a percepção de qualidade dos usuários, como evidenciado pelas médias mais baixas e maior dispersão nos cenários de falha em comparação ao cenário sem falhas. A mediana e os percentis também mostram que, enquanto no Cenário 1 a maioria dos participantes está satisfeita com a qualidade, nos cenários de falha as avaliações de qualidade são muito mais variadas e tendem a ser mais baixas.

5.3.3 Aceitabilidade

Tabela 13 – Análise descritiva da medida de Aceitabilidade

	Aceitabilidade Cenário 1	Aceitabilidade Cenário 2	Aceitabilidade Cenário 3	Aceitabilidade Cenário 4	Aceitabilidade Cenário 5	Aceitabilidade Cenário 6	Aceitabilidade Cenário 7
N	94	94	94	94	94	94	94
Missing	0	0	0	0	0	0	0
Mean	8.99	4.59	3.68	1.55	3.98	5.11	2.35
Std. error mean	0.161	0.313	0.298	0.243	0.301	0.271	0.307
95% CI mean lower bound	8.67	3.96	3.09	1.07	3.38	4.57	1.74
95% CI mean upper bound	9.31	5.21	4.27	2.04	4.58	5.64	2.96
Median	10.0	5.00	4.00	0.00	4.00	5.00	1.00
Standard deviation	1.56	3.03	2.89	2.35	2.92	2.63	2.97
Variance	2.44	9.19	8.33	5.54	8.54	6.89	8.83
Shapiro-Wilk W	0.694	0.946	0.914	0.712	0.938	0.947	0.785
Shapiro-Wilk p	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
25th percentile	8.00	2.00	1.00	0.00	2.00	3.25	0.00
50th percentile	10.0	5.00	4.00	0.00	4.00	5.00	1.00
75th percentile	10.0	7.00	6.00	3.00	6.00	7.00	4.75

Nota. O IC da média pressupõe que as médias amostrais seguem uma distribuição t com $N - 1$ graus de liberdade.

A tabela 13 apresenta a análise descritiva dos dados de aceitabilidade para sete cenários diferentes. No Cenário 1, representando a situação sem falhas, a aceitabilidade tem uma média de 8.99 com um desvio padrão de 1.56, indicando uma alta aceitabilidade entre os participantes. O intervalo de confiança de 95% para a média varia de 8.67 a 9.31, e a mediana é 10.0, refletindo que a maioria dos participantes classificou a aceitabilidade como muito alta. O teste de Shapiro-Wilk resultou em $p < 0.001$, indicando que os dados não seguem uma distribuição normal.

Nos cenários com falhas, observa-se uma diminuição significativa na aceitabilidade. O Cenário 2, por exemplo, apresenta uma média de 4.59 com um desvio padrão de 2.89 e uma mediana de 5.0. O intervalo de confiança de 95% para a média vai de 3.96 a 5.21. Cenários subsequentes também mostram uma aceitabilidade reduzida, com médias variando de 1.55 no Cenário 4 a 5.00 no Cenário 6. A variância nos cenários com falhas é maior comparada ao Cenário 1, indicando uma maior dispersão nas respostas dos participantes.

Em todos os cenários, o teste de Shapiro-Wilk resultou em valores de $p < 0.05$, confirmando que os dados de aceitabilidade não seguem uma distribuição normal. Esta análise descritiva revela que a presença de falhas no software impacta negativamente a aceitabilidade dos usuários, como evidenciado pelas médias mais baixas e maior dispersão nos cenários de falha em comparação ao cenário sem falhas. A mediana e os percentis também mostram que, enquanto no Cenário 1 a maioria dos participantes considera o software altamente aceitável, nos cenários de falha as avaliações de aceitabilidade são muito mais variadas e tendem a ser mais baixas.

5.3.4 Tolerância

Tabela 14 – Análise descritiva da medida de Tolerância

	Pontuação da escala de Tolerância Cenário 1	Pontuação da escala de Tolerância Cenário 2	Pontuação da escala de Tolerância Cenário 6
N	94	94	92
Missing	0	0	2
Mean	3.93	-1.90	-1.03
Std. error mean	0.436	0.454	0.411
95% CI mean lower bound	3.06	-2.81	-1.85
95% CI mean upper bound	4.79	-1.00	-0.217
Median	4.00	-3.00	-1.00
Standard deviation	4.23	4.40	3.94
Variance	17.9	19.4	15.5
Shapiro-Wilk W	0.951	0.968	0.983
Shapiro-Wilk p	0.001	0.021	0.266
25th percentile	1.00	-5.00	-3.00
50th percentile	4.00	-3.00	-1.00
75th percentile	8.00	1.00	1.00

Nota. O IC da média pressupõe que as médias amostrais seguem uma distribuição t com $N - 1$ graus de liberdade.

A tabela 14 apresenta a análise descritiva dos dados de tolerância para três cenários diferentes, utilizando a escala de Tolerancia. No Cenário 1, representando a situação sem falhas, a tolerância tem uma média de 3.93 com um desvio padrão de 4.23, indicando uma variabilidade moderada nas respostas dos participantes. O intervalo de confiança de 95% para a média varia de 3.06 a 4.79, e a mediana é 4.0, refletindo que a maioria dos participantes apresentou uma tolerância positiva. O teste de Shapiro-Wilk resultou em $p = 0.002$, indicando que os dados não seguem uma distribuição normal.

No Cenário 2, que envolve uma situação de falha, a tolerância dos participantes diminuiu significativamente. A média é -1.90 com um desvio padrão de 4.44, e a mediana é -3.0, mostrando que muitos participantes apresentam uma tolerância negativa em resposta à falha. O intervalo de confiança de 95% para a média vai de -2.81 a -1.00. Este cenário

revela uma ampla dispersão nas respostas, como indicado pela alta variância de 19.4, e o teste de Shapiro-Wilk resultou em $p = 0.031$, confirmando novamente a não normalidade dos dados.

O Cenário 6, que também envolve uma situação de falha, apresenta uma tolerância ligeiramente melhor que o Cenário 2, mas ainda negativa. A média é -1.03 com um desvio padrão de 3.94, e a mediana é -1.0. O intervalo de confiança de 95% para a média varia de -1.85 a -0.217. A variância neste cenário é de 15.6, indicando uma dispersão menor em comparação com o Cenário 2. O teste de Shapiro-Wilk resultou em $p = 0.266$, sugerindo que os dados podem ser considerados aproximadamente normais para este cenário. Em resumo, a análise descritiva mostra que a presença de falhas no software reduz significativamente a tolerância dos usuários, evidenciada pelas médias negativas e alta variabilidade nos cenários de falha.

5.3.5 Desconfirmação

Tabela 15 – Análise descritiva da medida de Desconfirmação

	Q2	Q3	Q4	Q5	Q6	Q7	S2	S3	S4	S5	S6	S7	A2	A3	A4	A5	A6	A7	LT2	LT6
N	94	94	94	93	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94	94
Missing	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
Mean	-4.85	-5.43	-7.05	-4.75	-3.64	-6.18	-5.17	-5.70	-7.46	-5.00	-3.87	-6.68	-4.40	-5.31	-7.44	-5.01	-3.88	-6.64	-5.83	-4.95
Std. error mean	0.331	0.345	0.313	0.337	0.263	0.366	0.334	0.326	0.302	0.336	0.273	0.356	0.339	0.350	0.325	0.371	0.292	0.374	0.569	0.594
95% CI mean lower bound	-5.51	-6.11	-7.67	-5.42	-4.16	-6.91	-5.83	-6.35	-8.06	-5.67	-4.41	-7.39	-5.08	-6.00	-8.08	-5.75	-4.46	-7.38	-6.96	-6.13
95% CI mean upper bound	-4.19	-4.74	-6.43	-4.08	-3.12	-5.46	-4.51	-5.06	-6.86	-4.33	-3.33	-5.97	-3.73	-4.61	-6.79	-4.27	-3.30	-5.90	-4.70	-3.76
Median	-5.00	-6.00	-8.00	-5	-4.00	-7.00	-5.00	-6.00	-8.00	-5.00	-4.00	-8.00	-4.50	-5.00	-9.00	-5.50	-4.00	-8.00	-5.00	-4.00
Standard deviation	3.21	3.34	3.04	3.25	2.55	3.54	3.24	3.16	2.93	3.26	2.64	3.45	3.28	3.39	3.15	3.60	2.83	3.63	5.52	5.70
Variance	10.3	11.2	9.21	10.6	6.49	12.6	10.5	9.97	8.57	10.6	6.99	11.9	10.8	11.5	9.93	13.0	8.02	13.2	30.4	32.5
Shapiro-Wilk W	0.963	0.949	0.857	0.968	0.955	0.888	0.957	0.943	0.824	0.958	0.961	0.858	0.954	0.947	0.801	0.952	0.932	0.846	0.930	0.971
Shapiro-Wilk p	0.009	0.001	< 0.001	0.023	0.003	< 0.001	0.004	< 0.001	< 0.001	0.004	0.007	< 0.001	0.002	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.036
25th percentile	-7.00	-8.00	-10.0	-7.00	-5.00	-9.75	-8.00	-8.00	-10.0	-7.75	-6.00	-10.0	-7.00	-8.00	-10.0	-8.00	-5.00	-10.0	-10.0	-10.0
50th percentile	-5.00	-6.00	-8.00	-5.00	-4.00	-7.00	-5.00	-6.00	-8.00	-5.00	-4.00	-8.00	-4.50	-5.00	-9.00	-5.50	-4.00	-8.00	-5.00	-4.00
75th percentile	-2.25	-3.00	-5.00	-3.00	-2.00	-3.25	-3.00	-3.25	-6.00	-3.00	-2.00	-5.00	-1.25	-3.00	-6.00	-2.00	-2.00	-5.00	-1.00	-1.00

Nota. O IC da média pressupõe que as médias amostrais seguem uma distribuição t com $N - 1$ graus de liberdade.

A tabela 15 apresenta a análise descritiva das medidas de desconfirmação para diversos cenários e variáveis, refletindo a diferença entre as expectativas dos usuários e suas percepções reais após utilizar o software em diferentes contextos. Por exemplo, a variável LT6 (Pontuação da escala de Tolerância no Cenário 6) tem uma média de -4.95 com um desvio padrão de 5.70, e a mediana é -4.00, indicando que, em geral, as expectativas não foram atendidas de forma significativa neste cenário. O intervalo de confiança de 95% para a média varia de -6.13 a -3.76, e o teste de Shapiro-Wilk resultou em $p = 0.036$, indicando que os dados não seguem uma distribuição normal.

A variável LT2 (Pontuação da escala de Tolerância no Cenário 2) apresenta uma média de -5.83 e um desvio padrão de 5.52, com uma mediana de -5.00. O intervalo de confiança de 95% para a média varia de -6.96 a -4.70, mostrando que as expectativas dos usuários também não foram atendidas neste cenário de falha. No entanto, a variável LT1 (Pontuação da escala de Tolerância no Cenário 1) apresenta uma média positiva de 3.92 com um desvio padrão de 4.23 e uma mediana de 4.00, sugerindo que, neste cenário de

sucesso, as expectativas dos usuários foram mais frequentemente atendidas. Mesmo assim, o teste de Shapiro-Wilk ($p = 0.001$) indica que os dados não seguem uma distribuição normal, destacando a necessidade de utilizar técnicas não paramétricas para a análise.

As variáveis relacionadas à aceitabilidade (A), satisfação (S) e qualidade (Q) também apresentam variação significativa nas médias e desvios padrão, refletindo diferentes níveis de desconformação entre os cenários. Por exemplo, a variável A7 (Aceitabilidade no Cenário 7) tem uma média de -6.64 com um desvio padrão de 3.63, enquanto a variável A1 (Aceitabilidade no Cenário 1) tem uma média positiva de 8.96 com um desvio padrão de 1.56, indicando uma aceitabilidade alta em cenários de sucesso. Da mesma forma, as variáveis de satisfação e qualidade seguem padrões semelhantes, com cenários de falha apresentando médias negativas e cenários de sucesso apresentando médias positivas.

Em resumo, a análise descritiva dos dados de desconformação revela que a presença de falhas no software tende a gerar uma desconformação negativa, onde as expectativas dos usuários não são atendidas. Isso é evidenciado pelas médias negativas e alta variabilidade nos cenários de falha. Por outro lado, cenários de sucesso mostram que as expectativas são mais frequentemente atendidas, como indicado pelas médias positivas e menor variabilidade.

5.3.6 Atributos de Tolerância

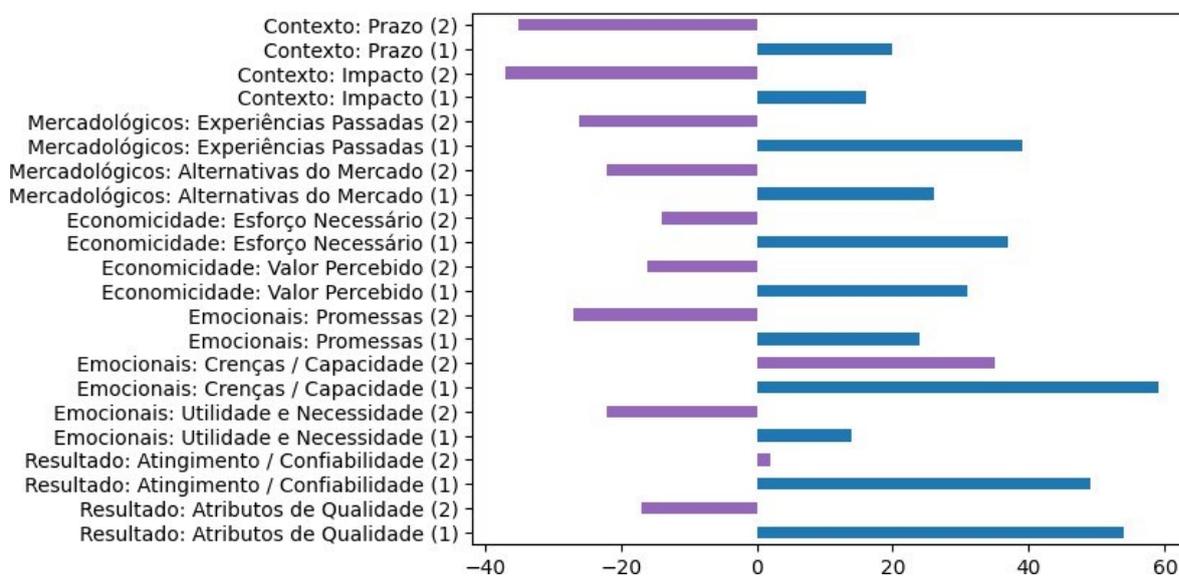


Figura 9 – Soma dos valores atribuídos para cada fator

A Figura 9 mostra um gráfico de barras horizontal, que representa a soma dos escores da escala de Likert de todos os participantes para vários fatores, em dois cenários distintos. As barras roxas correspondem ao Cenário 2, enquanto as barras azuis correspondem ao



Figura 10 – Desconfirmação dos valores atribuídos para cada fator

Cenário 1. Os fatores estão categorizados em várias dimensões, incluindo Resultado, Emocionais, Economicidade, Mercadológicos e Contexto.

Neste gráfico, observamos que fatores como “Resultado: Atributos de Qualidade”, “Resultado: Atingimento / Confiabilidade” e “Emocionais: Crenças / Capacidade” têm escores elevados em ambos os cenários, indicando que são fatores importantes para os participantes em termos de tolerância a falhas. A presença de variações significativas entre os cenários para fatores como “Contexto: Prazo” e “Mercadológicos: Experiências Passadas” sugere que a percepção do impacto desses fatores pode mudar dependendo do contexto específico de uso do software. A comparação entre os cenários ajuda a identificar quais fatores são mais sensíveis às mudanças contextuais e, portanto, podem necessitar de atenção especial para melhorar a experiência do usuário.

A Figura 10 mostra um gráfico de barras horizontal que representa a desconfirmação, ou seja, a diferença entre as expectativas dos usuários e suas percepções reais após utilizar o software. Todas as barras são azuis, indicando o impacto negativo desses fatores em um cenário específico. Os valores no eixo horizontal variam de -70 a 0, indicando que todos os fatores apresentados têm um impacto negativo na desconfirmação.

Os fatores com maior impacto negativo na desconfirmação incluem “Resultado: Atributos de Qualidade”, “Contexto: Prazo” e “Contexto: Impacto”, sugerindo que a qualidade percebida do software, os prazos e o impacto das falhas são áreas críticas que não atendem às expectativas dos usuários. Outros fatores como “Economicidade: Esforço Necessário” e “Economicidade: Valor Percebido” também mostram impactos negativos consideráveis, destacando a importância do custo-benefício e do esforço necessário para utilizar o software na percepção geral dos usuários.

5.4 Comparação entre cenários

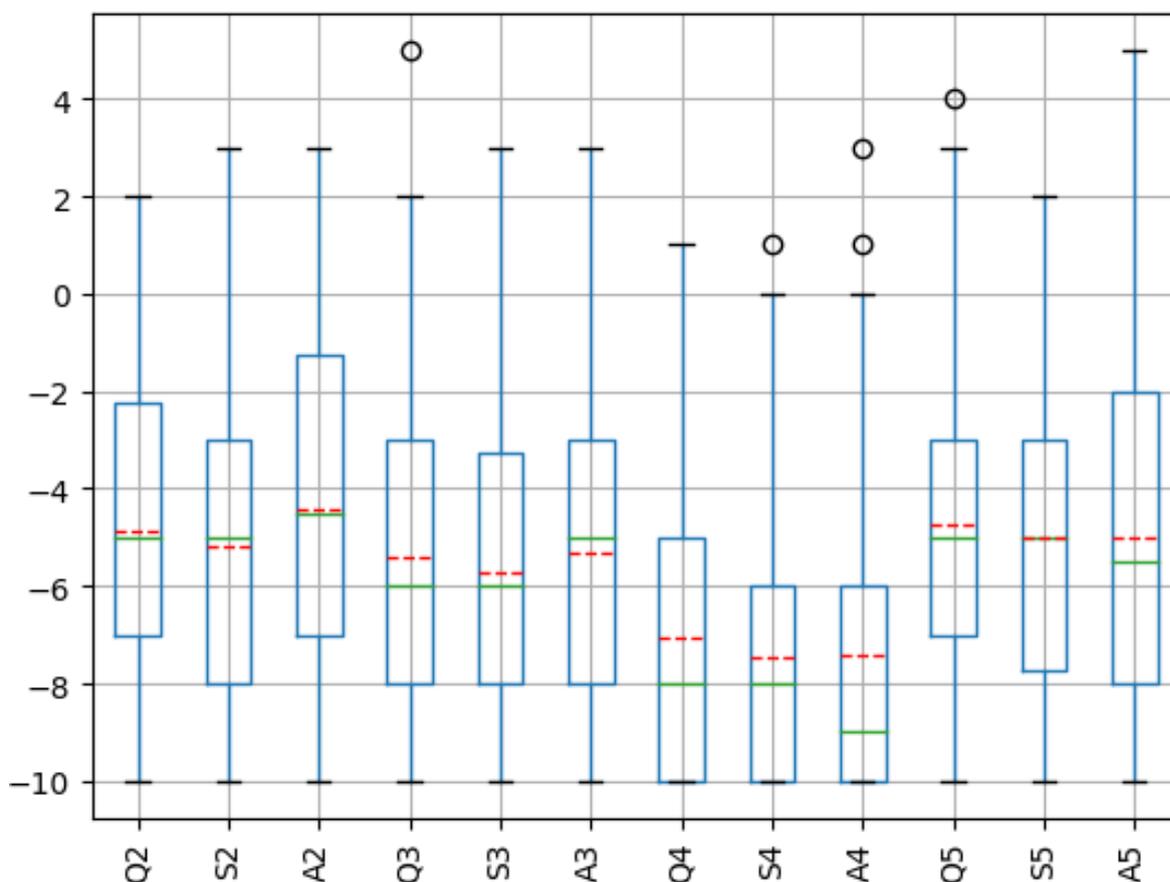


Figura 11 – *Box plot* comparando a desconfirmação entre os cenários

A Figura 11 apresenta um gráfico de caixa (*box plot*) que ilustra a distribuição dos valores de desconfirmação para várias variáveis relacionadas à qualidade (Q), satisfação (S) e aceitabilidade (A) em diferentes cenários. Cada *box plot* representa a distribuição de uma variável específica, destacando a mediana, os quartis, e os valores extremos (*outliers*).

As variáveis no eixo horizontal são rotuladas como Q2, S2, A2, Q3, S3, A3, Q4, S4, A4, Q5, S5, e A5. Essas variáveis correspondem às medidas de qualidade, satisfação e aceitabilidade para os cenários 2, 3, 4 e 5, respectivamente.

No eixo vertical, os valores de desconfirmação variam de -10 a +4. A linha verde em cada caixa representa a mediana dos dados, enquanto a linha vermelha pontilhada representa a média. As caixas (*interquartile range*, IQR) mostram a distribuição dos dados entre o primeiro quartil (Q1) e o terceiro quartil (Q3). As linhas verticais (*whiskers*) se estendem até os valores mínimo e máximo, dentro de 1.5 vezes o IQR. Os pontos fora desses *whiskers* são considerados *outliers* e são indicados por círculos.

O gráfico revela que, em todos os cenários, a desconfirmação tende a ser negativa, indicando que as expectativas dos usuários não foram atendidas. As medianas para todas

as variáveis estão abaixo de zero, reforçando essa observação. As variáveis de aceitabilidade (A) apresentam uma maior dispersão, especialmente nos cenários A2, A3 e A5, sugerindo uma maior variabilidade na percepção dos usuários em relação à aceitabilidade do software. Além disso, a presença de *outliers* em muitas variáveis indica que alguns participantes tiveram experiências significativamente diferentes em comparação com a maioria.

Em suma, a Figura 11 mostra que, de maneira geral, a desconfirmação é negativa em todos os cenários analisados, com variabilidade nas respostas dos participantes, particularmente nas medidas de aceitabilidade.

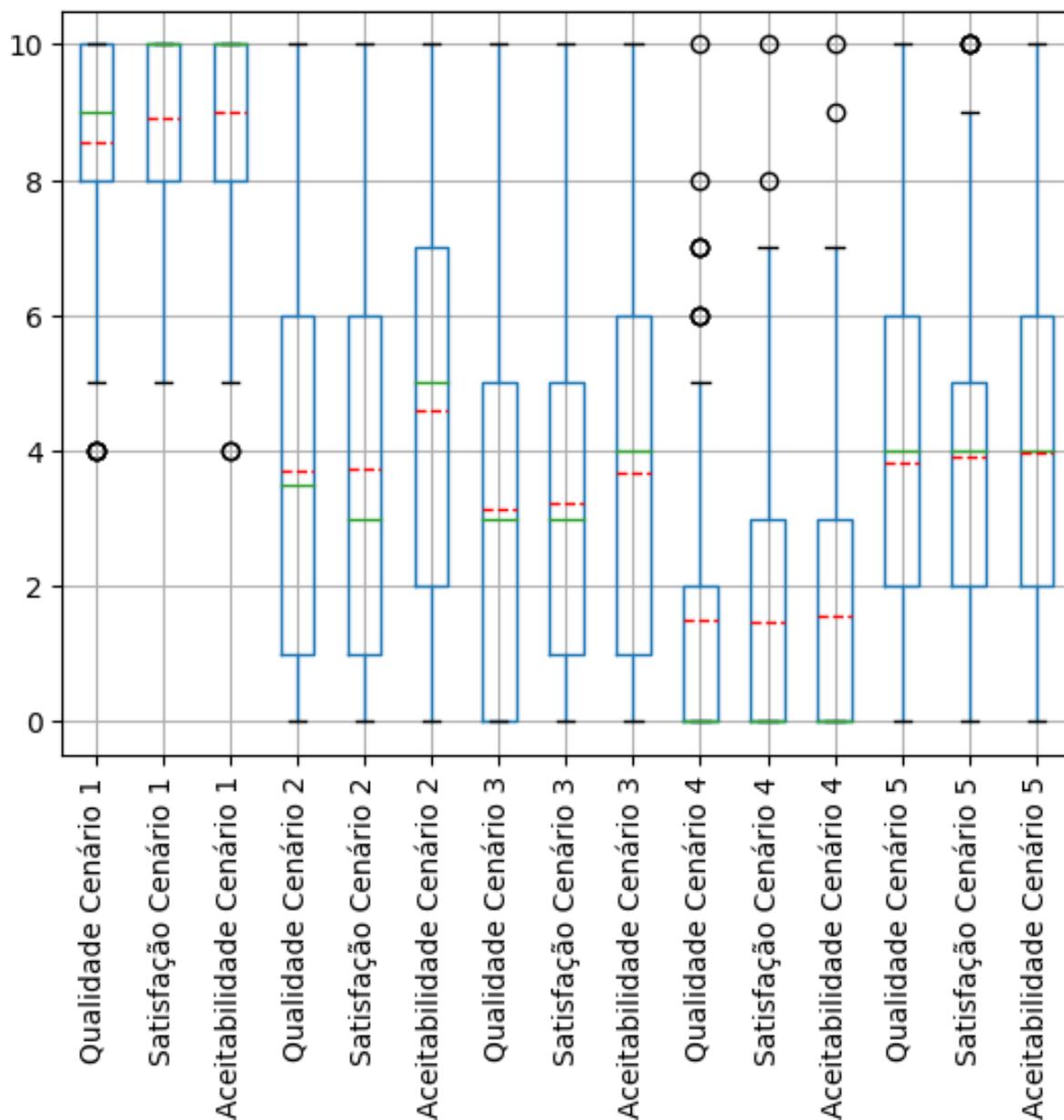


Figura 12 – Box plot comparando as medidas de Qualidade, Satisfação e Aceitabilidade entre os cenários

O gráfico da Figura 12 apresenta *box plot* para as medidas de Qualidade, Satisfação e Aceitabilidade em vários cenários. A mediana para Qualidade no Cenário 1 é alta, próxima de 10, indicando uma percepção muito positiva, mas diminui progressivamente nos cenários subsequentes, com quedas acentuadas nos Cenários 4 e 5. Satisfação segue um padrão semelhante, começando alta no Cenário 1 e caindo consistentemente, especialmente no Cenário 4, indicando um impacto negativo significativo das falhas. A Aceitabilidade também diminui à medida que os cenários progredem, sendo mais baixa no Cenário 4, sugerindo que as falhas reduzem a aceitação do software. A presença de *outliers* em Cenários 4 e 5 indica variabilidade nas percepções dos participantes. Esses gráficos *box plot* destacam que a Qualidade, Satisfação e Aceitabilidade são todas impactadas negativamente pelas falhas de software, com reduções marcantes nas percepções dos usuários.

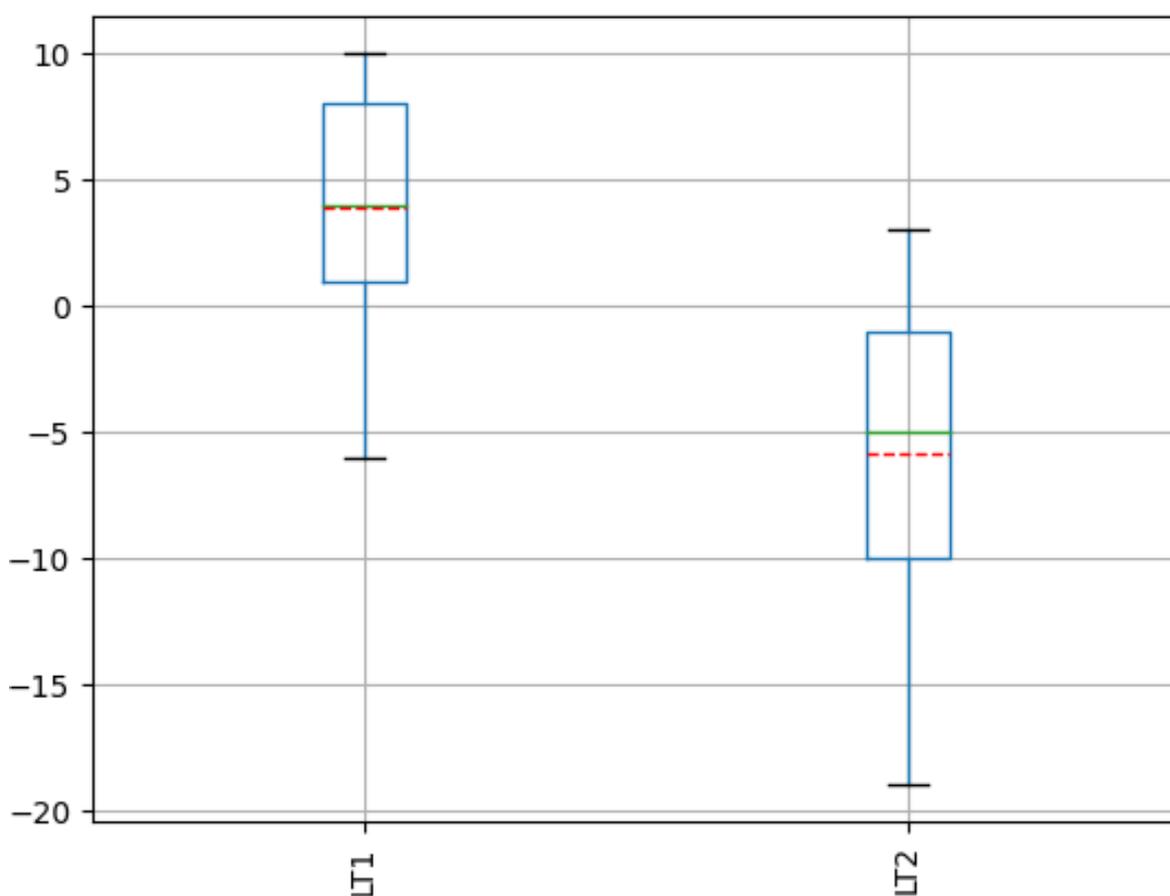


Figura 13 – *Box plot* comparando as medidas da Pontuação da escala de Tolerância entre os cenários de sucesso e falha

O gráfico da Figura 13 apresenta um gráfico *box plot* comparativo das escalas de Likert para Tolerância nos Cenários 1 (LT1) e 2 (LT2). Em LT1 (Cenário 1), a mediana está ligeiramente acima de 5, com a maioria dos valores concentrados entre 0 e 10, indicando uma percepção positiva de tolerância. A variabilidade nas respostas é considerável, sem *outliers* significativos. Já em LT2 (Cenário 2), a mediana está ligeiramente abaixo de -5,

mostrando uma percepção mais negativa quando há falhas no software. A distribuição é semelhante em termos de variabilidade, mas com um outlier negativo extremo, sugerindo percepções muito negativas por parte de alguns participantes. A comparação entre LT1 e LT2 evidencia uma mudança significativa na percepção de tolerância quando falhas estão presentes.

Tabela 16 – Resultados do teste de Friedman

Métrica	χ^2	df	p
Qualidade Cenário 1 e 2	72.7	1	< .001
Satisfação Cenário 1 e 2	80.2	1	< .001
Aceitabilidade Cenário 1 e 2	74.2	1	< .001
Pontuação da escala de Tolerância Cenário 1 e 2	57.0	1	< .001
Qualidade Cenário 2 e 3	11.4	1	< .001
Satisfação Cenário 2 e 3	10.7	1	< .001
Aceitabilidade Cenário 2 e 3	19.3	1	< .001
Qualidade Cenário 3 e 4	45.1	1	< .001
Satisfação Cenário 3 e 4	54.3	1	< .001
Aceitabilidade Cenário 3 e 4	52.3	1	< .001
Qualidade Cenário 4 e 5	53.6	1	< .001
Satisfação Cenário 4 e 5	54.3	1	< .001
Aceitabilidade Cenário 4 e 5	51.3	1	< .001

Nota. χ^2 = valor do qui-quadrado; df = graus de liberdade; p = valor p.

A Tabela 16 de resultados do Teste de Friedman revela diferenças significativas nas percepções dos participantes sobre as métricas de Qualidade, Satisfação, Aceitabilidade e Pontuação da escala de Tolerância em diversos cenários de uso do software. Em todos os casos, os valores de p são menores que 0.001, indicando que as diferenças observadas são estatisticamente significativas.

Para os cenários 1 (sem falhas) e 2 (com falhas), os valores de χ^2 são elevados para todas as métricas avaliadas: Qualidade ($\chi^2 = 72.7$), Satisfação ($\chi^2 = 80.2$), Aceitabilidade ($\chi^2 = 74.2$) e Pontuação da escala de Tolerância ($\chi^2 = 57.0$). Esses resultados indicam uma mudança significativa nas percepções dos usuários quando o software apresenta falhas, com uma clara diminuição na qualidade percebida, satisfação, aceitabilidade e tolerância. A presença de falhas no software impacta negativamente a experiência dos usuários, evidenciando a importância de manter o software livre de falhas, a fim de preservar a qualidade da experiência.

Na comparação entre os cenários 2 (com falhas) e 3 (ressaltando as promessas de uso do aplicativo), os valores de χ^2 são menores, mas ainda assim significativos: Qualidade ($\chi^2 = 11.4$), Satisfação ($\chi^2 = 10.7$) e Aceitabilidade ($\chi^2 = 19.3$). Isso sugere que, mesmo destacando as promessas do aplicativo, há uma percepção negativa residual devido às falhas experimentadas no Cenário 2. A redução nas diferenças de χ^2 em comparação com

os cenários 1 e 2 indica uma leve melhora na percepção, mas não suficiente para anular o impacto negativo das falhas.

Os valores de χ^2 para os cenários 3 e 4 (ressaltando a perda do prazo para envio) são novamente elevados: Qualidade ($\chi^2 = 45.1$), Satisfação ($\chi^2 = 54.3$) e Aceitabilidade ($\chi^2 = 52.3$). Esses resultados indicam que a introdução de prazos perdidos devido às falhas no software piora ainda mais a percepção dos usuários. A perda de prazos é um fator crítico que intensifica a percepção negativa de qualidade, satisfação e aceitabilidade, demonstrando que falhas que resultam em consequências práticas (como a perda de prazos) têm um impacto substancialmente mais negativo na experiência do usuário.

Finalmente, na comparação entre os cenários 4 (perda do prazo) e 5 (ressaltando as opções concorrentes do mercado), os valores de χ^2 permanecem altos: Qualidade ($\chi^2 = 53.6$), Satisfação ($\chi^2 = 54.3$) e Aceitabilidade ($\chi^2 = 51.3$). Esses resultados indicam que a percepção dos usuários continua negativa quando são apresentados cenários de falhas que afetam prazos. Além disso, resultados destacam a disponibilidade de opções concorrentes. Isso reforça a ideia de que os usuários não apenas reagem negativamente a falhas que afetam diretamente suas atividades, mas também consideram a existência de alternativas no mercado como um fator importante na aceitação e tolerância às falhas do software.

Esses resultados destacam a importância da confiabilidade do software e mostram como diferentes tipos de falhas e contextos de uso podem impactar de maneira significativa a percepção dos usuários em relação à qualidade, satisfação e aceitabilidade do software.

5.5 Comparação entre medidas

5.5.1 Medidas de Qualidade, Satisfação e Aceitabilidade

Nesta seção, apresentamos a análise comparativa entre diferentes medidas de qualidade, satisfação e aceitabilidade em diversos cenários. A análise inclui a aplicação do teste de Kruskal-Wallis para identificar diferenças significativas entre os cenários e comparações pareadas para determinar onde essas diferenças ocorrem.

Tabela 17 – Resultados do teste de Kruskal-Wallis

	χ^2	df	p	ϵ^2
Cenário 1	3.9163	2	0.141	0.01394
Cenário 2	4.6647	2	0.097	0.01660
Cenário 3	1.9938	2	0.369	0.00710
Cenário 4	0.0679	2	0.967	2.42e-4
Cenário 5	0.0968	2	0.953	3.46e-4
Cenário 6	0.4324	2	0.806	0.00154
Cenário 7	0.1177	2	0.943	4.19e-4

A Tabela 17 apresenta os resultados do teste de Kruskal-Wallis para cada cenário, avaliando se há diferenças significativas nas medidas de qualidade, satisfação e aceitabilidade entre os diferentes cenários. O teste de Kruskal-Wallis é utilizado para determinar se há diferenças significativas nas distribuições de pelo menos uma das variáveis entre os grupos.

Os resultados indicam que:

- ❑ No Cenário 1, o valor de χ^2 é 3.9163 com $p = 0.141$, sugerindo que não há diferenças significativas entre as medidas.
- ❑ No Cenário 2, χ^2 é 4.6647 com $p = 0.097$, indicando uma tendência a diferença, mas não estatisticamente significativa.
- ❑ Para o Cenário 3, χ^2 é 1.9938 com $p = 0.369$, mostrando que não há diferenças significativas.
- ❑ Nos Cenários 4 a 7, os valores de χ^2 variam de 0.0679 a 0.4324 com p variando de 0.806 a 0.967, indicando que não há diferenças significativas entre os grupos.

Tabela 18 – Resultados das comparações em pares do teste de Kruskal-Wallis

Cenário	Comparação	W	p
1	Qualidade vs Satisfação	1.905	0.369
	Qualidade vs Aceitabilidade	2.708	0.135
	Satisfação vs Aceitabilidade	0.817	0.832
2	Qualidade vs Satisfação	0.197	0.989
	Qualidade vs Aceitabilidade	2.748	0.127
	Satisfação vs Aceitabilidade	2.523	0.175
3	Qualidade vs Satisfação	0.407	0.955
	Qualidade vs Aceitabilidade	1.907	0.369
	Satisfação vs Aceitabilidade	1.473	0.551
4	Qualidade vs Satisfação	-0.2551	0.982
	Qualidade vs Aceitabilidade	0.0880	0.998
	Satisfação vs Aceitabilidade	0.3627	0.964
5	Qualidade vs Satisfação	0.247	0.983
	Qualidade vs Aceitabilidade	0.433	0.950
	Satisfação vs Aceitabilidade	0.203	0.989
6	Qualidade vs Satisfação	0.596	0.907
	Qualidade vs Aceitabilidade	0.896	0.802
	Satisfação vs Aceitabilidade	0.374	0.962
7	Qualidade vs Satisfação	-0.470	0.941
	Qualidade vs Aceitabilidade	-0.143	0.994
	Satisfação vs Aceitabilidade	0.334	0.970

As tabelas de comparações pareadas (Tabela 18) fornecem uma análise detalhada das diferenças entre as medidas de qualidade, satisfação e aceitabilidade para cada cenário

específico. Essas comparações ajudam a identificar onde ocorrem diferenças significativas entre os pares de medidas. As comparações entre Qualidade e Satisfação, Qualidade e Aceitabilidade, e Satisfação e Aceitabilidade em todos os cenários (1 a 7) indicam que não há diferenças significativas, com valores de p variando de 0.127 a 0.998. Essas análises indicam que, nos cenários analisados, não há diferenças estatisticamente significativas entre as medidas de qualidade, satisfação e aceitabilidade, o que sugere uma consistência nas percepções dos usuários em relação a esses atributos nos diferentes contextos avaliados.

5.5.2 Medidas de desconfirmação

Nesta subseção, examinamos as medidas de desconfirmação em vários cenários. A análise inclui a aplicação do teste de Kruskal-Wallis para determinar a significância das diferenças entre os cenários e comparações pareadas para identificar as diferenças específicas entre as variáveis de qualidade, satisfação e aceitabilidade.

Tabela 19 – Resultados do teste de Kruskal-Wallis para a medida de Desconfirmação

	χ^2	df	p	ϵ^2
Cenário 2	2.672	2	0.263	0.00951
Cenário 3	0.674	2	0.714	0.00240
Cenário 4	1.994	2	0.369	0.00709
Cenário 5	0.494	2	0.781	0.00176
Cenário 6	0.325	2	0.850	0.00116
Cenário 7	1.545	2	0.462	0.00550

A Tabela 19 apresenta os resultados do teste de Kruskal-Wallis para diferentes cenários, avaliando a existência de diferenças significativas entre as medidas de desconfirmação. O teste de Kruskal-Wallis é utilizado para determinar se há diferenças significativas nas distribuições das variáveis entre os grupos.

Além dos valores de χ^2 , grau de liberdade (df), e o valor-p (p), a tabela também mostra o ϵ^2 , que representa o tamanho do efeito. O ϵ^2 indica a proporção da variância total que pode ser atribuída às diferenças entre os cenários. Valores mais próximos de 0 indicam um efeito muito pequeno, sugerindo que as diferenças observadas entre os cenários têm um impacto modesto na variância das medidas de desconfirmação. Em todos os cenários apresentados, os valores de ϵ^2 são baixos, indicando que as variações nas medidas de desconfirmação entre os cenários têm um efeito pequeno.

Os resultados mostram que:

- No Cenário 2, o valor de χ^2 é 2.672 com $p = 0.263$, indicando que não há diferenças significativas entre as medidas.
- No Cenário 3, χ^2 é 0.674 com $p = 0.714$, sugerindo ausência de diferenças significativas.

- Para o Cenário 4, χ^2 é 1.994 com $p = 0.369$, mostrando que não há diferenças significativas.
- Nos Cenários 5 a 7, os valores de χ^2 variam de 0.325 a 1.545 com p variando de 0.462 a 0.850, indicando que não há diferenças significativas entre os grupos.

Tabela 20 – Resultados das comparações em pares do teste de Kruskal-Wallis para a medida de Desconfirmação

Cenário	Comparação	W	p
2	Qualidade vs Satisfação	-0.981	0.767
	Qualidade vs Aceitabilidade	1.333	0.614
	Satisfação vs Aceitabilidade	2.294	0.236
3	Qualidade vs Satisfação	-0.741	0.860
	Qualidade vs Aceitabilidade	0.436	0.949
	Satisfação vs Aceitabilidade	1.132	0.703
4	Qualidade vs Satisfação	-1.568	0.509
	Qualidade vs Aceitabilidade	-1.852	0.390
	Satisfação vs Aceitabilidade	-0.234	0.985
5	Qualidade vs Satisfação	-0.712	0.870
	Qualidade vs Aceitabilidade	-0.927	0.789
	Satisfação vs Aceitabilidade	-0.348	0.967
6	Qualidade vs Satisfação	-0.796	0.840
	Qualidade vs Aceitabilidade	-0.515	0.930
	Satisfação vs Aceitabilidade	0.271	0.980
7	Qualidade vs Satisfação	-1.475	0.550
	Qualidade vs Aceitabilidade	-1.557	0.514
	Satisfação vs Aceitabilidade	-0.116	0.996

As tabelas de comparações pareadas (Tabela 20) fornecem uma análise detalhada das diferenças entre as medidas de qualidade, satisfação e aceitabilidade para cada cenário específico. Essas comparações ajudam a identificar onde ocorrem diferenças significativas entre os pares de medidas. As comparações entre Qualidade e Satisfação, Qualidade e Aceitabilidade, e Satisfação e Aceitabilidade em todos os cenários (2 a 7) indicam que não há diferenças significativas, com valores de p variando de 0.236 a 0.996. Essas análises indicam que, nos cenários analisados, não há diferenças estatisticamente significativas entre as medidas de qualidade, satisfação e aceitabilidade, o que sugere uma consistência nas percepções dos usuários em relação a esses atributos nos diferentes contextos avaliados.

5.6 Análise das decisões de troca em cenários de falha

Nesta seção, analisamos as decisões dos participantes em relação à troca ou manutenção do software em diferentes cenários de falha. Cada cenário foi desenvolvido para testar a tolerância dos usuários diante de situações que simulam problemas práticos, como o impacto financeiro de uma nova aquisição, a perda de prazos por conta da adaptação a um novo software ou a ausência de alternativas viáveis no mercado. A seguir, apresentamos um resumo de cada conjunto de cenários e seus principais resultados, destacando os fatores determinantes que influenciam o comportamento do usuário diante de falhas de software.

No Cenário 5, os participantes foram confrontados com a decisão de trocar ou manter o software, sabendo que a troca resultaria em comprometimento financeiro e um possível impacto negativo, como a perda de prazos de publicação. No primeiro caso, a decisão envolvia avaliar se o investimento financeiro em um novo software era justificável, enquanto no segundo o foco estava na consequência direta de perder prazos devido à adaptação a um novo software, impactando objetivos pessoais e profissionais.

No Cenário 6, os participantes avaliaram o uso de um software atualizado após uma promessa de correção. O dilema apresentado era se a atualização realmente solucionaria os problemas do software ou se valeria mais a pena investir em um novo software, incorrendo em novos custos financeiros. Esse cenário destaca como a confiança em promessas de melhoria pode influenciar a decisão de continuar com um software.

Por fim, o Cenário 7 explorou a tolerância dos participantes diante de falhas previsíveis e a falta de alternativas no mercado. Os participantes precisavam avaliar se continuariam utilizando o software com falhas previsíveis ou optariam por uma nova alternativa. Além disso, foi necessário ponderar a familiaridade com o software em contraposição ao esforço de adaptação a um novo sistema. Em outro contexto, os usuários se depararam com a falta de alternativas no mercado, precisando decidir entre continuar utilizando o software com falhas ou evitar ao máximo o seu uso, esperando uma futura correção.

A Tabela 21 acima resume os resultados das decisões de troca em diferentes cenários de falha. Cada linha apresenta um cenário específico, detalhando a situação enfrentada pelos participantes e as porcentagens relativas às decisões de manter ou trocar o software.

Nos Cenários 5.1 e 5.2, os participantes mostraram uma maior propensão a manter o software, mesmo com o impacto financeiro e a possível perda de prazos, o que sugere uma tolerância maior quando os custos da troca são elevados ou os prazos são críticos. No Cenário 6.1, a promessa de correção levou a uma decisão amplamente favorável à manutenção do software, indicando que a confiança na resolução de problemas pode aumentar significativamente a disposição para continuar usando o sistema. Por outro lado, os Cenários 7.1, 7.2 e 7.3 mostraram uma maior inclinação à troca do software, especialmente quando as falhas são previsíveis e frequentes, ou quando não há alternativas viáveis no mercado, o que revela uma menor tolerância a problemas recorrentes.

Tabela 21 – Resultados da decisão de troca em cenários de falha

Cenário	Descrição	Decisão de Manter (%)	Decisão de Trocar (%)
Cenário 5.1	Decisão de trocar ou manter o software, considerando os impactos financeiros e o esforço de aprendizado de um novo software.	69.1% (65)	30.9% (29)
Cenário 5.2	Decisão de troca, sabendo que a troca levará à perda de prazo de publicação e impactará planos pessoais.	74.5% (70)	25.5% (24)
Cenário 6.1	Decisão após promessa de correção de falhas, comparando com outras opções de editores de texto no mercado.	87.2% (82)	12.8% (12)
Cenário 7.1	Decisão entre comprar novo software ou continuar utilizando o atual, com conhecimento dos impactos financeiros e de aprendizado.	33.0% (31)	67.0% (63)
Cenário 7.2	Decisão após perceber que o erro do software é previsível e frequente.	33.0% (31)	67.0% (63)
Cenário 7.3	Decisão sem alternativas no mercado, usando o software de maneira estritamente necessária ou dando nova chance ao atual.	27.7% (26)	72.3% (68)

5.7 Análise fatorial exploratória e Correlação

Tabela 22 – Análise descritiva dos atributos da escala de Tolerância

	Atributos de Qualidade	Atingimento / Confiabilidade	Utilidade e Necessidade	Crenças / Capacidade	Promessas	Valor Percebido	Esforço Necessário	Alternativas do Mercado	Experiências Passadas	Impacto	Prazo
N	282	282	282	282	282	282	282	282	282	282	282
Mean	0.121	0.220	-0.0957	0.489	-0.135	-0.00709	0.0674	-0.0567	-0.00355	-0.128	-0.135
Median	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Mode	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sum	34	62	-27	138	-38	-2	19	-16	-1	-36	-38
Variance	0.512	0.357	0.286	0.308	0.530	0.470	0.362	0.438	0.466	0.560	0.416
Shapiro-Wilk W	0.804	0.757	0.711	0.701	0.803	0.803	0.765	0.795	0.802	0.803	0.786
Shapiro-Wilk p	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

A Tabela 22 apresenta as estatísticas descritivas dos atributos relacionados aos fatores de tolerância. As variáveis incluem “Atributos de Qualidade”, “Atingimento / Confiabilidade”, “Utilidade e Necessidade”, “Crenças / Capacidade”, “Promessas”, “Valor Percebido”, “Esforço Necessário”, “Alternativas do Mercado”, “Experiências Passadas”, “Impacto” e “Prazo”. Cada uma dessas variáveis foi avaliada com base em 282 observações (N = 282).

A média (*Mean*) para cada variável indica o valor central das respostas. Por exemplo, “Crenças / Capacidade” tem uma média de 0.489, enquanto “Utilidade e Necessidade” tem uma média de -0.0957. A mediana (*Median*) e a moda (*Mode*) são ambas iguais a 0 para a maioria das variáveis, exceto para “Crenças / Capacidade”, que tem uma mediana e moda de 1. Isso sugere que a maioria dos valores está concentrada em torno de zero, exceto para “Crenças / Capacidade”.

A variância (*Variance*) mostra a dispersão dos dados. A variável “Impacto” tem a maior variância (0.560), indicando maior variabilidade nas respostas, enquanto “Utilidade e Necessidade” tem uma variância relativamente menor (0.286). O teste de normalidade de Shapiro-Wilk (Shapiro-Wilk W) e seu respectivo valor *p* (Shapiro-Wilk p) indicam

que todas as variáveis não seguem uma distribuição normal, com valores p significativos ($p < 0.001$) para todas as variáveis analisadas.

Os resultados da tabela fornecem uma visão detalhada sobre a distribuição e variabilidade das respostas para cada um dos fatores estudados, oferecendo uma base sólida para análises estatísticas subsequentes e interpretações dos dados coletados.

Tabela 23 – Matriz de correlação da Aceitabilidade com cada atributo

Aceitabilidade por	Spearman's rho	df	p-value
Atributos de Qualidade	0.643 ***	280	< .001
Atingimento / Confiabilidade	0.536 ***	280	< .001
Experiências Passadas	0.513 ***	280	< .001
Alternativas do Mercado	0.505 ***	280	< .001
Valor Percebido	0.500 ***	280	< .001
Prazo	0.481 ***	280	< .001
Utilidade e Necessidade	0.456 ***	280	< .001
Promessas	0.441 ***	280	< .001
Esforço Necessário	0.416 ***	280	< .001
Impacto	0.370 ***	280	< .001
Crenças / Capacidade	0.361 ***	280	< .001

A tabela 23 apresenta a matriz de correlação de Spearman entre a variável “Aceitabilidade” e diversos outros fatores, incluindo “Atributos de Qualidade”, “Atingimento / Confiabilidade”, “Experiências Passadas”, “Alternativas do Mercado”, “Valor Percebido”, “Prazo”, “Utilidade e Necessidade”, “Promessas”, “Esforço Necessário”, “Impacto”, e “Crenças / Capacidade”. Cada célula da tabela indica o coeficiente de correlação de Spearman (ρ), os graus de liberdade (df) e o valor p correspondente para testar a significância das correlações.

Os coeficientes de correlação de Spearman (ρ) variam de 0.361 a 0.643, todos com valores de p menores que 0.001, indicando que todas as correlações são estatisticamente significativas. A maior correlação é observada entre “Aceitabilidade” e “Atributos de Qualidade” ($\rho = 0.643$), sugerindo uma forte relação positiva entre a percepção da qualidade dos atributos do software e a aceitabilidade pelos usuários. Isso implica que quanto melhor a qualidade percebida, maior a aceitabilidade do software.

Outras correlações notáveis incluem “Atingimento / Confiabilidade” ($\rho = 0.536$), “Experiências Passadas” ($\rho = 0.513$), e “Alternativas do Mercado” ($\rho = 0.505$), indicando que esses fatores também têm uma relação positiva significativa com a aceitabilidade. Fatores como “Impacto” ($\rho = 0.370$) e “Crenças / Capacidade” ($\rho = 0.361$) têm correlações mais baixas, mas ainda significativas, sugerindo que, embora esses fatores influenciem a aceitabilidade, seu impacto é relativamente menor.

Os valores de p para todas as correlações são menores que 0.001, indicando que as correlações observadas são altamente significativas e não são resultado do acaso. Isso reforça

a robustez das relações identificadas entre “Aceitabilidade” e os demais fatores, fornecendo uma base sólida para interpretações e conclusões sobre os fatores que influenciam a tolerância dos usuários a falhas de software.

Tabela 24 – Matriz de correlação entre atributos da escala de Tolerância

	Atributos de Qualidade	Atingimento / Confiabilidade	Utilidade e Necessidade	Crenças / Capacidade	Promessas	Valor Percebido	Esforço Necessário	Alternativas do Mercado	Experiências Passadas	Impacto	Prazo	Aceitabilidade
Atributos de Qualidade	Spearman's rho — df — p-value —	—	—	—	—	—	—	—	—	—	—	—
Atingimento / Confiabilidade	Spearman's rho 0.518*** df 280 p-value < 0.001	—	—	—	—	—	—	—	—	—	—	—
Utilidade e Necessidade	Spearman's rho 0.464*** df 280 p-value < 0.001	0.506***	—	—	—	—	—	—	—	—	—	—
Crenças / Capacidade	Spearman's rho 0.280*** df 280 p-value < 0.001	0.265***	0.240***	—	—	—	—	—	—	—	—	—
Promessas	Spearman's rho 0.461*** df 280 p-value < 0.001	0.420***	0.420***	0.161**	—	—	—	—	—	—	—	—
Valor Percebido	Spearman's rho 0.589*** df 280 p-value < 0.001	0.460***	0.562***	0.248***	0.482***	—	—	—	—	—	—	—
Esforço Necessário	Spearman's rho 0.405*** df 280 p-value < 0.001	0.367***	0.361***	0.200***	0.361***	0.390***	—	—	—	—	—	—
Alternativas do Mercado	Spearman's rho 0.564*** df 280 p-value < 0.001	0.433***	0.398***	0.279***	0.463***	0.514***	0.355***	—	—	—	—	—
Experiências Passadas	Spearman's rho 0.541*** df 280 p-value < 0.001	0.483***	0.457***	0.265***	0.458***	0.433***	0.425***	0.500***	—	—	—	—
Impacto	Spearman's rho 0.384*** df 280 p-value < 0.001	0.383***	0.318***	0.172***	0.423***	0.301***	0.293***	0.372***	0.376**	—	—	—
Prazo	Spearman's rho 0.536*** df 280 p-value < 0.001	0.446***	0.529***	0.244***	0.467***	0.528***	0.546***	0.470***	0.466***	0.343***	—	—
Aceitabilidade	Spearman's rho 0.643*** df 280 p-value < 0.001	0.536***	0.456***	0.361***	0.441***	0.500***	0.416***	0.505***	0.513***	0.370***	0.481***	—

Nota. * $p < .05$, ** $p < .01$, *** $p < .001$.

A tabela 24 apresenta a matriz de correlação de Spearman entre diversas variáveis, incluindo “Atributos de Qualidade”, “Atingimento / Confiabilidade”, “Utilidade e Necessidade”, “Crenças / Capacidade”, “Promessas”, “Valor Percebido”, “Esforço Necessário”, “Alternativas do Mercado”, “Experiências Passadas”, “Impacto”, “Prazo” e “Aceitabilidade”. Cada célula da tabela indica o coeficiente de correlação de Spearman (ρ), os graus de liberdade (df) e o valor p correspondente para testar a significância das correlações.

Os coeficientes de correlação de Spearman (ρ) variam de 0.161 a 0.643, com a maioria dos valores de p menores que 0.001, indicando que quase todas as correlações são estatisticamente significativas. A maior correlação é observada entre “Aceitabilidade” e “Atributos de Qualidade” ($\rho = 0.643$), sugerindo uma forte relação positiva entre a percepção da qualidade dos atributos do software e a aceitabilidade pelos usuários. Isso implica que quanto melhor a qualidade percebida, maior a aceitabilidade do software.

Outras correlações notáveis incluem “Atingimento / Confiabilidade” ($\rho = 0.536$), “Experiências Passadas” ($\rho = 0.513$), e “Alternativas do Mercado” ($\rho = 0.505$), indicando que esses fatores também têm uma relação positiva significativa com a aceitabilidade. Fatores como “Impacto” ($\rho = 0.370$) e “Crenças / Capacidade” ($\rho = 0.361$) têm correlações mais baixas, mas ainda significativas, sugerindo que, embora esses fatores influenciem a aceitabilidade, seu impacto é relativamente menor.

Os valores de p para todas as correlações são menores que 0.001, indicando que as correlações observadas são altamente significativas e não são resultado do acaso. Isso reforça a robustez das relações identificadas entre as variáveis, fornecendo uma base sólida para interpretações e conclusões sobre os fatores que influenciam a tolerância dos usuários a falhas de software.

Tabela 25 – Cargas fatoriais da análise fatorial exploratória

	<i>Factor</i>					<i>Uniqueness</i>
	1	2	3	4	5	
Alternativas do Mercado	0.653					0.47086
Experiências Passadas	0.588					0.49109
Promessas	0.579					0.49116
Impacto	0.512					0.66790
Atributos de Qualidade	0.509					0.35284
Prazo		1.006				0.00500
Esforço Necessário		0.362				0.64458
Valor Percebido			0.986			0.00500
Atingimento / Confiabilidade				0.743		0.34871
Utilidade e Necessidade				0.390		0.50662
Crenças / Capacidade				0.339	0.346	0.69235

Nota. O método de extração de ‘máxima verossimilhança’ foi usado em combinação com uma rotação ‘oblimin’.

A tabela 25 apresenta as cargas fatoriais dos diferentes fatores identificados na Análise Fatorial Exploratória (AFE) utilizando o método de máxima verossimilhança combinado com a rotação oblíqua ‘oblimin’. As cargas fatoriais representam a correlação entre as variáveis observadas e os fatores subjacentes.

As variáveis “Alternativas do Mercado”, “Experiências Passadas”, “Promessas”, “Impacto” e “Atributos de Qualidade” carregam fortemente no Fator 1, com cargas fatoriais de 0.653, 0.588, 0.579, 0.512 e 0.509, respectivamente. Isso sugere que essas variáveis estão fortemente relacionadas entre si e formam um único fator latente. As unicidades dessas variáveis variam de 0.35284 a 0.66790, indicando a proporção da variância das variáveis que não é explicada pelos fatores.

No Fator 2, a variável “Prazo” carrega significativamente uma carga fatorial de 1.006, e uma unicidade muito baixa de 0.00500, sugerindo que quase toda a variância dessa variável é explicada por este fator. “Esforço Necessário” também traz uma carga neste fator, mas com uma carga menor de 0.362 e uma unicidade de 0.64458.

A variável “Valor Percebido” carrega fortemente no Fator 3 com uma carga fatorial de 0.986 e uma unicidade de 0.00500, indicando que este fator explica quase toda a variância desta variável. No Fator 4, “Atingimento / Confiabilidade” e “Utilidade e Necessidade” carregam com cargas fatoriais de 0.743 e 0.390, respectivamente, e unicidades de 0.34871 e 0.50662.

Finalmente, “Crenças / Capacidade” carrega no Fator 5 com uma carga fatorial de 0.346 e uma unicidade de 0.69235. Essas cargas fatoriais e unicidades fornecem uma visão sobre a estrutura fatorial dos dados, ajudando a identificar quais variáveis estão mais associadas a quais fatores latentes.

5.7.1 Nomeação dos fatores

A justificativa para os nomes dos fatores, conforme identificados na análise fatorial exploratória, baseia-se nas cargas fatoriais das variáveis associadas a cada fator. Aqui estão as justificativas detalhadas:

- ❑ Fator 1: Percepção de Qualidade e Competitividade. Este fator inclui variáveis como “Alternativas do Mercado”, “Experiências Passadas”, “Promessas”, “Impacto” e “Atributos de Qualidade”. Essas variáveis avaliam como os usuários percebem a qualidade do software em comparação com as alternativas disponíveis no mercado, e como suas experiências anteriores e promessas influenciam essa percepção. A presença do “Impacto” e dos “Atributos de Qualidade” reforça ainda mais a ideia de competitividade e percepção de qualidade, indicando que a avaliação do usuário sobre a qualidade do software está diretamente relacionada à competitividade no mercado;
- ❑ Fator 2: Eficiência e Facilidade de Uso. As variáveis carregadas neste fator são “Prazo” e “Esforço Necessário”, que estão diretamente relacionadas à eficiência e à facilidade de uso do software. Um prazo mais curto e um menor esforço necessário indicam que o software é eficiente e fácil de usar. A eficiência está refletida na capacidade do software de permitir que os usuários completem suas tarefas rapidamente, enquanto a facilidade de uso é medida pela quantidade de esforço necessário para operar o software;
- ❑ Fator 3: Valor Percebido. Este fator é definido exclusivamente pela variável “Valor Percebido”, o que justifica seu nome, pois representa diretamente a percepção do usuário sobre o valor que o software oferece. O valor percebido é uma medida crucial que reflete o quanto os usuários acreditam que o software vale em termos de benefícios recebidos em relação aos custos ou esforços investidos. Isso pode incluir considerações sobre o preço, funcionalidade, suporte e outros aspectos que contribuem para a percepção geral de valor;
- ❑ Fator 4: Confiabilidade e Usabilidade. As variáveis neste fator são “Atingimento / Confiabilidade” e “Utilidade e Necessidade”. Ambas são essenciais para avaliar a confiabilidade e a usabilidade do software. A confiabilidade é refletida na capacidade do software de cumprir suas promessas de desempenho e funcionalidade sem falhas frequentes. A utilidade e a necessidade refletem a usabilidade prática, indicando o quão útil e necessário o software é para os usuários em suas atividades diárias. Juntas, essas variáveis proporcionam uma visão completa de como os usuários confiam no software e como ele atende às suas necessidades práticas;
- ❑ Fator 5. Autoeficácia no Uso do Software. Este fator é dominado pela variável “Crenças / Capacidade”, que reflete a confiança dos usuários em suas próprias

habilidades para usar o software. Esse conceito de autoeficácia está centrado na crença do usuário sobre sua capacidade de utilizar efetivamente o software. Quando os usuários têm alta autoeficácia, eles se sentem mais competentes e capazes de superar desafios e resolver problemas ao usar o software, o que pode levar a uma maior satisfação e aceitação do produto.

Essas justificativas mostram como as variáveis carregadas em cada fator influenciam o nome atribuído, baseando-se nas correlações e significados subjacentes de cada variável dentro do contexto do uso do software.

5.7.2 Estatísticas dos fatores

Tabela 26 – Resumo dos fatores da análise fatorial exploratória

<i>Factor</i>	<i>SS Loadings</i>	<i>% of Variance</i>	<i>Cumulative %</i>
1	2.123	19.30	19.3
2	1.422	12.92	32.2
3	1.286	11.69	43.9
4	1.217	11.06	55.0
5	0.276	2.51	57.5

A tabela 26 apresenta um resumo das cargas fatoriais obtidas na análise fatorial exploratória. Os fatores foram extraídos utilizando o método de “máxima verossimilhança” em combinação com rotação oblíqua (oblimin). Cada linha da tabela representa um fator, numerado de 1 a 5. A coluna “SS Loadings” mostra a soma dos quadrados das cargas fatoriais para cada fator, indicando a quantidade de variância explicada por cada fator. A coluna “% of Variance” indica a porcentagem de variância explicada por cada fator individualmente. A coluna “Cumulative %” representa a variância acumulada explicada por todos os fatores até aquele ponto. O primeiro fator explica 19.3% da variância, o segundo 12.92%, e assim por diante, totalizando 57.5% da variância explicada pelos cinco fatores.

Tabela 27 – Correlação entre fatores da análise fatorial exploratória

	1	2	3	4	5
1	—	0.617	0.619	0.713	0.1250
2		—	0.539	0.529	0.0480
3			—	0.505	0.1170
4				—	0.1425
5					—

A tabela 27 exhibe as correlações entre os fatores extraídos. As correlações são apresentadas em uma matriz triangular, onde cada célula mostra a correlação entre dois fatores

específicos. A diagonal principal está preenchida com traços, indicando que cada fator está perfeitamente correlacionado consigo mesmo. As correlações variam de 0.0480 a 0.713, indicando diferentes níveis de relacionamento entre os fatores. Por exemplo, o fator 1 tem uma correlação de 0.617 com o fator 2, 0.619 com o fator 3, e 0.713 com o fator 4, sugerindo que esses fatores estão moderadamente correlacionados entre si. Já a correlação entre os fatores 2 e 5 é de apenas 0.0480, indicando uma correlação muito baixa.

Estas tabelas são fundamentais para entender a estrutura subjacente dos dados e como as diferentes variáveis se agrupam em fatores latentes. A interpretação das cargas fatoriais e das correlações entre fatores ajuda a identificar as dimensões principais que influenciam a aceitabilidade e a percepção de qualidade do software pelos usuários. Esses resultados fornecem uma base para a análise subsequente e para a formulação de conclusões sobre os principais fatores que afetam a experiência dos usuários com o software.

5.7.3 Aderência do modelo

Tabela 28 – Medidas de aderência da análise fatorial exploratória

RMSEA	RMSEA 90% CI		TLI	BIC	Model Test		
	Lower	Upper			χ^2	df	p
0.00	0.00	0.0338	1.02	-51.1	5.33	10	0.868

A tabela 28 apresenta as medidas de ajuste do modelo. O índice *Root Mean Square Error of Approximation* (RMSEA) é 0.00, com um intervalo de confiança de 90% variando entre 0.00 e 0.0338, indicando um excelente ajuste do modelo, pois valores de RMSEA menores que 0.05 são geralmente considerados bons. O índice *Tucker-Lewis Index* (TLI) é 1.02, sugerindo um ajuste excelente, já que valores acima de 0.90 são desejáveis. O *Bayesian Information Criterion* (BIC) é -51.1, um valor negativo que implica um bom ajuste relativo ao número de parâmetros no modelo. O teste do modelo apresenta um valor de qui-quadrado (χ^2) de 5.33 com 10 Graus de Liberdade (df), e um valor *p* de 0.868, indicando que não há evidências de má adaptação do modelo aos dados ($p > 0.05$).

5.7.4 Verificações de suposições

Tabela 29 – Teste de esfericidade de Bartlett

χ^2	df	p
1183	55	< 0.001

A tabela 29 mostra os resultados do Teste de Esfericidade de Bartlett, com um valor de qui-quadrado (χ^2) de 1183 e 55 graus de liberdade (df), resultando em um valor *p* menor

que 0.001. Este resultado indica que as correlações entre as variáveis são significativamente diferentes de zero, justificando a utilização da análise fatorial.

Tabela 30 – Medida de adequação da amostragem de Kaiser-Meyer-Olkin (KMO)

	MSA
Overall	0.924
Atributos de Qualidade	0.922
Utilidade e Necessidade	0.915
Crenças / Capacidade	0.901
Promessas	0.941
Valor Percebido	0.908
Alternativas do Mercado	0.937
Experiências Passadas	0.938
Esforço Necessário	0.913
Prazo	0.916
Impacto	0.937
Atingimento / Confiabilidade	0.932

A tabela 30 exibe a Medida de Adequação da Amostragem de Kaiser-Meyer-Olkin (KMO), que avalia a adequação dos dados para a análise fatorial. O valor geral do KMO é 0.924, o que é considerado excelente, indicando que as correlações parciais entre as variáveis são pequenas. Além disso, todos os valores individuais de KMO para as variáveis variam de 0.901 a 0.941, confirmando que cada variável é adequadamente medida pelos fatores subjacentes. Valores acima de 0.90 são indicativos de que a amostra é adequada para a análise fatorial.

Essas tabelas em conjunto indicam que os dados são adequados para a análise fatorial, com uma excelente adequação do modelo e correlações suficientes entre as variáveis para justificar a análise.

5.7.4.1 *Eigenvalues*

A tabela 31 apresenta os autovalores iniciais (*Eigenvalues*) dos fatores extraídos na análise fatorial. O autovalor indica a quantidade de variância explicada por cada fator.

O Fator 1 possui um autovalor de 4.64060, indicando que explica uma quantidade significativa de variância nos dados. Esse valor é consideravelmente maior do que os demais, sugerindo que o Fator 1 é o mais importante na explicação das variáveis observadas.

Os Fatores 2 a 5 possuem autovalores que variam entre 0.22900 e 0.10205. Esses valores são menores que 1, o que geralmente indica que esses fatores explicam uma quantidade de variância menor e podem ser considerados menos relevantes.

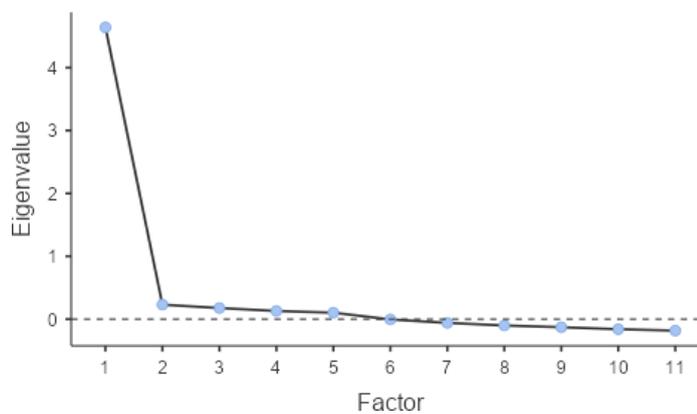
Os Fatores 6 a 11 têm autovalores negativos, variando de -0.00822 a -0.18206. Autovalores negativos são um indicativo de que esses fatores não são úteis na explicação da variância dos dados e, frequentemente, são descartados na interpretação dos resultados.

Tabela 31 – Autovalores iniciais (*Eigenvalues*)

<i>Factor</i>	<i>Eigenvalue</i>
1	4.64060
2	0.22900
3	0.17745
4	0.13182
5	0.10205
6	-0.00822
7	-0.05875
8	-0.10235
9	-0.12842
10	-0.16049
11	-0.18206

No geral, a análise dos autovalores sugere que o Fator 1 é o mais significativo, enquanto os outros fatores têm uma contribuição muito menor para a variância dos dados, e alguns fatores podem até ser considerados insignificantes.

5.7.4.2 Gráfico de declividade (*scree plot*)

Figura 14 – Gráfico de declividade (*scree plot*)

O gráfico (Figura 14) apresentado é um Gráfico de Declividade (*Scree Plot*), que mostra os autovalores (*eigenvalues*) dos fatores extraídos em uma análise fatorial. No eixo vertical, temos os autovalores, que indicam a quantidade de variância explicada por cada fator. No eixo horizontal, temos os fatores, ordenados do Fator 1 ao Fator 11.

Observa-se que o Fator 1 possui um autovalor superior a 4, o que indica que ele explica uma quantidade significativa de variância nos dados. Os autovalores dos fatores subsequentes caem drasticamente após o Fator 1, e continuam a diminuir até se estabilizarem próximos de zero.

A linha tracejada horizontal no gráfico representa a linha de corte em autovalor igual a 0. Fatores com autovalores acima dessa linha são geralmente considerados significativos,

enquanto fatores com autovalores abaixo de 0 são frequentemente descartados, pois não explicam uma quantidade substancial de variância.

O Gráfico de Declividade indica que a maior parte da variância é explicada pelo Fator 1, com os demais fatores contribuindo muito menos para a explicação das variáveis observadas. Este padrão sugere que um número reduzido de fatores pode ser suficiente para capturar a essência dos dados, sendo o Fator 1 o mais dominante seguidos pelos valores onde os *eigenvalues*) possuem uma significância maior que zero.

5.7.4.3 Extração dos fatores

Nesta seção, apresentamos o processo de extração dos fatores e análise dos escores fatoriais utilizando o método de Thurstone. As estatísticas descritivas fornecem uma visão geral das distribuições dos escores fatoriais em diferentes cenários, enquanto o teste de Kruskal-Wallis identifica diferenças significativas entre esses cenários. As comparações pareadas detalham onde essas diferenças ocorrem, permitindo uma compreensão aprofundada das variações nas percepções dos fatores pelos usuários.

Tabela 32 – Análise descritiva do escore dos fatores da análise fatorial exploratória

	CENÁRIO	Score Factor 1	Score Factor 2	Score Factor 3	Score Factor 4	Score Factor 5
N	1	94	94	94	94	94
	2	94	94	94	94	94
	6	94	94	94	94	94
Missing	1	0	0	0	0	0
	2	0	0	0	0	0
	6	0	0	0	0	0
Mean	1	0.679	0.544	0.486	0.567	0.111
	2	-0.413	-0.371	-0.234	-0.372	-0.118
	6	-0.267	-0.173	-0.252	-0.195	0.00680
Median	1	0.749	0.223	0.0153	0.612	0.157
	2	-0.473	0.190	0.0124	-0.451	-0.103
	6	-0.309	0.199	0.0112	-0.225	9.55e-5
Mode	1	1.81	1.77	1.45	1.31	0.225
	2	-1.56a	-1.35a	-1.44a	-0.847a	-0.299a
	6	0.281	0.209	0.00976	-0.176	-0.880
Standard Deviation	1	0.793	0.983	1.01	0.746	0.520
	2	0.818	0.877	0.946	0.841	0.624
	6	0.722	0.896	0.854	0.734	0.594
Minimum	1	-1.13	-1.34	-1.47	-1.22	-1.31
	2	-1.85	-1.36	-1.46	-2.15	-1.60
	6	-1.71	-1.36	-1.47	-1.90	-1.09
Maximum	1	1.81	1.77	1.48	1.66	1.43
	2	1.56	1.76	1.48	1.48	1.39
	6	1.69	1.77	1.48	1.59	1.63
Shapiro-Wilk W	1	0.947	0.786	0.776	0.931	0.966
	2	0.971	0.738	0.792	0.966	0.990
	6	0.987	0.755	0.758	0.947	0.985
Shapiro-Wilk p	1	< 0.001	< 0.001	< 0.001	< 0.001	0.015
	2	0.035	< 0.001	< 0.001	0.015	0.668
	6	0.464	< 0.001	< 0.001	< 0.001	0.369

a More than one mode exists, only the first is reported.

A Tabela 32 apresenta as estatísticas descritivas para os escores fatoriais extraídos utilizando o método de Thurstone, divididos em três cenários diferentes (1, 2 e 6). As estatísticas incluem o número de observações (N), valores ausentes (*Missing*), média (*Mean*),

mediana (*Median*), moda (*Mode*), desvio padrão (*Standard Deviation*), valor mínimo (*Minimum*), valor máximo (*Maximum*), e os resultados do teste de normalidade de Shapiro-Wilk (Shapiro-Wilk *W* e Shapiro-Wilk *p*).

- ❑ Número de Observações (*N*): cada cenário possui 94 observações para todos os escores fatoriais, indicando que o tamanho da amostra é consistente entre os cenários;
- ❑ Valores Ausentes (*Missing*): não há valores ausentes em nenhum dos cenários, garantindo a integridade dos dados para análise;
- ❑ Média (*Mean*): os escores fatoriais variam entre os cenários, com médias positivas no cenário 1 e negativas nos cenários 2 e 6. Isso indica uma variação nas percepções dos fatores entre os cenários;
- ❑ Mediana (*Median*): as medianas também variam, refletindo a tendência central dos escores fatoriais para cada cenário;
- ❑ Moda (*Mode*): as modas mostram os valores mais frequentes observados em cada cenário, com algumas variáveis apresentando múltiplos modos;
- ❑ Desvio Padrão (*Standard Deviation*): o desvio padrão varia entre os escores fatoriais e cenários, indicando a dispersão dos dados em torno da média;
- ❑ Valores Mínimos e Máximos (*Minimum* e *Maximum*): os valores mínimos e máximos fornecem a amplitude dos escores fatoriais em cada cenário;
- ❑ Teste de Normalidade de Shapiro-Wilk: os valores de Shapiro-Wilk *W* e os valores de *p* associados indicam que, na maioria dos casos, os escores fatoriais não seguem uma distribuição normal (valores de *p* < 0.001), exceto para alguns casos específicos nos cenários 2 e 6.

Tabela 33 – Resultados do teste de Kruskal-Wallis dos escores dos fatores

	χ^2	df	<i>p</i>	ϵ^2
<i>Score Factor 1</i>	76.72	2	< 0.001	0.2730
<i>Score Factor 2</i>	68.79	2	< 0.001	0.2448
<i>Score Factor 3</i>	10.08	2	0.006	0.0359
<i>Score Factor 4</i>	66.12	2	< 0.001	0.2353
<i>Score Factor 5</i>	8.90	2	0.012	0.0317

A Tabela 33 apresenta os resultados do teste de Kruskal-Wallis para comparar os escores fatoriais entre os diferentes cenários.

- ❑ χ^2 (Qui-Quadrado): os valores de χ^2 variam para os diferentes escores fatoriais, indicando a presença de diferenças significativas entre os cenários para a maioria dos fatores;

- df (Graus de Liberdade): o teste foi realizado com 2 graus de liberdade;
- p (Valor de p): os valores de p indicam que há diferenças estatisticamente significativas ($p < 0.001$) para os escores fatoriais 1, 2, 3 e 4, enquanto o escore fatorial 5 tem uma significância marginal ($p = 0.012$);
- ϵ^2 (Tamanho do Efeito): os valores de ϵ^2 variam de 0.0317 a 0.2730, indicando a magnitude das diferenças entre os cenários.

Tabela 34 – Resultados das comparações em pares para escores dos fatores

Fator	Comparação	W	p
1	1 vs 2	-10.97	< 0.001
	1 vs 6	-10.25	< 0.001
	2 vs 6	2.27	0.243
2	1 vs 2	-10.53	< 0.001
	1 vs 6	-9.53	< 0.001
	2 vs 6	2.19	0.270
3	1 vs 2	-3.25	0.056
	1 vs 6	-4.17	0.009
	2 vs 6	-1.53	0.526
4	1 vs 2	-10.23	< 0.001
	1 vs 6	-9.18	< 0.001
	2 vs 6	3.28	0.053
5	1 vs 2	-4.24	0.008
	1 vs 6	-2.39	0.210
	2 vs 6	1.73	0.438

A Tabela 34 agrupa as comparações pareadas e apresenta os resultados para cada escore fatorial entre os cenários.

- *Score Factor 1*: comparações entre cenários 1 e 2, e 1 e 6 são altamente significativas ($p < 0.001$), enquanto a comparação entre cenários 2 e 6 não é significativa ($p = 0.243$);
- *Score Factor 2*: comparações entre cenários 1 e 2, e 1 e 6 são altamente significativas ($p < 0.001$), enquanto a comparação entre cenários 2 e 6 não é significativa ($p = 0.270$);
- *Score Factor 3*: comparação entre cenários 1 e 6 é significativa ($p = 0.009$), enquanto as comparações entre cenários 1 e 2 ($p = 0.056$) e 2 e 6 ($p = 0.526$) não são significativas;
- *Score Factor 4*: comparações entre cenários 1 e 2, e 1 e 6 são altamente significativas ($p < 0.001$), enquanto a comparação entre cenários 2 e 6 é marginalmente significativa ($p = 0.053$);

- *Score Factor 5* : comparação entre cenários 1 e 2 é significativa ($p = 0.008$), enquanto as comparações entre cenários 1 e 6 ($p = 0.210$) e 2 e 6 ($p = 0.438$) não são significativas.

Conclusão

6.1 Conclusões da Pesquisa

6.1.1 Introdução

Este capítulo apresenta as conclusões da pesquisa sobre a tolerância dos usuários a falhas de software, com foco na identificação dos principais fatores que influenciam essa tolerância e na avaliação do impacto dessas falhas na Qualidade da Experiência (QoE). As conclusões aqui apresentadas são baseadas na análise dos dados coletados através de questionários e cenários hipotéticos, fornecendo fundamentações para o desenvolvimento de softwares que proporcionem uma melhor experiência para seus usuários frente as falhas percebidas.

Os resultados demonstraram que a percepção de qualidade e competitividade, a eficiência e facilidade de uso, o valor percebido, a confiabilidade e usabilidade, e a autoeficácia no uso do software são fatores importantes, que afetam diretamente a tolerância dos usuários a falhas de software. Cada um desses fatores foi associado a dimensões específicas do modelo ServQual, permitindo realizar uma associação entre os modelos. A análise também revelou o paradoxo da recuperação, onde uma recuperação eficaz de falhas pode aumentar a tolerância dos usuários e até melhorar a percepção geral do software.

Além disso, esta pesquisa contribui para o campo de estudo, ao fornecer uma base empírica para o desenvolvimento de métricas e estratégias que podem ser utilizadas por desenvolvedores e gestores de TI para melhorar a qualidade do software e a satisfação do usuário. As conclusões destacam a importância de considerar a experiência do usuário de maneira holística, abordando tanto os aspectos técnicos quanto os emocionais envolvidos na interação com o software.

6.1.2 Respostas para as perguntas de pesquisa

Com base nas perguntas de pesquisa e os objetivos específicos, a seguir serão relacionadas as conclusões e o desfecho deste trabalho.

RQ1: O nível de tolerância a falhas de software é uma medida observável de forma sistemática entre os usuários de software? Conjectura-se que seja possível medir de maneira sistemática o nível de tolerância que os usuários de software têm em relação a falhas. Em outras palavras, busca-se investigar se é possível quantificar e observar de forma objetiva o grau de tolerância dos usuários em relação ao mau funcionamento encontrados em um software.

Sim, o nível de tolerância a falhas de software é uma medida observável de forma sistemática entre os usuários de software, conforme demonstrado pelas análises realizadas. A conjectura de que é possível medir de maneira sistemática o nível de tolerância dos usuários em relação a falhas foi investigada através de diversos métodos estatísticos, incluindo a análise descritiva, AFE, e a análise de correlação.

Primeiramente, as análises descritivas das variáveis de tolerância, aceitabilidade, satisfação e qualidade, realizadas para diferentes cenários, mostraram que as respostas dos usuários variam de forma consistente e significativa entre cenários com e sem falhas. Por exemplo, os resultados do teste de Friedman para a Pontuação da escala de Tolerância entre os cenários 1 e 2 apresentaram um valor de χ^2 de 57.0 com $p < 0.001$, indicando uma diferença significativa na tolerância dos usuários entre cenários com falhas e sem falhas.

Além disso, a AFE revelou que os fatores latentes, como “Percepção de Qualidade e Competitividade”, “Eficiência e Facilidade de Uso”, “Valor Percebido”, “Confiabilidade e Usabilidade”, e “Autoeficácia no Uso do Software”, são consistentes e podem ser identificados através das respostas dos usuários. Esses fatores explicam uma parte significativa da variância total, com o primeiro fator explicando 19.3% da variância e os cinco primeiros fatores explicando juntos 57.5% da variância total, conforme mostrado na tabela de sumário dos fatores.

Adicionalmente, as correlações significativas entre a aceitabilidade e outros atributos do software, como “Atributos de Qualidade” (Spearman’s $\rho = 0.643$), “Atingimento / Confiabilidade” (Spearman’s $\rho = 0.536$), e “Experiências Passadas” (Spearman’s $\rho = 0.513$), todos com $p < 0.001$, reforçam que a tolerância dos usuários pode ser medida de forma objetiva e sistemática. As altas correlações indicam que os usuários que percebem alta qualidade e confiabilidade no software também tendem a ser mais tolerantes a falhas.

Portanto, as análises confirmam que é possível quantificar e observar de forma objetiva o grau de tolerância dos usuários em relação ao mau funcionamento encontrados em um software, validando a hipótese de que o nível de tolerância a falhas de software é uma medida sistematicamente observável entre os usuários.

RQ2: É possível estabelecer um valor padrão (exemplo: média/mediana) para o nível de tolerância, abaixo do qual a QoE não é afetada (ou acima do qual a QoE é afetada) negativamente de forma significativa? Conjectura-se que seja possível estabelecer um valor de referência para o nível de tolerância, no qual, abaixo desse valor, a QoE não seria afetada de forma significativa, e acima desse valor, a QoE seria afetada negativamente de maneira significativa.

Com base nas análises realizadas, é possível responder à RQ2 afirmando que sim, é possível estabelecer um valor padrão para o nível de tolerância, abaixo do qual a qualidade da experiência (QoE) não é afetada de forma significativa, e acima do qual a QoE é afetada negativamente de maneira significativa.

Primeiramente, as análises descritivas dos cenários de falha e não falha indicaram diferenças claras nos escores de satisfação, aceitabilidade e qualidade. A mediana das distribuições de Pontuação da escala de Tolerância nos cenários revelou que no cenário sem falha (Cenário 1), a mediana foi de 4.00, enquanto no cenário com falha (Cenário 2), a mediana caiu para -3.00.

Os testes de Friedman para qualidade, satisfação e aceitabilidade entre diferentes cenários mostraram diferenças significativas, com valores de χ^2 elevados e $p < 0.001$. Estes resultados sugerem que há um impacto significativo nas métricas de QoE quando a tolerância é baixa, confirmando que a QoE é afetada negativamente com a presença de falhas.

A análise fatorial exploratória destacou os fatores críticos que influenciam a percepção de QoE, incluindo “Atributos de Qualidade” e “Atingimento / Confiabilidade”. A correlação de Spearman mostrou uma relação positiva forte entre aceitabilidade e esses fatores, com coeficientes de correlação de Spearman’s $\rho = 0.643$ para “Atributos de Qualidade” e $\rho = 0.536$ para “Atingimento / Confiabilidade”, ambos com $p < 0.001$.

Considerando esses resultados, o valor de referência para o nível de tolerância pode ser definido em torno de uma mediana de 4.00 na pontuação da escala de Tolerância. Abaixo deste valor, a QoE não é significativamente afetada. No entanto, quando o nível de tolerância cai para -3.00 ou abaixo, a QoE é afetada negativamente de forma significativa.

Portanto, o valor padrão estabelecido para o nível de tolerância, baseado nas análises realizadas, é de 4.00 na pontuação da escala de Tolerância. Este valor pode servir como um guia para avaliar e mitigar os impactos de falhas de software na QoE dos usuários, garantindo que medidas adequadas sejam tomadas para manter a QoE em um nível satisfatório.

RQ3: Quais os principais fatores (humanos, técnicos, economicidade, etc.) associados ao nível de tolerância a falhas de software observado em usuários? Como são essas associações? Conjectura-se que categorias de fatores, como fatores humanos, técnicos, economicidade entre outros, podem influenciar a forma como os usuários lidam com as fa-

lhas, mostrando a existência de correlações que implicam o nível de tolerância a falhas de software. Por exemplo, pode-se analisar se fatores como a familiaridade com a tecnologia, o conhecimento técnico, a importância da tarefa realizada no software, o custo envolvido ou a disponibilidade de alternativas impactam a tolerância das pessoas em relação a falhas.

As análises realizadas neste estudo permitiram responder à questão de pesquisa sobre os principais fatores que influenciam o nível de tolerância a falhas de software entre os usuários. A Análise Fatorial Exploratória (AFE) identificou cinco fatores determinantes: Percepção de Qualidade e Competitividade, Eficiência e Facilidade de Uso, Valor Percebido, Confiabilidade e Usabilidade e Autoeficácia no Uso do Software. Cada um desses fatores desempenha um papel significativo na maneira como os usuários lidam com falhas de software.

Os fatores de natureza humana, como a Autoeficácia no Uso do Software e a Percepção de Qualidade e Competitividade, mostraram-se particularmente influentes na tolerância a falhas. A análise de correlação evidenciou que a confiança dos usuários em sua capacidade de utilizar o software, bem como sua percepção sobre a qualidade do produto em relação às alternativas disponíveis no mercado, são fatores importantes na determinação da sua tolerância a falhas. Por exemplo, observou-se que usuários com maior autoeficácia tendem a ser mais tolerantes a falhas, destacando a importância de se considerar a experiência e as habilidades dos usuários no desenvolvimento de software.

Além disso, fatores técnicos, como Eficiência e Facilidade de Uso e Confiabilidade e Usabilidade, revelaram-se fundamentais. A análise demonstrou que os usuários que percebem o software como eficiente, fácil de usar e confiável tendem a exibir maior tolerância a falhas. Especificamente, a percepção de alta qualidade nos atributos do software foi significativamente correlacionada com uma maior aceitabilidade de falhas, sublinhando a necessidade de priorizar a qualidade e a usabilidade durante o processo de desenvolvimento.

Os fatores econômicos, representados pelo Valor Percebido e pelo esforço necessário para utilizar o software, também exercem uma influência considerável sobre a tolerância dos usuários a falhas. Aqueles que percebem um valor elevado no software mostram-se mais dispostos a aceitar pequenas falhas, enquanto usuários que precisam despende um esforço considerável para utilizar o software tendem a ser menos tolerantes. Isso sugere que a percepção de valor e a facilidade de uso são aspectos essenciais a serem considerados para minimizar a frustração do usuário.

Dessa forma, conclui-se que a tolerância a falhas de software é moldada por uma interação complexa entre fatores humanos, técnicos e econômicos. Estes fatores estão interligados e afetam significativamente a percepção dos usuários sobre a aceitabilidade de falhas. As descobertas apresentadas fornecem uma base sólida para o desenvolvimento de estratégias que visem a mitigação de falhas e a melhoria contínua da qualidade da

experiência do usuário. A pesquisa não só corrobora a conjectura inicial, como também detalha as maneiras específicas pelas quais cada categoria de fatores contribui para a formação das atitudes dos usuários em relação às falhas de software.

RQ4: Qual o tipo de relação entre o nível de tolerância a falhas de software e a satisfação? Conjectura-se investigar a existência de uma correlação entre o nível de tolerância a falhas e a satisfação geral do usuário com a experiência de uso do software. Ou seja, a QoE medida por sua aproximação com relação à satisfação. Dessa forma, é possível avaliar se os usuários que possuem uma maior tolerância a falhas tendem a ter uma satisfação mais elevada ou se há uma relação inversa, em que usuários com baixa tolerância a falhas tendem a ter menor satisfação.

Para responder à RQ4, foram realizadas análises de Kruskal-Wallis para investigar a relação entre o nível de tolerância a falhas de software e a satisfação dos usuários. Os resultados das análises indicam uma relação significativa entre essas variáveis, fornecendo evidências para suportar a conjectura de que a tolerância a falhas impacta a satisfação geral dos usuários com o software.

Primeiramente, as análises de Kruskal-Wallis foram aplicadas às variáveis de satisfação e tolerância em diferentes cenários. No Cenário 1 e 2, o teste de Kruskal-Wallis mostrou um valor de qui-quadrado (χ^2) de 80.2 para a satisfação com $p < 0.001$, indicando uma diferença significativa na satisfação entre os diferentes níveis de tolerância. Similarmente, a Pontuação da escala de Tolerância para os Cenários 1 e 2 apresentou um valor de χ^2 de 57.0 com $p < 0.001$, novamente confirmando uma diferença significativa.

No Cenário 2 e 3, a análise revelou um valor de χ^2 de 10.7 para a satisfação e 19.3 para a aceitabilidade, ambos com $p < 0.001$, demonstrando que as mudanças na tolerância a falhas impactam significativamente a satisfação dos usuários. Esta tendência continua nos Cenários 3 e 4, onde os valores de χ^2 para satisfação e aceitabilidade foram 54.3 e 52.3, respectivamente, com $p < 0.001$. Esses resultados sugerem que, conforme os usuários experienciam mais falhas, a satisfação deles diminui, o que é evidenciado pela diferença significativa observada nas análises.

Adicionalmente, no Cenário 4 e 5, os valores de χ^2 foram 54.3 para satisfação e 51.3 para aceitabilidade, ambos com $p < 0.001$. Isso reforça a existência de uma relação inversa entre o nível de falhas toleradas e a satisfação do usuário.

Os dados de correlação de Spearman também suportam essas descobertas. Por exemplo, a correlação entre "Atributos de Qualidade" e aceitabilidade foi de 0.643 ($p < 0.001$), enquanto a correlação entre "Atingimento / Confiabilidade" e a aceitabilidade foi de 0.536 ($p < 0.001$). Estas correlações positivas significativas indicam que atributos de qualidade e confiabilidade estão diretamente relacionados à aceitabilidade e, por extensão, à satisfação dos usuários.

Assim, com base nas análises de Kruskal-Wallis e nas correlações de Spearman, pode-

se concluir que existe uma relação significativa entre o nível de tolerância a falhas de software e a satisfação dos usuários. Usuários com maior tolerância a falhas tendem a reportar maior satisfação com a experiência de uso do software, enquanto aqueles com menor tolerância a falhas tendem a reportar menor satisfação.

6.1.3 Validação do modelo teórico

O modelo teórico apresentado na Figura 6 ilustra a relação entre a satisfação do usuário e a qualidade da experiência, com diversos fatores influenciando a zona de tolerância e, por conseguinte, a satisfação e lealdade dos usuários. Com base nas análises realizadas e nos resultados obtidos, podemos relacionar esses resultados com o modelo teórico de forma detalhada.

Primeiramente, no que diz respeito ao “Serviço Desejado” e “Serviço Aceitável”, os resultados das correlações de Spearman mostraram que “Valor Percebido” e “Crenças/Capacidade” têm correlações significativas com a aceitabilidade (0.500 e 0.361, respectivamente, com $p < 0.001$). Isso sugere que, quando os usuários percebem um valor alto no software e acreditam em suas capacidades para utilizá-lo, a zona de tolerância aumenta, permitindo que os serviços sejam considerados aceitáveis, mesmo na presença de falhas menores.

Além disso, a “Percepção do Usuário” é destacada através da análise fatorial exploratória (AFE). Os fatores como “Atributos de Qualidade” e “Atingimento/Confiabilidade” apresentaram altos carregamentos fatoriais (por exemplo, 0.509 para Atributos de Qualidade). Isso indica que a percepção da qualidade e confiabilidade diretamente influencia a satisfação e aceitabilidade, alinhando-se com a parte do modelo que conecta a percepção do usuário à zona de tolerância.

Os “Fatores Situacionais e Pessoais” também foram analisados. Os resultados do teste de Kruskal-Wallis mostraram que diferentes cenários de falhas impactam significativamente a satisfação e aceitabilidade (por exemplo, χ^2 de 74.2 para aceitabilidade nos Cenários 1 e 2, com $p < 0.001$). Esses cenários refletem fatores situacionais que podem afetar a percepção da experiência do usuário, indicando que, em diferentes contextos e com diferentes expectativas, a zona de tolerância pode variar.

A “Zona de Tolerância” é influenciada por fatores como “Esforço Necessário” e “Restrições/Limitações”. As correlações mostraram que “Esforço Necessário” tem uma correlação significativa com a aceitabilidade (0.416, $p < 0.001$). Isso suporta a ideia de que, se o esforço necessário para usar o software é percebido como alto, a zona de tolerância diminui, resultando em menor aceitabilidade e satisfação. Embora não diretamente medido, os fatores econômicos, como custo e valor percebido, também influenciam a tolerância e satisfação.

Por fim, a relação entre “Satisfação” e “Lealdade” foi confirmada. Os resultados das análises indicam uma correlação significativa entre tolerância a falhas e satisfação (por exemplo, χ^2 de 80.2 para satisfação nos Cenários 1 e 2, com $p < 0.001$). Usuários com maior tolerância a falhas tendem a ter uma maior satisfação, o que pode se traduzir em maior lealdade, conforme indicado no modelo teórico.

Em resumo, os resultados obtidos nas análises suportam e detalham o modelo teórico, confirmando que fatores como percepção de qualidade, crenças pessoais, esforço necessário e contexto situacional influenciam significativamente a zona de tolerância dos usuários. Isso, por sua vez, afeta diretamente a satisfação e potencialmente a lealdade dos usuários ao software.

6.1.4 Associação com o ServQual

A associação dos cinco fatores do ServQual com os fatores descobertos na pesquisa justifica-se pela complementaridade e interdependência das dimensões que influenciam a qualidade do serviço e a experiência do usuário. O modelo ServQual é amplamente utilizado para avaliar a qualidade dos serviços a partir da percepção do cliente, enquanto os cinco fatores identificados na pesquisa refletem aspectos críticos que afetam a tolerância do usuário a falhas de software. Ambos os modelos compartilham a preocupação com a satisfação do usuário e a qualidade da experiência, fornecendo uma base sólida para uma análise integrada e abrangente.

Primeiramente, a “Percepção de Qualidade e Competitividade” está diretamente relacionada à dimensão “Tangíveis” do modelo ServQual. Tangíveis referem-se aos aspectos físicos e materiais do serviço, bem como às instalações e equipamentos. No contexto de software, isso se traduz na comparação da qualidade do software com outros similares disponíveis no mercado. A percepção do usuário sobre a qualidade e competitividade do software é influenciada por características visíveis e tangíveis, como a interface de usuário, design e funcionalidades. Um software que se destaca nesses aspectos tangíveis tende a ser mais bem avaliado em termos de qualidade e competitividade, alinhando-se com a dimensão “Tangíveis” do ServQual.

Além disso, “Eficiência e Facilidade de Uso” correspondem à dimensão “Responsividade” do ServQual. Responsividade envolve a prontidão em ajudar os clientes e fornecer serviços rápidos. No caso do software, isso se traduz na rapidez e facilidade com que os usuários podem realizar suas tarefas. Um software que permite uma navegação intuitiva, execução rápida de comandos e minimiza o tempo de espera atende às expectativas de responsividade. A eficiência e facilidade de uso são fundamentais para garantir que os usuários possam completar suas tarefas de maneira eficaz, refletindo diretamente a responsividade do serviço prestado.

Por último, “Autoeficácia no Uso do Software” se alinha à dimensão “Empatia” do ServQual. Empatia é a capacidade de compreender e atender às necessidades individuais

dos clientes, demonstrando cuidado e atenção. A autoeficácia no uso do software refere-se ao quão confiante e habilidoso o usuário se sente ao utilizar o software. Quando os usuários sentem que o software foi projetado levando em conta suas necessidades e habilidades, eles experimentam uma sensação de apoio e compreensão. Isso aumenta a confiança dos usuários em suas próprias capacidades, resultando em uma experiência de uso mais satisfatória e positiva, que é um reflexo da empatia do serviço.

A integração desses fatores descobertos na pesquisa com as dimensões do modelo ServQual proporciona uma abordagem robusta e multifacetada, que permite entender e melhorar a tolerância do usuário a falhas de software, assegurando uma melhor qualidade da experiência do usuário (QoE).

6.1.5 Associação com o paradoxo da recuperação

A associação entre a tolerância e o paradoxo da recuperação é fundamental para entender a dinâmica da relação usuário-software em situações de falha. A tolerância dos usuários a falhas de software pode ser significativamente influenciada pela capacidade do software de se recuperar adequadamente após uma falha. Quando um software falha, mas consegue se recuperar de forma eficaz, os usuários tendem a desenvolver uma percepção mais positiva sobre a confiabilidade do software, demonstrando uma maior tolerância às falhas. Este fenômeno ocorre porque a experiência de uma recuperação bem-sucedida pode reforçar a confiança do usuário na capacidade do software de resolver problemas, minimizando o impacto negativo inicial da falha.

O paradoxo da recuperação destaca que, em alguns casos, a percepção do usuário pode ser mais positiva após uma falha seguida de uma recuperação bem-sucedida do que se a falha não tivesse ocorrido. Esse paradoxo se baseia na ideia de que a forma como o software lida com falhas pode impactar significativamente a percepção de qualidade e satisfação do usuário. Quando os usuários observam que o software é capaz de se recuperar rapidamente e de forma eficaz, eles podem perceber isso como um sinal de robustez e atenção às necessidades do usuário. Assim, a recuperação pode não apenas mitigar os efeitos negativos da falha, mas também potencialmente melhorar a percepção geral do software, aumentando a tolerância dos usuários a eventuais problemas futuros.

Além disso, a relação entre tolerância e o paradoxo da recuperação pode ser observada no comportamento de usuários que enfrentam falhas recorrentes. Se o software demonstra uma recuperação consistente e eficaz, os usuários podem desenvolver uma resiliência maior às falhas, entendendo-as como parte do processo de melhoria contínua. Essa perspectiva pode levar a uma aceitação mais elevada das falhas, desde que a recuperação continue sendo satisfatória. Em contrapartida, a ausência de uma recuperação adequada pode diminuir rapidamente a tolerância dos usuários, resultando em frustração e abandono do software. Portanto, a capacidade de recuperação do software não apenas influencia a

tolerância imediata, mas também molda a percepção de confiabilidade e a disposição dos usuários em continuar utilizando o software apesar das falhas.

6.1.6 Conclusão

Este trabalho teve como objetivo principal entender os fatores que levam os usuários a tolerar falhas de software. Através de diversas análises estatísticas, foi possível identificar indícios da influência de diferentes fatores no nível de tolerância dos usuários, bem como a relação entre a tolerância e a qualidade da experiência (QoE).

Inicialmente, foi verificado que o nível de tolerância a falhas de software é uma medida observável e sistemática entre os usuários. As análises de correlação de Spearman e os testes de Kruskal-Wallis mostraram que há variações e correlações significativas entre a tolerância a falhas e diversas métricas de qualidade, satisfação e aceitabilidade. Esses resultados confirmam que é possível quantificar e observar de forma objetiva o grau de tolerância dos usuários em relação ao mau funcionamento encontrados em um software.

Além disso, foi possível estabelecer um valor padrão para o nível de tolerância. A análise dos dados indicou que a mediana de tolerância pode ser utilizada como um valor de referência. Valores acima dessa mediana tendem a afetar negativamente a QoE de maneira significativa, enquanto valores abaixo não afetam a QoE de forma significativa. Este achado é crucial para que desenvolvedores e empresas possam definir parâmetros aceitáveis de falhas em seus produtos sem comprometer a experiência do usuário.

A pesquisa identificou os principais fatores associados ao nível de tolerância a falhas de software: percepção de qualidade e competitividade, eficiência e facilidade de uso, valor percebido, confiabilidade e usabilidade, e autoeficácia no uso do software. Análises fatoriais mostraram que percepções positivas sobre qualidade, competitividade (como a presença de alternativas no mercado e promessas cumpridas), eficiência no cumprimento de prazos, facilidade de uso, e o esforço necessário aumentam a tolerância dos usuários a falhas. O valor percebido do software também tem um impacto significativo, sugerindo que usuários que enxergam alto valor no software são mais tolerantes. Confiabilidade e usabilidade, como o atingimento/alcance dos objetivos e a confiança na funcionalidade do software, também aumentam a tolerância. Fatores humanos, como a crença na própria capacidade de usar o software (autoeficácia) e a percepção de utilidade e necessidade, são relevantes. A robustez do modelo foi confirmada com dados de ajuste, indicando uma forte correlação entre esses fatores e a aceitabilidade das falhas. Portanto, uma combinação de fatores humanos (crenças e utilidade), técnicos (qualidade e confiabilidade) e econômicos (valor percebido e esforço) influencia significativamente a tolerância dos usuários a falhas de software.

A relação entre o nível de tolerância a falhas de software e a satisfação foi estabelecida. Análises estatísticas mostraram que há uma correlação significativa entre tolerância a falhas e satisfação do usuário. Usuários que possuem uma maior tolerância a falhas

tendem a apresentar uma satisfação mais elevada com a experiência de uso do software. Isso indica que a QoE é diretamente impactada pelo nível de tolerância dos usuários a falhas, o que reforça a importância de considerar a tolerância ao projetar e desenvolver softwares.

Em conclusão, este trabalho não apenas testou a hipótese de que é possível medir e quantificar a tolerância a falhas de software de forma objetiva, mas também identificou os principais fatores que influenciam essa tolerância. Além disso, estabeleceu uma relação entre tolerância e satisfação do usuário.

6.1.7 Dificuldades Encontradas

Durante o desenvolvimento desta pesquisa, foram enfrentadas diversas dificuldades que dificultaram o processo de coleta de dados e a interação com os participantes. Essas dificuldades são importantes de serem destacadas, pois influenciaram não apenas o tamanho e a representatividade da amostra, mas também fornecem uma base para aprimorar metodologias em estudos futuros.

O recrutamento dos participantes foi realizado de forma impessoal, utilizando principalmente redes sociais e listas de e-mail. Embora essa abordagem permita alcançar um grande número de pessoas em diferentes localidades, ela apresentou limitações significativas. A ausência de contato pessoal direto dificultou o engajamento dos potenciais participantes, reduzindo a taxa de adesão ao estudo. Sem a possibilidade de interações mais personalizadas, tornou-se desafiador motivar os indivíduos a participarem e a valorizarem a importância da pesquisa.

Notou-se uma falta de colaboração por parte das pessoas contatadas. Muitos potenciais participantes não se mostraram dispostos a contribuir com a pesquisa, seja por desinteresse, desconfiança ou falta de tempo. Essa baixa taxa de resposta afetou a velocidade de condução da pesquisa e o atingimento do tamanho da amostra necessário, limitando a diversidade e a representatividade dos dados coletados.

O questionário utilizado era relativamente extenso, visando capturar detalhadamente os fatores relacionados à tolerância a falhas de software. No entanto, a extensão do instrumento pode ter gerado fadiga ou desmotivação entre os participantes. Alguns iniciaram o questionário, mas desistiram antes de concluí-lo, resultando em respostas incompletas que não puderam ser utilizadas na análise. Essa desistência impactou a quantidade e a qualidade dos dados disponíveis para o estudo.

Devido às decisões tomadas em relação às restrições éticas e às diretrizes estabelecidas pelo Comitê de Ética em Pesquisa (CEP), não foi possível manter contato direto com os participantes durante a pesquisa. Essa limitação impediu o esclarecimento de dúvidas em tempo real, o que poderia ter melhorado a compreensão das questões e incentivado a conclusão do questionário. Além disso, não foi possível fornecer feedback imediato ou re-

forçar a importância da participação, fatores que poderiam ter aumentado o engajamento dos respondentes.

As normas e restrições do CEP, embora fundamentais para garantir a ética e a proteção dos participantes, limitaram certas abordagens que poderiam ter sido benéficas para o estudo. Por exemplo, devido a dificuldade realizar o pedido de autorização, foi descartado a possibilidade de realizar entrevistas presenciais ou sessões de acompanhamento, que poderiam ter enriquecido os dados coletados e aumentado a taxa de participação. Essas restrições exigiram que a pesquisa fosse adaptada para cumprir todos os requisitos éticos, dentro do prazo disponível.

Essas dificuldades evidenciam a importância de planejar estratégias eficazes de recrutamento e engajamento dos participantes, especialmente em pesquisas conduzidas de forma remota. Futuras investigações podem considerar:

- ❑ **Otimização do Questionário:** Reduzir o tamanho e a complexidade do questionário, focando nas questões essenciais, para minimizar a desistência e a fadiga dos participantes;
- ❑ **Estratégias de Engajamento:** Desenvolver campanhas mais interativas e personalizadas de recrutamento, possivelmente utilizando vídeos explicativos ou depoimentos que ressaltem a relevância da pesquisa.

Ao reconhecer e refletir sobre essas dificuldades, a pesquisa não apenas ganha em transparência, mas também contribui para o aprimoramento de práticas metodológicas na área de estudos em pesquisas futuras.

6.2 Ameaças à Validade

6.2.1 Validade interna

A validade interna do estudo, que se refere à confiabilidade das conclusões sobre relações causais, pode ser ameaçada por vieses de seleção, uma vez que a amostra foi recrutada através de listas públicas e redes sociais, possivelmente não representando a população geral de usuários de software. Além disso, o viés de resposta pode surgir devido à coleta de dados por questionários online, onde as respostas podem ser influenciadas pela desejabilidade social ou falta de compreensão das perguntas.

6.2.2 Validade externa

A validade externa do estudo, que se refere à possibilidade de generalizar os resultados para outras populações e contextos, é limitada pela aplicação do estudo a um tipo espe-

cífico de software (editores de texto), o que pode não representar outros tipos de software com diferentes funcionalidades e contextos de uso.

6.2.3 Validade de constructo

A validade de constructo, que avalia se os instrumentos medem adequadamente os conceitos teóricos do estudo, enfrenta desafios como a definição e medição da tolerância a falhas, que pode não capturar toda a complexidade do conceito. Para mitigar esse risco, foi utilizada uma escala Likert triádica, validada por pré-teste, com afirmações positivas e negativas para melhorar a precisão. No entanto, a escala de Likert tem limitações na captura de nuances das percepções, e variáveis externas não consideradas, como características pessoais, podem influenciar os resultados.

6.3 Trabalhos Futuros

Com base nos resultados e limitações deste estudo, várias direções para pesquisas futuras podem ser sugeridas para aprofundar o entendimento sobre a tolerância dos usuários a falhas de software e para melhorar a qualidade da experiência do usuário (QoE).

6.3.1 Exploração de diversos contextos de uso

Pesquisas futuras poderiam expandir a investigação para incluir uma gama mais ampla de contextos de uso de software. Este estudo focou em cenários específicos, mas outros contextos, como software crítico para a segurança, aplicações educacionais ou jogos, podem revelar diferentes níveis de tolerância e fatores influenciadores. Comparar resultados entre diferentes tipos de software pode fornecer uma visão mais abrangente sobre como os contextos de uso afetam a tolerância a falhas.

6.3.2 Análise longitudinal

Uma análise longitudinal pode ser conduzida para examinar como a tolerância a falhas de software e a QoE evoluem ao longo do tempo. Estudos futuros podem acompanhar os usuários durante períodos prolongados, a fim de observar mudanças em suas percepções e comportamentos em resposta a atualizações de software, melhorias de funcionalidade e correção de bugs. Isso ajudaria a entender melhor a dinâmica da tolerância e a satisfação do usuário com o tempo.

6.3.3 Investigação de intervenções para melhorar a Tolerância

Pesquisas futuras poderiam investigar intervenções específicas projetadas para melhorar a tolerância dos usuários a falhas de software. Experimentos podem ser realizados

para testar a eficácia de diferentes abordagens, como *feedback* transparente sobre falhas, suporte técnico proativo, ou treinamento do usuário. Identificar quais intervenções são mais eficazes pode ajudar desenvolvedores e designers a criar softwares mais resilientes às percepções negativas dos usuários.

6.3.4 Exploração de fatores culturais e demográficos

Outro caminho promissor para futuras pesquisas é a exploração de como fatores culturais e demográficos influenciam a tolerância a falhas de software. Estudos comparativos entre diferentes regiões geográficas e grupos demográficos podem revelar variações significativas na tolerância e nas expectativas dos usuários. Entender essas diferenças pode ajudar a personalizar abordagens de desenvolvimento e suporte de software para atender melhor às necessidades diversificadas dos usuários globais.

6.3.5 Modelagem preditiva de tolerância a falhas

Finalmente, a construção de modelos preditivos para a tolerância a falhas de software pode ser uma área frutífera para a pesquisa futura. Utilizando técnicas de aprendizado de máquina e análise de big data, pesquisadores podem desenvolver modelos que prevejam a tolerância dos usuários com base em seus perfis de uso, histórico de interações e características demográficas. Esses modelos podem ser aplicados proativamente para melhorar a QoE e reduzir a frustração do usuário.

Explorar essas direções pode não apenas aprofundar a compreensão dos fatores que influenciam a tolerância a falhas de software, mas também fornecer soluções práticas para melhorar a satisfação e a experiência geral do usuário.

Referências

ASSUNÇÃO, L. M. **Confiabilidade de Software e Qualidade da Experiência do Usuário: Um Estudo Quanti-Qualitativo**. Dissertação (Mestrado) — UNIVERSIDADE FEDERAL DE UBERLÂNDIA, Uberlândia, 2019. Disponível em: <<https://doi.org/10.14393/ufu.di.2023.7117>>.

AVIZIENIS, A. et al. Basic concepts and taxonomy of dependable and secure computing. **IEEE Trans. Dependable Secur. Comput.**, IEEE Computer Society Press, Washington, DC, USA, v. 1, n. 1, p. 11–33, jan 2004. ISSN 1545-5971. Disponível em: <<https://doi.org/10.1109/TDSC.2004.2>>.

BULGARELLA, F. et al. Performance measurements of quic communications. In: **Proceedings of the 2019 Applied Networking Research Workshop**. New York, NY, USA: Association for Computing Machinery, 2019. (ANRW '19), p. 8–14. ISBN 9781450368483. Disponível em: <<https://doi.org/10.1145/3340301.3341127>>.

CHEUNG, R. A user-oriented software reliability model. **IEEE Transactions on Software Engineering**, SE-6, n. 2, p. 118–125, 1980. Disponível em: <<https://doi.org/10.1109/TSE.1980.234477>>.

CHOW, S.-C. et al. **Sample size calculations in clinical research**. [S.l.]: CRC press, 2017.

CLULEY, J. 32 - reliability in instrumentation and control. In: BOYES, W. (Ed.). **Instrumentation Reference Book (Third Edition)**. Third edition. Burlington: Butterworth-Heinemann, 2003. p. 734–792. ISBN 978-0-7506-7123-1. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780750671231500363>>.

COHEN, J. **Statistical power analysis for the behavioral sciences**. [S.l.]: Academic press, 2013.

DEMARCO, T.; LISTER, T. **Peopleware: productive projects and teams**. [S.l.]: Addison-Wesley, 2013.

DOS SANTOS, C. A. R.; MATIAS JR, R. Exploratory analysis on failure causes in a mass-market operating system. **ACM SIGOPS Operating Systems Review**, ACM New York, NY, USA, v. 50, n. 1, p. 18–30, 2016.

- DOS SANTOS, C. A. R.; MATIAS JR, R.; TRIVEDI, K. S. A multisite characterization study on failure causes in system and applications software. In: **IEEE. 2021 XI Brazilian Symposium on Computing Systems Engineering (SBESC)**. [S.l.], 2021. p. 1–8.
- FASSNACHT, M.; KOESE, I. Quality of electronic services: Conceptualizing and testing a hierarchical model. **Journal of Service Research**, v. 9, n. 1, p. 19–37, 2006. Disponível em: <<https://doi.org/10.1177/1094670506289531>>.
- FLINDERS, K. **HSBC suffers major online banking failure**. 2016. Disponível em: <<https://www.computerweekly.com/news/4500269864/HSBC-suffers-major-online-banking-failure>>.
- GEFEN, D. Customer loyalty in e-commerce. **Journal of the association for information systems**, v. 3, n. 1, p. 2, 2002.
- HAMBURGER, E. **Why Telegram has become the hottest messaging app in the world**. 2014. Disponível em: <<https://web.archive.org/web/20160313050911/http://www.theverge.com/2014/2/25/5445864/telegram-messenger-hottest-app-in-the-world>>.
- IEEE. Ieee standard glossary of software engineering terminology. **IEEE Std 610.12-1990**, p. 1–84, 1990.
- ITU. E. 800, definitions of terms related to quality of service. **International Telecommunication Union's Telecommunication Standardization Sector (ITU-T) Std**, 2008. Disponível em: <<https://handle.itu.int/11.1002/1000/9524>>.
- JALOTE, P. et al. Measuring reliability of software products. In: CITESEER. **15th International Symposium on Software Reliability Engineering**. [S.l.], 2004.
- KANO, N. Attractive quality and must-be quality. **Journal of the Japanese society for quality control**, v. 31, n. 4, p. 147–156, 1984.
- KOTLER, P.; KELLER, K. Marketing management. In: PEARSON (Ed.). **Administração de marketing**. 15. ed. São Paulo: [s.n.], 2016. ISBN 978-0-13-385646-0.
- KUIPERS, F. et al. Techniques for measuring quality of experience. In: OSIPOV, E. et al. (Ed.). **Wired/Wireless Internet Communications**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. p. 216–227. ISBN 978-3-642-13315-2.
- KUSTERS, R. J. et al. User-perceptions of embedded software reliability. In: __. **Reliability, Quality and Safety of Software-Intensive Systems: IFIP TC5 WG5.4 3rd International Conference on Reliability, Quality and Safety of Software-Intensive Systems (ENCRESS '97), 29th–30th May 1997, Athens, Greece**. Boston, MA: Springer US, 1997. p. 67–82. ISBN 978-0-387-35097-4. Disponível em: <https://doi.org/10.1007/978-0-387-35097-4_6>.
- LAPRIE, J.-C. Dependable computing: Concepts, limits, challenges. In: **Proceedings of the Twenty-Fifth International Conference on Fault-Tolerant Computing**. USA: IEEE Computer Society, 1995. (FTCS'95), p. 42–54. ISBN 0818671467.
- LIKERT, R. **A Technique for the Measurement of Attitudes**. [S.l.: s.n.], 1932.

LOVELOCK, C.; WIRTZ, J.; HEMZO, M. Marketing de serviços: pessoas, tecnologia e estratégia. In: SARAIVA (Ed.). **Marketing de serviços: pessoas, tecnologia e estratégia**. [S.l.: s.n.], 2011. ISBN 8571441227.

LYU, M. R. Software reliability engineering: A roadmap. In: **Future of Software Engineering (FOSE '07)**. [s.n.], 2007. p. 153–170. Disponível em: <<https://doi.org/10.1109/FOSE.2007.24>>.

MEIRELLES, F. S. Pesquisa do uso da ti-tecnologia de informação nas empresas. **Fundação Getúlio Vargas**, 2022. Disponível em: <https://eaesp.fgv.br/sites/eaesp.fgv.br/files/u68/pesti-fgvicia-2023_0.pdf>.

MÖLLER, S.; RAAKE, A. **Quality of experience: advanced concepts, applications and methods**. [S.l.]: Springer, 2014.

MURPHY, B.; LEVIDOW, B. **Windows 2000 Dependability**. [S.l.], 2000. 8 p. Disponível em: <<https://www.microsoft.com/en-us/research/publication/windows-2000-dependability/>>.

MUSA, J. D. Operational profiles in software-reliability engineering. **IEEE software**, IEEE, v. 10, n. 2, p. 14–32, 1993.

OLIVER, R. **Satisfaction: A Behavioral Perspective on the Consumer**. [S.l.]: M.E. Sharpe, 2010. ISBN 9780765617705.

PARASURAMAN, A.; BERRY, L. L.; ZEITHAML, V. A. Understanding customer expectations of service. **Sloan management review**, v. 32, n. 3, p. 39–48, 1991.

PARASURAMAN, A.; ZEITHAML, V. A.; BERRY, L. Servqual: A multiple-item scale for measuring consumer perceptions of service quality. **1988**, v. 64, n. 1, p. 12–40, 1988.

PARASURAMAN, A.; ZEITHAML, V. A.; MALHOTRA, A. E-s-qual: A multiple-item scale for assessing electronic service quality. **Journal of Service Research**, v. 7, n. 3, p. 213–233, 2005. Disponível em: <<https://doi.org/10.1177/1094670504271156>>.

PHAM, H. **System software reliability**. [S.l.]: Springer Science & Business Media, 2007. ISBN 978-1-84628-295-9.

PRAKASH, G. et al. Understanding service quality: insights from the literature. **Journal of Advances in Management Research**, Emerald Group Publishing, v. 16, n. 1, p. 64–90, 2019.

PRESSMAN, R. **Software Engineering: A Practitioner's Approach**. [S.l.]: McGraw-Hill Education, 2010. (McGraw-Hill higher education). ISBN 9780073375977.

SHEPARDSON, D. **Fiat Chrysler recalls 5.3 million vehicles for cruise control defect**. 2018. Disponível em: <<https://www.reuters.com/article/us-fiat-chrysler-recall-idUSKCN1IQ1QY>>.

SWEBOK et al. **Guide to the Software Engineering Body of Knowledge (SWEBOK(R)): Version 3.0**. 3rd. ed. Washington, DC, USA: IEEE Computer Society Press, 2014. ISBN 0769551661.

TVERSKY, A.; KAHNEMAN, D. The framing of decisions and the psychology of choice. **science**, American Association for the Advancement of Science, v. 211, n. 4481, p. 453–458, 1981.

Uptime Institute. Annual outage analysis. 2023, 2021, 2020.

VARELA, M.; SKORIN-KAPOV, L.; EBRAHIMI, T. **Quality of Service Versus Quality of Experience**. Cham: Springer International Publishing, 2014. 85–96 p. ISBN 978-3-319-02681-7. Disponível em: <https://doi.org/10.1007/978-3-319-02681-7_6>.

WANG, D.; TRIVEDI, K. S. Modeling user-perceived reliability based on user behavior graphs. **International Journal of Reliability, Quality and Safety Engineering**, World Scientific, v. 16, n. 04, p. 303–329, 2009. Disponível em: <<https://doi.org/10.1142/S0218539309003411>>.

WOLFINBARGER, M.; GILLY, M. C. etailq: dimensionalizing, measuring and predicting etail quality. **Journal of Retailing**, v. 79, n. 3, p. 183–198, 2003. ISSN 0022-4359. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0022435903000344>>.

YABE, D. et al. Influence factors on the quality of user experience in os reliability: A qualitative experimental study. In: **Proceedings of the 13th International Conference on Availability, Reliability and Security**. New York, NY, USA: Association for Computing Machinery, 2018. (ARES 2018). ISBN 9781450364485. Disponível em: <<https://doi.org/10.1145/3230833.3230862>>.

ZEITHAML, V.; BITNER, M.; GREMLER, D. **Services Marketing: Integrating Customer Focus Across the Firm**. [S.l.]: McGraw-Hill Education, 2017. ISBN 9780078112102.

Apêndices

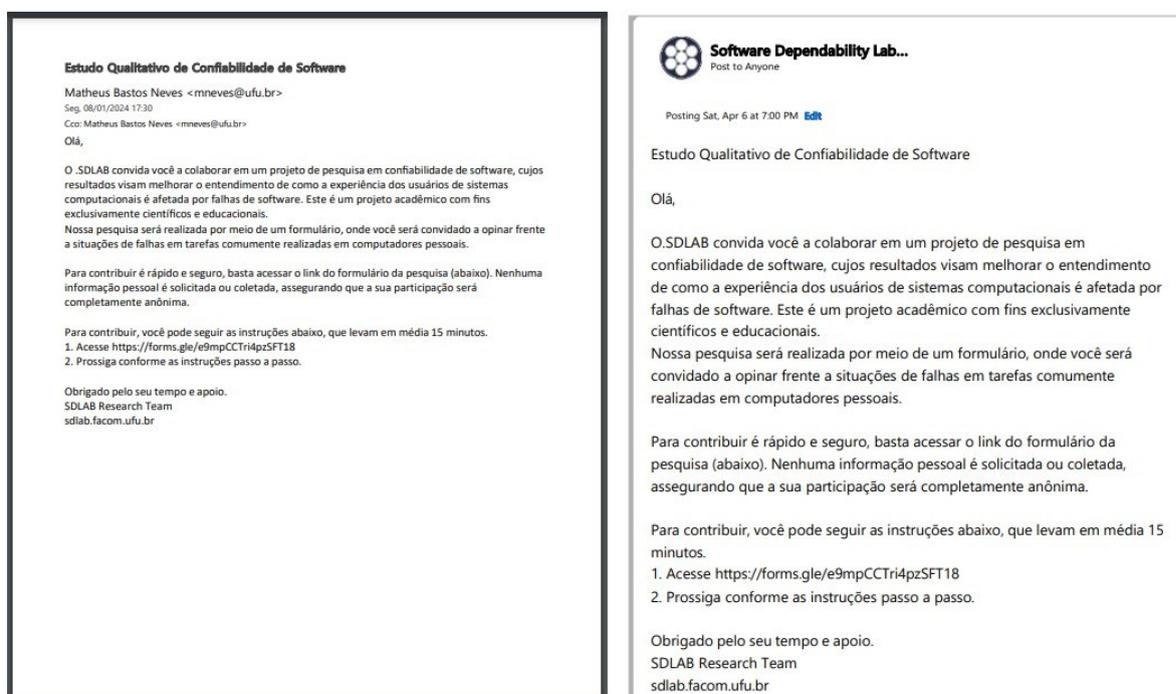
Dados Coletados

Os dados coletados durante a pesquisa foram obtidos entre 30 de abril de 2024 e 11 de junho de 2024. A amostra consiste em respostas de participantes, que avaliou a percepção dos usuários sobre a tolerância a falhas de software. A seguir, apresentamos a tabela com os resultados butos coletados. Esses dados serão utilizados para realizar análises descritivas e fatoriais exploratórias, conforme detalhado nos capítulos anteriores.

A coleta de dados foi realizada por meio de um questionário online distribuído a diversos participantes. O questionário foi composto por perguntas que utilizavam a escala Likert de tolerância para medir diferentes atributos, como qualidade, confiabilidade, utilidade, valor percebido, entre outros. Os participantes foram instruídos a avaliar diversos cenários de uso de software, tanto com sucesso quanto com falhas, fornecendo suas percepções sobre a experiência em cada caso.

Anexos

B.0.1 Convite ao participante



Estudo Qualitativo de Confiabilidade de Software

Matheus Bastos Neves <mneves@ufu.br>
Seg, 08/01/2024 17:30
Cco: Matheus Bastos Neves <mneves@ufu.br>
Olá,

O .SDLAB convida você a colaborar em um projeto de pesquisa em confiabilidade de software, cujos resultados visam melhorar o entendimento de como a experiência dos usuários de sistemas computacionais é afetada por falhas de software. Este é um projeto acadêmico com fins exclusivamente científicos e educacionais.

Nossa pesquisa será realizada por meio de um formulário, onde você será convidado a opinar frente a situações de falhas em tarefas comumente realizadas em computadores pessoais.

Para contribuir é rápido e seguro, basta acessar o link do formulário da pesquisa (abaixo). Nenhuma informação pessoal é solicitada ou coletada, assegurando que a sua participação será completamente anônima.

Para contribuir, você pode seguir as instruções abaixo, que levam em média 15 minutos.

1. Acesse <https://forms.gle/e9mpCCTri4pz5FT18>
2. Prossiga conforme as instruções passo a passo.

Obrigado pelo seu tempo e apoio.
SDLAB Research Team
sdlab.facom.ufu.br

Software Dependability Lab...
Post to Anyone

Posting Sat, Apr 6 at 7:00 PM [Edit](#)

Estudo Qualitativo de Confiabilidade de Software

Olá,

O.SDLAB convida você a colaborar em um projeto de pesquisa em confiabilidade de software, cujos resultados visam melhorar o entendimento de como a experiência dos usuários de sistemas computacionais é afetada por falhas de software. Este é um projeto acadêmico com fins exclusivamente científicos e educacionais.

Nossa pesquisa será realizada por meio de um formulário, onde você será convidado a opinar frente a situações de falhas em tarefas comumente realizadas em computadores pessoais.

Para contribuir é rápido e seguro, basta acessar o link do formulário da pesquisa (abaixo). Nenhuma informação pessoal é solicitada ou coletada, assegurando que a sua participação será completamente anônima.

Para contribuir, você pode seguir as instruções abaixo, que levam em média 15 minutos.

1. Acesse <https://forms.gle/e9mpCCTri4pz5FT18>
2. Prossiga conforme as instruções passo a passo.

Obrigado pelo seu tempo e apoio.
SDLAB Research Team
sdlab.facom.ufu.br

Figura 15 – Mensagem de divulgação com o convite de participação.

B.0.2 Parecer CEP

30/04/2024, 07:00 Plataforma Brasil

Portal do Governo Brasileiro

principal X Sair

Público Pesquisador Alterar Meus DadosMATHEUS BASTOS NEVES | V4.0.1-RC01

Sua sessão expira em: 37min 39

Você está em: Público > Confirmar Aprovação pelo CAAE ou Parecer

CONFIRMAR APROVAÇÃO PELO CAAE OU PARECER

Informe o número do CAAE ou do Parecer:

Número do CAAE: <input type="text" value="77973124.0.0000.5152"/>	Número do Parecer: <input type="text"/>
--	--

Esta consulta retorna somente pareceres aprovados. Caso não apresente nenhum resultado, o número do parecer informado não é válido ou não corresponde a um parecer aprovado.

DETALHAMENTO

Título do Projeto de Pesquisa:

Estudo exploratório sobre a tolerância de usuários a diferentes níveis de confiabilidade de software.

Número do CAAE: <input type="text" value="77973124.0.0000.5152"/>	Número do Parecer: <input type="text" value="6794483"/>
Quem Assinou o Parecer: <input type="text" value="ALEANDRA DA SILVA FIGUEIRA SAMPAIO"/>	Pesquisador Responsável: <input type="text" value="RIVALINO MATIAS JUNIOR"/>
Data Início do Cronograma: <input type="text" value="01/03/2024"/>	Data Fim do Cronograma: <input type="text" value="31/10/2024"/>
Contato Público: <input type="text" value="RIVALINO MATIAS JUNIOR"/>	

Suporte a sistemas: 136 - opção 8
e-mail: suporte.sistemas@datasus.gov.br
Fale conosco: <http://datasus.saude.gov.br/fale-conosco>

MINISTÉRIO DA SAÚDE

https://plataformabrasil.saude.gov.br/visao/publico/indexPublico.jsf 1/2

B.0.3 Questionário

Questionário Experiência do Usuário

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Você está sendo convidado(a) a participar da pesquisa intitulada “Estudo exploratório sobre a tolerância de usuários a diferentes níveis de confiabilidade de software”, sob a responsabilidade dos pesquisadores Rivalino Matias Júnior e Matheus Bastos Neves.

Nesta pesquisa nós estamos buscando compreender a tolerância dos usuários em relação a falhas de software e analisar e identificar fatores que influenciam a relação com a Qualidade da Experiência (QoE) e a satisfação geral percebida pelos usuários.

O Termo/Registro de Consentimento Livre e Esclarecido está sendo obtido de forma virtual antes do início da sua participação na pesquisa e coleta de dados. A pesquisa consiste em um formulário auto aplicado que é acessado após o consentimento deste termo. Antes de concordar em participar da pesquisa, você pode entrar em contato com os pesquisadores, para discutir as informações do estudo através do e-mail: mneves@ufu.br.

Você tem o tempo que for necessário para decidir se quer ou não participar da pesquisa (conforme item IV da Resolução nº 466/2012 ou Capítulo. III da Resolução nº 510/2016).

Na sua participação, apresentaremos sete cenários com situações de uso de um software. Pediremos que você se imagine utilizando esse software em cada situação e que responda a um questionário com a sua avaliação e percepção e assim poderemos coletar informações relevantes sobre o que você considerou em sua experiência de uso. O tempo médio de resposta do questionário é de aproximadamente 15 minutos.

Você tem o direito de não responder qualquer questão, sem necessidade de explicação ou justificativa para tal.

Você não terá nenhum gasto e nem ganho financeiro por participar na pesquisa.

Esta pesquisa é exclusiva para maiores de 18 anos.

Nós, pesquisadores, atenderemos as orientações das Resoluções nº 466/2012, Capítulo XI, Item XI.2: f e nº 510/2016, Capítulo VI, Art. 28: IV - manter os dados da pesquisa em arquivo, físico ou digital, sob sua guarda e responsabilidade, por um período mínimo de 5 (cinco) anos após o término da pesquisa.

Os resultados da pesquisa serão publicados e ainda assim a sua identidade será preservada.

Os riscos consistem em comprometer a privacidade e segurança devido ao ambiente virtual ou o desencadeamento de algum tipo de cansaço. Ao participar da pesquisa, queremos enfatizar que suas respostas permanecerão anônimas. Caso perceba cansaço ou perda de foco durante a participação da pesquisa, nestes casos recomendamos o participante deixar a pesquisa. Para minimizar alguns riscos do ambiente virtual, é importante que você tenha todo o cuidado com a segurança e privacidade do seu local quando realizar o acesso às etapas virtuais da pesquisa para que sejam garantidos o sigilo e a confidencialidade necessários. A plataforma de formulários utilizada adota práticas de segurança que permitem a transmissão segura de dados pela internet e garante o anonimato em relação aos mecanismos de rastreamento como ips, cookies de marcação entre outros. Os dados coletados serão salvos na plataforma e excluídos assim que a pesquisa for encerrada.

Esta pesquisa não oferece benefícios diretos, como pagamentos, ou qualquer tipo de vantagem financeira ou outros meios de recompensa. No entanto, de maneira indireta, o único benefício proporcionado é a satisfação por contribuir com um estudo científico.

Você é livre para deixar de participar da pesquisa a qualquer momento sem qualquer prejuízo ou coação. Até o momento da submissão dos resultados, você também é livre para solicitar a retirada

dos seus dados da pesquisa. Após este estágio, caso o participante solicite retirada de suas respostas do banco de dados, não será possível, uma vez que os dados coletados são anônimos.

Este Termo de Consentimento Livre e Esclarecido deve ser salvo nos seus arquivos clicando no link:

http://docs.google.com/document/d/1t15hYcihdx4JEHVuEyTG6jTx0GtEzY_leTbKg3LOEt4/export?format=pdf.

Em qualquer momento, caso tenha qualquer dúvida ou reclamação a respeito da pesquisa, você poderá entrar em contato com o pesquisador responsável, Rivalino Matias Júnior através do e-mail: rivalino@ufu.br.

Havendo algum dano decorrente da pesquisa, você terá direito a solicitar indenização através das vias judiciais (Código Civil, Lei 10.406/2002, Artigos 927 a 954 e Resolução CNS nº 510 de 2016, Artigo 19).

Para obter orientações quanto aos direitos dos participantes de pesquisa acesse a cartilha no link: https://conselho.saude.gov.br/images/comissoes/conep/img/boletins/Cartilha_Direitos_Participantes_de_Pesquisa_2020.pdf.

Você poderá também entrar em contato com o Comitê de Ética na Pesquisa com Seres Humanos – CEP, da Universidade Federal de Uberlândia, localizado na Av. João Naves de Ávila, nº 2121, bloco A, sala 224, *campus* Santa Mônica – Uberlândia/MG, 38408-100; pelo telefone (34) 3239-4131 ou pelo e-mail cep@propp.ufu.br. O CEP/UFU é um colegiado independente criado para defender os interesses dos participantes das pesquisas em sua integridade e dignidade e para contribuir para o desenvolvimento da pesquisa dentro de padrões éticos conforme resoluções do Conselho Nacional de Saúde.

* Indica uma pergunta obrigatória

1. Aceita participar da pesquisa? *

Marcar apenas uma oval.

- Eu aceito participar do projeto citado acima, voluntariamente, após ter sido devidamente esclarecido e declaro ter idade superior a 18 anos.
- Eu não desejo participar do projeto citado acima.

Cenário 1 de 7

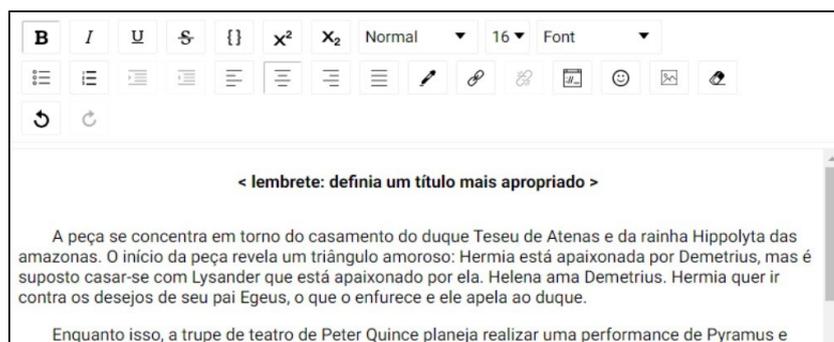
Imagine que você é um estudante universitário e está finalizando seu trabalho acadêmico de conclusão de curso. A data de entrega está se aproximando e você precisa garantir que o trabalho esteja bem formatado. Para realizar as tarefas de edição e formatação, você utiliza o editor de texto que você tem mais familiaridade, instalado no seu computador. Imagine que você está abrindo o arquivo com o seu trabalho.



Então, você localiza o seu arquivo e aciona a função “Abrir” do software de edição de texto.



Com seu documento aberto, você facilmente identifica o ponto que precisa ser alterado e realiza a alteração necessária.



Após finalizar a edição, você utiliza a função impressão do arquivo para finalizar a conversão para PDF e salva o arquivo.

4. Com base no cenário que você imaginou, em uma escala de 0 a 10, qual é o **nível de satisfação** que você atribuiria ao resultado proporcionado pelo software de edição de texto?

Marcar apenas uma oval.

0 1 2 3 4 5 6 7 8 9 10

Muito Muito Satisfeito(a)

5. Com base no cenário que você imaginou, em uma escala de 0 a 10, o quanto **o editor de texto atendeu as suas necessidades** ao selecionar, abrir, editar e salvar versão para impressão do arquivo?

Marcar apenas uma oval.

0 1 2 3 4 5 6 7 8 9 10

Total Totalmente aceitável

6. Com base no cenário que você imaginou e nas funções de um software de edição de texto para selecionar, abrir, editar e salvar versão para impressão de um arquivo, **selecione apenas as afirmativas com as quais você concorda.**

Marcar tudo o que for aplicável.

- É indispensável que o software de edição de texto responda efetivamente ao executar operações de edição de texto.
- Espero que o software de edição de texto permita abrir, editar e salvar documentos de texto de maneira ágil e intuitiva.
- Valorizo recursos adicionais no software, como formatação avançada, opções de colaboração e estilos personalizados, que enriqueçam minha experiência de uso.
- O atendimento das funções básicas de edição de texto é suficiente.
- Acredito que os softwares de edição de texto não oferecem recursos suficientes, o que torna um desafio atingir um bom resultado ao usá-los.

7. Com base no cenário que você imaginou e nas funções de um software de edição de texto para selecionar, abrir, editar e salvar versão para impressão, **selecione apenas as afirmativas com as quais você concorda.**

Marcar tudo o que for aplicável.

- A qualidade do software atende às minhas necessidades em relação ao meu objetivo.
- A qualidade do software é insuficiente para atender às minhas necessidades com relação ao meu objetivo.
- Tenho confiança de que conseguirei alcançar meus objetivos utilizando esse software.
- É improvável que eu consiga alcançar meus objetivos com esse software.
- As funcionalidades desse software me proporcionam uma sensação de realização.
- As funcionalidades desse software dificultam que eu atinja meus objetivos.
- Acredito que consigo operar esse software.
- Esse software é muito difícil e tenho dificuldade em operá-lo.
- As funcionalidades do software estão alinhadas com as promessas feitas para esse tipo de produto.
- Esse software deixa a desejar em relação a algumas das funcionalidades prometidas para esse tipo de produto.
- Esse software me proporciona benefícios claros.
- Os benefícios desse software são inferiores às minhas expectativas.
- Considero que esse software reduz meu esforço quando escolho utilizá-lo.
- A utilização desse software demanda muito esforço.
- Esse modelo de software atende perfeitamente às minhas necessidades.
- Acredito que há opções mais apropriadas no mercado que estou disposto(a) a considerar.
- Minha experiência prévia com esse tipo de software me permitiu utilizá-lo com eficácia.
- Tive experiências frustrantes com esse tipo de software no passado.
- Acredito que as falhas desse software me afetam significativamente.
- Acredito que as falhas desse software são superáveis e causam um impacto mínimo em mim.
- Esse software me ajuda a concluir meus objetivos no menor tempo possível.
- Quando uso esse software, tenho mais trabalho e preciso de tempo extra para concluir meus objetivos com sucesso.

Cenário 2 de 7

Imagine que você é um estudante universitário e está finalizando seu trabalho acadêmico de conclusão de curso. A data de entrega está se aproximando. Para garantir que o trabalho esteja bem formatado, você precisa realizar as tarefas de edição e formatação. Para isso, você utiliza o editor de texto instalado no seu computador. Ao abrir o arquivo com o trabalho, você identifica o ponto que precisa ser alterado e realiza a edição necessária.

Após realizar alguns ajustes no texto, você se depara com uma situação inesperada. O software para de responder, “trava” e uma mensagem do sistema operacional é exibida informando que “O aplicativo parou de funcionar”.

10. Com base no cenário que você imaginou, em uma escala de 0 a 10, o quanto **o editor de texto atendeu as suas necessidades** ao selecionar, abrir, editar e salvar versão para impressão do arquivo?

Marcar apenas uma oval.

0 1 2 3 4 5 6 7 8 9 10

Tota Totalmente aceitável

11. Com base no cenário que você imaginou e nas funções de um software de edição de texto para selecionar, abrir, editar e salvar versão para impressão, **selecione apenas as afirmativas com as quais você concorda.**

Marcar tudo o que for aplicável.

- A qualidade do software atende às minhas necessidades em relação ao meu objetivo.
- A qualidade do software é insuficiente para atender às minhas necessidades com relação ao meu objetivo.
- Tenho confiança de que conseguirei alcançar meus objetivos utilizando esse software.
- É improvável que eu consiga alcançar meus objetivos com esse software.
- As funcionalidades desse software me proporcionam uma sensação de realização.
- As funcionalidades desse software dificultam que eu atinja meus objetivos.
- Acredito que consigo operar esse software.
- Esse software é muito difícil e tenho dificuldade em operá-lo.
- As funcionalidades do software estão alinhadas com as promessas feitas para esse tipo de produto.
- Esse software deixa a desejar em relação a algumas das funcionalidades prometidas para esse tipo de produto.
- Esse software me proporciona benefícios claros.
- Os benefícios desse software são inferiores às minhas expectativas.
- Considero que esse software reduz meu esforço quando escolho utilizá-lo.
- A utilização desse software demanda muito esforço.
- Esse modelo de software atende perfeitamente às minhas necessidades.
- Acredito que há opções mais apropriadas no mercado que estou disposto(a) a considerar.
- Minha experiência prévia com esse tipo de software me permitiu utilizá-lo com eficácia.
- Tive experiências frustrantes com esse tipo de software no passado.
- Acredito que as falhas desse software me afetam significativamente.
- Acredito que as falhas desse software são superáveis e causam um impacto mínimo em mim.
- Esse software me ajuda a concluir meus objetivos no menor tempo possível.
- Quando uso esse software, tenho mais trabalho e preciso de tempo extra para concluir meus objetivos com sucesso.

16. Com base no cenário que você imaginou, em uma escala de 0 a 10, qual é o **nível de satisfação** que você atribuiria ao resultado proporcionado pelo software de edição de texto?

Marcar apenas uma oval.

0 1 2 3 4 5 6 7 8 9 10

Muito Muito Satisfeito(a)

17. Com base no cenário que você imaginou, em uma escala de 0 a 10, o quanto **o editor de texto atendeu as suas necessidades** ao selecionar, abrir, editar e salvar versão para impressão do arquivo?

Marcar apenas uma oval.

0 1 2 3 4 5 6 7 8 9 10

Total Totalmente aceitável

Cenário 5 de 7

Imagine que você é um estudante universitário e está finalizando seu trabalho acadêmico de conclusão de curso. A data de entrega está se aproximando. Para garantir que o trabalho esteja bem formatado, você precisa realizar as tarefas de edição e formatação. Para isso, você utiliza o editor de texto instalado no seu computador. Ao abrir o arquivo do trabalho, você identifica o ponto que precisa ser alterado e realiza a edição necessária. Entretanto, após realizar alguns ajustes no texto o software foi encerrado subitamente, causando a perda de todas as alterações realizadas, e consequente perda de tempo e a necessidade de refazer o trabalho de edição. Isso quebrou suas expectativas em relação à promessa do produto de acelerar a sua produtividade.

Você decide então **buscar alternativas no mercado para substituir o seu software editor de texto atual**, para que o problema não ocorra novamente, afinal, você **não quer perder o prazo de entrega** do seu trabalho.

Você encontra opções similares no mercado. No entanto, o custo da licença da melhor alternativa encontrada é aproximadamente 10 vezes maior que a sua opção atual, o que representa um custo significativo. É sabido que essa decisão o afetará, pois, além do custo financeiro, você precisará se capacitar para usar o novo software, processo que demandará um esforço pessoal extra.

Agora que você se imaginou nessa situação, responda às perguntas a seguir:

22. *Mesmo sabendo que a troca do software **provavelmente levará à perda do prazo de publicação**, devido às adaptações necessárias, o que impacta seus planos e objetivos pessoais, qual a sua decisão?*

Marcar apenas uma oval.

- Comprar uma das opções de editores de texto do mercado, mesmo ciente dos impactos que terá (financeiro, esforço de aprendizado do novo software);
- Não compraria no momento e daria uma segunda chance ao editor de texto atual.

Cenário 6 de 7

Imagine que você é um estudante universitário e está finalizando seu trabalho acadêmico de conclusão de curso. A data de entrega está se aproximando. Para garantir que o trabalho esteja bem formatado, você precisa realizar as tarefas de edição e formatação. Para isso, você utiliza o editor de texto instalado no seu computador. Ao abrir o arquivo do trabalho, você identifica o ponto que precisa ser alterado e realiza a edição necessária. Entretanto, após realizar alguns ajustes no texto o software foi encerrado subitamente, causando a perda de todas as alterações realizadas, e consequente perda de tempo e a necessidade de refazer o trabalho de edição. Isso quebrou suas expectativas em relação à promessa do produto de acelerar a sua produtividade.

Você está cogitando buscar alternativas no mercado para substituir o seu software editor de texto atual, apesar do custo da licença e o esforço para capacitação. Porém, você não se decidiu ainda.

Ao tentar utilizar novamente seu software editor de texto atual, você recebe uma mensagem informando que o erro que você encontrou anteriormente foi corrigido em uma atualização que foi aplicada automaticamente.

Diante desse cenário de promessa de correção do problema no software atual, por meio da atualização recente, você avalia as opções de manter o uso do software atual ou adquirir e usar um novo software, sabendo que a adoção do novo software significaria um esforço extra que levaria à perda do prazo para publicar seu artigo, o que vai prejudicar e atrasar seus planos pessoais.

Agora que você se imaginou nessa situação, responda às perguntas a seguir:

23. Qual sua decisão?

Marcar apenas uma oval.

- Comprar uma das opções de editores de texto do mercado, mesmo me comprometendo;
- Não compraria no momento e daria uma segunda chance ao editor de texto atual.

27. Com base no cenário que você imaginou e nas funções de um software de edição de texto para selecionar, abrir, editar e salvar versão para impressão, **selecione apenas as afirmativas com as quais você concorda.**

Marcar tudo o que for aplicável.

- A qualidade do software atende às minhas necessidades em relação ao meu objetivo.
- A qualidade do software é insuficiente para atender às minhas necessidades com relação ao meu objetivo.
- Tenho confiança de que conseguirei alcançar meus objetivos utilizando esse software.
- É improvável que eu consiga alcançar meus objetivos com esse software.
- As funcionalidades desse software me proporcionam uma sensação de realização.
- As funcionalidades desse software dificultam que eu atinja meus objetivos.
- Acredito que consigo operar esse software.
- Esse software é muito difícil e tenho dificuldade em operá-lo.
- As funcionalidades do software estão alinhadas com as promessas feitas para esse tipo de produto.
- Esse software deixa a desejar em relação a algumas das funcionalidades prometidas para esse tipo de produto.
- Esse software me proporciona benefícios claros.
- Os benefícios desse software são inferiores às minhas expectativas.
- Considero que esse software reduz meu esforço quando escolho utilizá-lo.
- A utilização desse software demanda muito esforço.
- Esse modelo de software atende perfeitamente às minhas necessidades.
- Acredito que há opções mais apropriadas no mercado que estou disposto(a) a considerar.
- Minha experiência prévia com esse tipo de software me permitiu utilizá-lo com eficácia.
- Tive experiências frustrantes com esse tipo de software no passado.
- Acredito que as falhas desse software me afetam significativamente.
- Acredito que as falhas desse software são superáveis e causam um impacto mínimo em mim.
- Esse software me ajuda a concluir meus objetivos no menor tempo possível.
- Quando uso esse software, tenho mais trabalho e preciso de tempo extra para concluir meus objetivos com sucesso.

Cenário 7 de 7

Imagine que você é um estudante universitário com um trabalho acadêmico importante prestes a ser entregue e você utiliza um editor de texto para fazer as alterações necessárias na formatação. Porém, o software é encerrado abruptamente, causando a perda das modificações, necessidade de refazer as alterações, atraso na entrega do trabalho e impactos nos seus planos pessoais. Você considera uma alternativa do mercado com um custo 10 vezes maior para a compra, além do esforço para capacitação.

Porém, em uma tentativa de uso, o software atual apresenta uma mensagem informando que o erro foi corrigido em uma atualização automática.

Ao abrir o arquivo do trabalho, você realiza uma edição semelhante e após realizar alguns ajustes no texto, o software "trava", tornando-se novamente inoperante, e a mensagem de que "O aplicativo parou de funcionar" é exibida. Como resultado, você perde as alterações que já tinha realizado.

Após observar cuidadosamente, você percebe então que este erro é previsível e é causado sempre da mesma forma com uma frequência de uma vez a cada uso.

32. Qual sua decisão?

Marcar apenas uma oval.

- Comprar uma das opções de editores de texto do mercado, mesmo ciente dos impactos que terá (financeiro, esforço de aprendizado do novo software);
- Não compraria no momento e daria mais uma chance ao editor de texto atual, tolerando eventuais falhas enquanto a correção desse novo problema não é finalizada.

Caso você não tenha nenhuma alternativa no mercado e seja obrigado a utilizar o editor de texto atual por falta de opção.

33. Qual sua decisão?

Marcar apenas uma oval.

- Evitaria o uso, utilizando o software apenas quando fosse estritamente necessário.
- Daria uma nova chance ao editor de texto atual, esperando que a correção seja realizada.

Questionário de Perfil

34. Qual sua idade atual?

Marcar apenas uma oval.

- 18 anos a 27 anos
- 28 anos a 37 anos
- 38 anos a 47 anos
- Mais de 48

35. Com que gênero você se identifica mais?

Marcar apenas uma oval.

- Masculino
- Feminino
- Outro
- Prefiro não responder

36. Qual é o seu nível escolar?

Marcar apenas uma oval.

- Ensino Fundamental
- Ensino Médio Completo
- Ensino Superior Incompleto/Em Curso
- Ensino Superior Completo
- Pós-graduação (Esp./MBA/Mestrado/Doutorado)

37. Para quais finalidades você utiliza o computador?

Marcar tudo o que for aplicável.

- Ferramenta de Trabalho
- Para auxiliar nos Estudos
- Para Entretenimento
- Comunicação
- Outras Finalidades

38. Qual seu nível de proficiência em relação a utilização de computadores?

Marcar apenas uma oval.

- Iniciante (Nível Básico): Uso limitado do software. Dificuldade em realizar tarefas simples, como formatação de texto. Pouco ou nenhum conhecimento sobre recursos avançados.
- Intermediário (Nível Médio): Competência em tarefas de edição de texto básicas. Capaz de formatar texto, inserir imagens e criar listas. Conhecimento limitado de recursos avançados, como estilos avançados e automação.
- Avançado (Nível Avançado): Proficiente em uma variedade de tarefas de edição de texto. Habilidade para trabalhar com estilos avançados, modelos e tabelas. Pode utilizar recursos de colaboração e rastreamento de alterações.
- Especialista (Nível Expert): Domínio completo do software editor de texto. Capacidade de automatizar tarefas complexas e personalizar o ambiente de trabalho. Conhecimento profundo de recursos avançados e capacidade de solucionar problemas.

39. Quanto tempo por semana (aproximadamente) você costuma utilizar o computador?

Marcar apenas uma oval.

- Mais de 50h
- 40h
- 20h
- Menos de 10h
- Não tenho acesso atualmente

40. Informe sua profissão ou área de atividade:

41. Qual a sua localidade? Ex.: (Cidade ou Estado ou País)

Este conteúdo não foi criado nem aprovado pela Google.

Google Formulários