
Análise quantitativa e qualitativa da experiência do usuário em confiabilidade de sistemas operacionais

Daniela Godinho Yabe



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

FACULDADE DE COMPUTAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia

2018

Daniela Godinho Yabe

**Análise quantitativa e qualitativa da experiência do usuário
em confiabilidade de sistemas operacionais**

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

2018

Y11a
2018 Yabe, Daniela Godinho, 1991-
Análise quantitativa e qualitativa da experiência do usuário em confiabilidade de sistemas operacionais [recurso eletrônico] / Daniela Godinho Yabe. - 2018.

Orientador: Rivalino Matias Júnior.
Dissertação (mestrado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Ciência da Computação.
Modo de acesso: Internet.
Disponível em: <http://dx.doi.org/10.14393/ufu.di.2018.985>
Inclui bibliografia.
Inclui ilustrações.

1. Computação. 2. Sistemas operacionais (Computadores). 3. Software - Avaliação. 4. Software - Confiança. I. Matias Júnior, Rivalino (Orient.) II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Ciência da Computação. III. Título.

CDU: 681.3

Maria Salete de Freitas Pinheiro - CRB-6/1262

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

FACULDADE DE COMPUTAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Os abaixo assinados, por meio deste, certificam que leram e recomendam para a Faculdade de Computação a aceitação da dissertação intitulada “**Análise quantitativa e qualitativa da experiência do usuário em confiabilidade de sistemas operacionais**” por Daniela Godinho Yabe como parte dos requisitos exigidos para a obtenção do título de **Mestre em Ciência da Computação**.

Uberlândia, de agosto de 2018.



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO



Ata da defesa de DISSERTAÇÃO DE MESTRADO junto ao Programa de Pós-graduação em Ciência da Computação da Faculdade de Computação da Universidade Federal de Uberlândia.

Defesa de Dissertação de Mestrado Acadêmico: PPGCO - 06/2018

Data: 16 de agosto de 2018

Hora de início: 9 h00 min

Discente: Daniela Godinho Yabe

Matrícula: 11612CCP007

Título do Trabalho: Análise Qualitativa da Experiência do Usuário em Confiabilidade de Sistemas Operacionais.

Área de concentração: Ciência da Computação

Linha de pesquisa: Engenharia de Software

Reuniu-se na sala 1B132, Bloco 1B, Campus Santa Mônica da Universidade Federal de Uberlândia, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em Ciência da Computação assim composta: Professores doutores: Marcelo de Almeida Maia - FACOM/UFU, Lúcio Borges de Araújo - FAMAT/UFU, Guilherme Horta Travassos - COS/UFRJ e Rivalino Matias Junior - FACOM/UFU, orientador da candidata.

Ressalta-se que o Prof. Dr. Guilherme Horta Travassos participou da defesa por meio de vídeo conferência desde a cidade do Rio de Janeiro - RJ (Brasil). Os demais membros da banca e a aluna participaram *in loco*.

Iniciando os trabalhos o presidente da mesa Prof. Dr. Rivalino Matias Junior apresentou a Banca Examinadora e a candidata, agradeceu a presença do público, e concedeu à 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 a candidata. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu os conceitos finais.

Em face do resultado obtido, a Banca Examinadora considerou a candidata **aprovada**.

Esta defesa de Dissertação de Mestrado Acadêmico é 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, legislação e regulamentação interna da Universidade Federal de Uberlândia.

Nada mais havendo a tratar foram encerrados os trabalhos às 12 horas e 00 minutos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.

Prof. Dr. Marcelo de Almeida Maia
FACOM/UFU

Prof. Dr. Lúcio Borges de Araújo
FAMAT/UFU

Participou por meio de vídeo conferência

Prof. Dr. Guilherme Horta Travassos
COS/UFRJ

Prof. Dr. Rivalino Matias Junior
FACOM/UFU (Orientador)

Orientador:

Prof. Dr. Rivalino Matias Júnior

Universidade Federal de Uberlândia

Banca examinadora:

Prof. Dr. Guilherme Horta Travassos

Universidade Federal do Rio de Janeiro

Prof. Dr. Lúcio Borges de Araújo

Universidade Federal de Uberlândia

Prof. Dr. Marcelo de Almeida Maia

Universidade Federal de Uberlândia

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Data: agosto de 2018.

Autor: Daniela Godinho Yabe

Título: Análise quantitativa e qualitativa da experiência do usuário em confiabilidade de sistemas operacionais

Faculdade: Faculdade de Computação

Grau: Mestrado

Fica garantido à Universidade Federal de Uberlândia o direito de circulação e impressão de cópias deste documento para propósitos exclusivamente acadêmicos, desde que o autor seja devidamente informado.

Autor

O AUTOR RESERVA PARA SI QUALQUER OUTRO DIREITO DE PUBLICAÇÃO DESTE DOCUMENTO, NÃO PODENDO O MESMO SER IMPRESSO E REPRODUZIDO, SEJA NA TOTALIDADE OU EM PARTES, SEM A PERMISSÃO ESCRITA DO AUTOR.

À minha mãe, por ter lutado ao meu lado para que este sonho pudesse se tornar realidade.

Agradecimentos

Agradeço imensamente a Deus, por me dar forças para correr atrás de todos os meus objetivos.

Ao meu orientador, Rivalino Matias, por ter-me conduzido durante esses dois anos. Seu apoio e dedicação foram fundamentais.

Agradeço à minha família (mãe, avó Maria, tio César, pai Valtinho e irmã Rayanne) pela dedicação, orações, força e apoio incondicional durante essa fase tão importante da minha vida. Vocês são o exemplo que quero seguir.

Ao meu pai, Márcio, e minhas irmãs, que, mesmo distantes, torcem por mim.

Ao meu tio Joel, por ter-me recebido em sua casa para me auxiliar durante a realização desta pesquisa.

À Lucimeire, por ter vivido cada etapa desta pesquisa ao meu lado, apoiando-me. Você, Lu, foi o presente que ganhei no mestrado. Obrigada por tudo.

Ao Lucas, por ter-me auxiliado durante todas as etapas deste trabalho. Obrigada pela amizade, parceria e por ter-se dedicado tanto para que esse sonho se tornasse realidade. A você, Lucas, minha sincera e eterna gratidão.

Aos meus amigos do laboratório, pela convivência, ajuda e cooperação. Aos meus amigos de sempre, por entenderem minha ausência e pela força, seja com o apoio, abraço ou conselho em momentos difíceis. Agradeço, principalmente, à Alauanda e à Lauana, por me ajudarem durante as correções de texto.

À Agnes e à Luciana Caldeira (Casa Cora Carolina), por muitas vezes terem assumido o lugar de família em Uberlândia. O carinho e as risadas foram fundamentais. Quero levar essa amizade para sempre comigo. Obrigada, meninas.

Ao Erisvaldo, pela dedicação e por sempre nos receber com paciência para sanar nossas dúvidas.

Aos antigos professores da graduação, Cláudio Sousa, Joicy Xavier, Jefferson, Dawson e Eduardo, por terem continuado me auxiliando durante esta pesquisa e pelos conselhos que me ajudaram a crescer tanto profissional quanto emocionalmente.

A todos que, direta ou indiretamente, me ajudaram a conquistar esse sonho, meu sincero obrigada.

“A verdadeira medida de um homem não é como ele se comporta em momentos de conforto e convivência, mas como ele se mantém em tempos de controvérsia e desafio.”

Martin Luther King

Resumo

Como consequência da crescente importância do software na rotina diária da sociedade moderna, as suas falhas podem causar um grande impacto na experiência dos usuários. Como a maioria desses software depende de um sistema operacional (SO) para sua execução, em termos práticos um SO com baixa confiabilidade, experimentando falhas frequentes, repercute diretamente na experiência do usuário ao usar o sistema computacional. Esta pesquisa avaliou a qualidade da experiência do usuário quanto à sua confiança no funcionamento do sistema operacional. Compreender os efeitos causados pelas falhas de SO na satisfação dos usuários é um importante requisito para desenvolver sistemas operacionais confiáveis de acordo com o ponto de vista do usuário, e não apenas levando em consideração métricas quantitativas, como tem sido a norma na literatura. Primeiramente, nesta pesquisa, foi aplicado um questionário online para 121 respondentes e, posteriormente, foi realizado um estudo de observação com 120 participantes, que utilizaram uma ferramenta de software capaz de simular falhas de kernel, aplicações de SO e serviços de SO. As evidências estatísticas encontradas nos resultados mostraram que a experiência do usuário é mais afetada pela consequência gerada pelo evento de falha do que pelo próprio tipo de falha vivenciado pelo usuário. Os resultados indicam que o nível de tolerância dos usuários para falha de SO bem como o nível de confiança no seu funcionamento dependem de diferentes fatores, tais como o contexto em que o usuário está inserido no momento da falha e a consequência gerada por ela. Esses dois fatores foram aqueles que mais impactaram a qualidade da experiência do usuário (QoE). Além disso, observou-se que a expectativa de solucionar o problema e, também, a familiaridade com o evento de falha afetaram positivamente a QoE dos participantes. Esse estudo visa a contribuir para uma melhor compreensão do impacto causado por falhas do SO na QoE do usuário, objetivando melhorar a modelagem de confiabilidade do sistema operacional futuro, adicionando percepções de confiabilidade dos usuários em vez de apenas elementos de qualidade de serviço (QoS).

Palavras-chave: Sistemas operacionais; falhas de *software*; qualidade da experiência do usuário; análise qualitativa

Abstract

As a consequence of the increasing importance of software in the daily routines of our modern society, its failures can have a huge impact on the user experience. Given that most of user applications depends on an operating system (OS) to execute, in practical terms, an OS with low reliability, experiencing frequent failures, has a direct impact on the user experience when considering the computer system as a whole. Understanding the effects of OS failures on user satisfaction is an important requirement for developing reliable operating systems from the user's point of view, not just taking into account quantitative metrics as has been the norm in the literature. Firstly, we applied an online questionnaire to 121 participants and, later, a face-to-face experiment was carried out with 120 participants using a software tool capable of simulating failures in kernel subsystems, OS applications, and OS services. The statistical evidence found in the results showed that the user experience is more affected by the consequence of the failure event than by the type of fault experienced by the user. The results indicate that the level of user tolerance for OS failure as well as the level of confidence in its functioning depends on different factors, such as the context in which the user is inserted at the moment of failure and the consequence of the failure. These two factors were the ones that most impacted the quality user experience (QoE). In addition, it was observed that the expectation of solving the problem and also the familiarity with the failure event positively affected the participants' QoE. This study aims to contribute to a better understanding of the impact caused by OS failures in the user's QoE, aiming to improve the reliability modeling of the future operating system, adding users' perceptions of reliability rather than just quality of service (QoS) elements.

Keywords: Operating systems; *softwares* failures; quality user experience; qualitative analysis

Lista de Ilustrações

Figura 1 – Confiabilidade sistêmica	37
Figura 2 - Opções utilizadas no questionário presencial	46
Figura 5 - Ambiente preparado após clicar no executável “Tarefa_01”	51
Figura 7 - Tela da falha da categoria kernel de SO	52
Figura 8 - Tela do bloco de notas (notepad.exe) original e o bloco de notas desenvolvido para realizar o estudo de observação	53
Figura 9 - Tela apresentada no momento em que o documento digitado no bloco de notas é salvo.....	53
Figura 10 - Tela de impressão do Windows.....	54
Figura 11 - Mensagem de falha apresentada para o usuário no momento em que ele clica em “ok” (Figura 10) para imprimir o documento.....	54
Figura 13 - Cenário II (Set 2) do questionário online (falha da categoria de falhas de kernel).....	56
Figura 14 - Exemplo de questões utilizadas para categorização dos usuários	60
Figura 15 - Fluxo de trabalho da aplicação do questionário online	61
Figura 16 - Fluxo de trabalho da aplicação do questionário presencial	62
Figura 18 - Exemplo de pergunta apresentada na primeira etapa do questionário online.....	66
Figura 19 - Exemplo de pergunta para avaliar o conhecimento do usuário em TI (após as alterações sugeridas no pré-teste).....	66
Figura 20 - Exemplo de pergunta apresentada no questionário presencial e online para caracterização do perfil dos participantes.....	67
Figura 21 - Distribuição dos participantes por gênero	68
Figura 22 - Distribuição dos sistemas operacionais utilizados pelos participantes desta pesquisa	70
Figura 23 - Roteiros de atividades realizadas pelo participante durante a sessão de observação	71
Figura 24 - Divisão dos participantes do estudo de observação.....	72
Figura 25 - Fluxo de trabalho do estudo de observação	72
Figura 26- Exemplo de pergunta apresentada no questionário presencial (após alterações do pré-teste).....	75
Figura 27 - Ambiente do estudo de observação montado em uma visitação residencial	76
Figura 28 - Distribuição dos participantes por gênero	77
Figura 29 - Distribuição dos participantes por faixa etária	78
Figura 30 - Distribuição dos participantes por área de estudo	78
Figura 31 - Distribuição dos participantes por escolaridade	79
Figura 32- Níveis de tolerância separados por contexto e por grupo de participantes.....	85
Figura 33 - Distribuição da avaliação dos serviços do SO de acordo com cada falha apresentada	95
Figura 34 - Avaliação da qualidade da experiência do usuário.....	96

Figura 35 - Avaliação da confiança do usuário	97
---	----

Lista de tabelas

Tabela 1 - Características dos cenários descritos no Set 1	56
Tabela 2 - Características dos cenários descritos no Set 2	57
Tabela 3 - Descrição dos contextos descritos nos cenários do	58
Tabela 4 - Número de participantes da etapa 1 – questionário online.....	59
Tabela 5 - Número de participantes da etapa 2 – questionário presencial	60
Tabela 6 - Exemplo de agendamento de horários e localização respeitando a.....	63
Tabela 7 - Amostra de respostas que os usuários colocaram na área de estudo e profissão	67
Tabela 8 - Cenários que causaram maior e menor impacto de acordo com os participantes ...	68
Tabela 9 - Distribuição dos participantes por formação profissional e faixa etária	69
Tabela 10 - Distribuição dos participantes por grau de escolaridade, importância do computador e conhecimento TI	69
Tabela 11 - Distribuição dos participantes por nível de importância do computador e conhecimento em TI	69
Tabela 12 - Distribuição dos participantes por conhecimento básico em TI e grupo de usuários	70
Tabela 13 - Variáveis monitoradas durante a aplicação do pré-teste	74
Tabela 14 - Distribuição dos participantes por nível de conhecimento em TI e escolaridade .	79
Tabela 15 - Sistemas operacionais utilizados	79
Tabela 17 - Conclusão da tabela anterior (continuação da Tabela 16).....	81
Tabela 18 - Erros e acertos dos participantes ao informarem qual das falhas era consequência de uma falha do sistema operacional	81
Tabela 19 - Resultados encontrados no teste exato de Fisher	82
Tabela 20 - Justificativas apresentadas pelos participantes.....	83
Tabela 21 - Justificativas encontradas para o questionamento sobre a falha ser consequência do SO	84
Tabela 22 - Resultado do teste Qui-Quadrado (influência dos prazos) – Usuários Leigos.....	86
Tabela 23 - Resultados do teste exato de Fisher (Influência da natureza da atividade)	87
Tabela 24 - Resultados do teste exato de Fisher (Intermediários vs. Leigos)	89
Tabela 25 - Cenários de maior impacto na qualidade da experiência dos participantes (questionário online).....	90
Tabela 26 - Cenários de menor impacto na qualidade da experiência dos participantes (questionário online).....	90
Tabela 27 - Resultados do Qui-Quadrado (Leigos vs. Intermediários).....	90
Tabela 28 - Ranking dos fatores de maior impacto	92
Tabela 29 - Ranking dos fatores de menor impacto	92
Tabela 30 - Resultados do Qui-Quadrado (influência das categorias de falha)	93
Tabela 31 - Evento de falha que causou maior impacto.....	93
Tabela 32 - Ranking dos fatores que influenciaram a escolha pela falha EF04	94
Tabela 33 - Ranking dos fatores que impactaram a QoE dos participantes	97

Tabela C.1 - Participantes descartados da amostra do questionário presencial.....	110
---	-----

Lista de abreviaturas, siglas e notação

BSOD	Blue Screen of Death – “Tela Azul”
EF01	Evento de Falha 1 – Categoria de Falhas de Serviço de SO (<i>sysmain.dll</i>)
EF02	Evento de Falha 2 – Falha de Aplicação do Usuário (Bloco de notas)
EF03	Evento de Falha 3 - Categoria de Falhas de Aplicação de SO (<i>explorer.exe</i>)
EF04	Evento de Falha 4 – Categoria de Falhas de Kernel de SO (<i>VIDEO_TDR_ERROR</i>)
CEP	Comitê de Ética em Pesquisas com Seres Humanos
FACOM	Faculdade de Computação
G1	Grupo de Usuários Leigos
G2	Grupo de Usuários Intermediários
IFTM	Instituto Federal do Triângulo Mineiro
MTTF	Tempo Médio até a Falha
QoE	<i>Quality of Experience</i> – Qualidade da Experiência do Usuário
QoS	<i>Quality of Service</i> – Qualidade do Serviço
RAC	Reliability Analysis Component
SBC	Sociedade Brasileira da Computação
Set 1	Conjunto de cenários de falha 1
Set 2	Conjunto de cenários de falha 2
SRE	<i>Software Reliability Engineering</i> - Engenharia de Confiabilidade de <i>Software</i>
SO	Sistema Operacional
TI	Tecnologia da Informação
UEQ	<i>User Experience Questionnaire</i> – Questionário de experiência do usuário
UFU	Universidade Federal de Uberlândia
UX	<i>User Experience</i> – Experiência do Usuário
Win8	Microsoft Windows 8
Win10	Microsoft Windows 10

Sumário

Introdução	30
1.1 Contextualização	30
1.2 Justificativa/Motivação	31
1.3 Objetivo da Pesquisa	32
1.3.1 Objetivo Geral	32
1.3.2 Objetivos Específicos	32
1.4 Perguntas de Pesquisa e Hipóteses	32
1.5 Desenvolvimento da Pesquisa	33
1.5.1 Revisão da Literatura	33
1.5.2 Material	34
1.5.3 Métodos	34
1.6 Estrutura do Documento	35
Fundamentação Teórica	36
2.1 Confiabilidade de <i>software</i>	36
2.2 Confiabilidade de Sistemas Operacionais	37
2.3 Experiência do Usuário	40
2.4 Métodos para Avaliar a Experiência do Usuário	42
2.5 Aplicação de Questionários	45
Metodologia	48
3.1 Introdução	48
3.2 Materiais	48
3.2.1 Taxonomia	48
3.2.2 Dados de Falha	49
Instrumental	50
3.2.3	50
Questionário utilizados para avaliar a experiência do usuário	54
3.2.4	54
3.2.5 Participantes da pesquisa	58
3.3 Método	61
3.3.1 Protocolo de pesquisa	61

3.3.2	Plano de recrutamento	62
3.3.3	Análise dos dados	63
	Planejamento e criação do estudo de observação	65
4.1	Introdução	65
4.2	Questionário <i>online</i>	65
4.3	Questionário <i>presencial</i>	71
	Resultados.....	80
5.1	Q1: O usuário, ao utilizar o computador para realizar tarefas de rotina (ex.: digitar textos, procurar arquivos, assistir vídeos), consegue identificar quando uma falha de operação do computador é consequência de uma falha de SO?	80
5.2	Q2: Qual o nível de tolerância do usuário para falhas de SO?	84
5.3	Q3: Qual categoria de falha de SO possui maior impacto negativo na experiência do usuário?.....	89
5.4	Qualidade da experiência do usuário	94
	Resultados.....	99
6.1	Síntese dos resultados	99
6.2	Contribuição para literatura	100
6.3	Ameaças à validade	100
6.4	Dificuldades encontradas	101
6.5	Trabalhos futuros	101
	Referências	103
	ANEXO A - PARECER DO COMITÊ DE ÉTICA DA UNIVERSIDADE FEDERAL DE UBERLÂNDIA (UFU) - UBERLÂNDIA/MG.....	108

Introdução

1.1 Contextualização

Com a crescente busca dos seres humanos por ferramentas que auxiliem na produtividade e praticidade, os sistemas computacionais têm-se tornado cada vez mais importantes para a sociedade moderna. Esses sistemas são utilizados tanto em contextos profissionais quanto em momentos de entretenimento, estando presentes em praticamente todas as tarefas do nosso cotidiano: na comunicação, no transporte e na segurança.

Como consequência dessa variedade de cenários em que esses sistemas são utilizados atualmente, uma falha em suas operações pode acarretar desde simples inconvenientes até grandes desastres [Matias et al., 2014]. Por isso a preocupação com o desenvolvimento de sistemas confiáveis tem crescido gradativamente.

Entre outros atributos da qualidade de software, tais como funcionalidade, usabilidade, capacidade e facilidade de manutenção, a confiabilidade é geralmente aceita como o principal fator na qualidade, uma vez que quantifica as falhas de software responsáveis pela maior parte da inatividade dos sistemas computacionais [Lyu, 2007]. Desse modo, a confiabilidade de software é um atributo de qualidade essencial que tem ganhado maior importância na pesquisa e na prática da engenharia de software.

De acordo com [Xavier et al., 2014], a maioria dos trabalhos de pesquisa nesta área tem-se concentrado apenas em confiabilidade de aplicações de software. Contudo é importante destacar que a grande maioria dessas aplicações depende de um sistema operacional (SO) para a sua execução. Assim, observa-se que a qualidade da aplicação é influenciada pela qualidade do SO que a suporta. Como mencionado em [Antunes e Matias, 2014], de nada adianta um sistema computacional com hardware e aplicações altamente confiáveis se o SO não apresentar nível de confiabilidade equivalente. Isso porque a aplicação para executar suas tarefas recai sobre diversos serviços oferecidos pelo SO, o que a torna completamente dependente desse software básico.

Em termos práticos, um sistema operacional com baixa confiabilidade, experimentando falhas frequentes, repercute diretamente na experiência do usuário (User Experience - UX). Para capturar a percepção de confiança do usuário no funcionamento do sistema, é necessário ter a capacidade de distinguir diferentes tipos de falhas por meio da medida de confiabilidade. Muitas vezes, há múltiplos grupos de usuários que usam o produto de maneira muito diferente e, por isso, o impacto das falhas varia entre eles [Jalote et al., 2004].

De acordo com a [ISO 9126, 1991], a principal consideração do usuário é a usabilidade do software, desempenho e seus efeitos, sem conhecer o que há dentro do produto, como funciona, como trabalha ou como foi desenvolvido. Sendo assim, a experiência do usuário ao utilizar dado software repercute diretamente na percepção de qualidade dele.

Segundo [Hassenzahl e Tractinsky, 2006], a experiência do usuário pode ser avaliada como uma consequência do estado interno do usuário (predisposições, expectativas, motivações, humor, tolerância), as características do sistema projetado (usabilidade, confiabilidade, funcionalidade, complexidade) e o ambiente dentro do qual a interação ocorre (ambiente organizacional/social/cultural). A UX leva em consideração o perfil do usuário e o tempo em que ele está envolvido com um produto ou serviço, buscando compreender os seus fatores emocionais antes ou depois da interação [Hassenzahl, 2007].

Observa-se, contudo, que o termo UX vem sendo discutido por diferentes autores e não possui um conceito predefinido, tornando-se amplo. Sendo assim, nesta pesquisa, objetivando compreender o impacto que diferentes tipos de falhas de sistema operacional causam na experiência do usuário ao utilizar o sistema computacional como um todo, foram analisados os fatores que influenciam a satisfação do usuário no momento em que essas falhas ocorrem.

Para atingir esse propósito, neste estudo a experiência do usuário foi avaliada por meio da aplicação de dois questionários, um online e outro presencial. O primeiro questionário foi compartilhado com a lista da rede de contatos da autora desta pesquisa. O intuito desse questionário foi estruturar o estudo de observação que permitiu aos participantes presenciar diferentes falhas simuladas no sistema operacional (Microsoft Windows 7) em contextos hipotéticos de uso do computador. O segundo questionário foi aplicado durante esse estudo e objetivou avaliar fatores que impactam a experiência do usuário ao presenciar essas falhas. Levando em consideração que o termo “contexto” é genérico, nesta pesquisa ele foi restrito à natureza da atividade (profissional e lazer) e prazo (curto ou longo). Por exemplo, digitar um relatório (profissional) que precisa ser enviado para seu chefe em três horas (prazo curto).

Além desses fatores, levou-se em consideração o nível de conhecimento dos participantes em sistema operacional e suas percepções quanto à ocorrência das falhas, isto é, a compreensão dos participantes ao perceberem que naquele momento ocorreu uma falha do sistema operacional. Por fim, a tolerância do usuário também foi avaliada, ou seja, o quanto ele aceita pacientemente a ocorrência das falhas sem desistir de realizar a ação desejada.

1.2 Justificativa/Motivação

Em virtude dos fatos mencionados anteriormente, é evidente a importância de uma operação livre de falhas do sistema operacional para ajudar a confiabilidade do sistema como um todo. Ao realizar uma extensa revisão da literatura de confiabilidade de software, foi possível observar que há, significativamente, menos estudos que analisam a confiabilidade do sistema operacional. Esses trabalhos avaliaram a confiabilidade do SO, seguindo, principalmente, uma das duas próximas abordagens.

A primeira investiga a confiabilidade de sistemas operacionais considerando como falhas de SO apenas aquelas que ocorreram em nível de kernel(ex.: [Xu et al., 1999], [Ganapathi and Patterson, 2005], [Ganapathi et al., 2006], [Murphy, 2008]). A segunda abordagem leva em consideração não apenas as falhas que ocorreram no espaço do kernel, mas também aquelas que aconteceram no espaço do usuário (ex.: [Matias et al., 2013], [Matias et al., 2014], [Antunes e Matias, 2014] [Dos Santos e Matias, 2015], [Dos Santos e Matias, 2016]). Essa abordagem está fundamentada no fato de que os sistemas operacionais modernos possuem funções sendo executadas tanto no nível de kernel quanto no nível de usuário. Dessa forma, analisar a confiabilidade do SO levando em consideração apenas as falhas que ocorreram em nível de kernel não demonstra a real confiabilidade do SO [Matias et al., 2013].

Neste estudo, foi utilizada a segunda abordagem, considerando as mesmas categorias de falhas do sistema operacional discutidas em [Dos Santos, 2016], que englobam falhas no kernel do sistema operacional, serviços de SO e aplicativos de software do sistema operacional. Os dois últimos representam falhas do sistema operacional no espaço do usuário. Além disso, neste trabalho avaliou-se a experiência do usuário diante das diferentes falhas do sistema operacional Microsoft Windows 7, pois, de acordo com [NetMarketShare, 2017] e [StatCounter Global Stats, 2017], esse ainda é o sistema operacional mais utilizado atualmente em desktops. Apesar de esse sistema ser o mais utilizado, ele não apareceu como o mais utilizado na amostra de dados coletada nesta pesquisa.

Até onde foi pesquisado, não se encontrou nenhum trabalho na literatura que tenha investigado a influência que as falhas de sistemas operacionais causam na qualidade da experiência (QoE) do usuário ao usar um sistema computacional como um todo. Visto que falhas de SO afetam diretamente a qualidade da execução das aplicações de software, a experiência do usuário, com respeito à confiabilidade de um dado software, sofre a influência do sistema operacional do qual esse software depende para executar suas funções. Analisar como essas falhas são percebidas pelos usuários, ao experimentarem tais eventos, contribui para um melhor entendimento sobre a confiança do usuário no funcionamento dos sistemas operacionais. O estudo desses fatos é necessário para criar modelos de confiabilidade que forneçam estimativas mais realistas, ou seja, mais próximas do usuário quanto à confiabilidade de SO.

Sendo assim, neste trabalho tal problema foi investigado avaliando como esses diferentes tipos de falhas afetam a experiência do usuário quanto à sua confiança no funcionamento do sistema operacional, a fim de entender como eles podem ser ponderados para calcular a confiabilidade do SO.

1.3 Objetivo da Pesquisa

1.3.1 Objetivo Geral

- Analisar a qualidade da experiência do usuário quanto à sua confiança no funcionamento do sistema operacional (Microsoft Windows 7) ao utilizar o sistema computacional como um todo.

1.3.2 Objetivos Específicos

- Avaliar a relação entre falhas de SO sobre a experiência do usuário e o contexto de uso do sistema operacional na ocasião da falha;
- Avaliar o nível de tolerância do usuário em relação às falhas de SO, isto é, o quanto o usuário aceita pacientemente a ocorrência das falhas;
- Investigar a percepção do usuário quanto à ocorrência das falhas de SO.
- Quantificar a avaliação do usuário quanto à sua confiança no funcionamento de um SO com diferentes tipos de falhas.

1.4 Perguntas de Pesquisa e Hipóteses

Para compreender a influência que diferentes tipos de falhas de SO causam na experiência do usuário bem como alcançar os objetivos deste trabalho, esta pesquisa teve como propósito responder a três perguntas:

Q1. O usuário, ao utilizar o computador para realizar tarefas de rotina (ex.: digitar textos, procurar arquivos, assistir a vídeos), consegue identificar quando uma falha de operação do computador é consequência de uma falha de SO?

Q2. Qual é o nível de tolerância do usuário (baixo, médio ou alto) para falhas de SO?

- Essa tolerância depende da categoria de falha de SO (ex. *kernel*, serviço ou aplicação)?
- Essa tolerância depende do contexto em que o usuário se encontra no momento da falha?
- Qual é a quantidade média de falhas consecutivas de SO que um usuário tolera?

Q3. Qual é a categoria de falha de SO que possui maior impacto negativo na experiência do usuário?

Diante dessas perguntas de pesquisa, as seguintes hipóteses foram formuladas e avaliadas:

- Usuários (não especialistas) são aqueles que não possuem conhecimento suficiente para conseguirem diferenciar uma falha da aplicação de uma falha do SO. Apesar de não conhecerem o sistema operacional detalhadamente, eles utilizam o sistema diariamente em seu trabalho ou em momentos de entretenimento. Acredita-se que, para esses usuários, independentemente do tipo de falha que ocorreu, o sistema operacional parou de funcionar. Com base nessa hipótese, conjectura-se que a maioria dos usuários (não especialistas) tenha uma percepção de baixa confiabilidade atribuída ao SO, independentemente de este ser ou não a principal fonte das falhas experimentadas pelos usuários.

- O nível de tolerância do usuário está correlacionado ao contexto em que o usuário se encontra ao experimentar as falhas bem como ao perfil do usuário (nível de conhecimento em TI de forma geral, nível de conhecimento em SO), ao tipo e à quantidade de falhas que ocorreram consecutivamente. Acredita-se que esses quatro fatores estão relacionados com o impacto que a falha causa na experiência do usuário e devem ser considerados.

- Sabe-se que falhas que ocorrem no núcleo (*kernel*) do SO geralmente interrompem a operação do sistema computacional como um todo, portanto é senso comum que essa categoria possa causar maior impacto na experiência do usuário. Porém nossa hipótese é que as falhas de SO que ocorrem fora do núcleo (em aplicações e serviços do SO) tenham impacto tão elevado ou até maior do que as falhas em nível de núcleo (*kernel*).

Para testar essas hipóteses, os dados foram coletados por meio de dois questionários, um online e outro presencial (Seção 3.2.4). Esses dados foram organizados em tabelas de contingência compostas pelas variáveis investigadas, isto é, nível de tolerância dos usuários, sua percepção quanto à falha que ocorreu bem como eventos de falha que causaram maior impacto nele. Esses dados foram separados em dois grupos de comparação independentes (ex.: usuários leigos e intermediários) e múltiplas respostas para as variáveis dependentes (ex.: baixo, médio, alto nível de tolerância).

Após realizar essa organização dos dados, os testes qui-quadrado (χ^2) e o teste exato de Fisher [Lyman, 2010] foram aplicados com o intuito de verificar se existia alguma associação entre os grupos e as variáveis investigadas (ver Seção 3.3.3).

1.5 Desenvolvimento da Pesquisa

Esta seção apresenta, de forma geral, as principais etapas que foram realizadas no trabalho descrito nesta dissertação de mestrado.

1.5.1 Revisão da Literatura

Inicialmente, foram localizados trabalhos cujo objetivo principal estava relacionado com a fundamentação teórica desta pesquisa. As fontes de pesquisa foram livros, publicações em anais de

conferências e revistas científicas obtidas por meio de bibliotecas digitais, tais como *ACM Digital Library*, *IEEE Digital Library*, *Taylor & Francis Online*, *SpringerLink* e *Google Scholar*.

Após realizar essa busca, foram estudados os trabalhos relacionados com confiabilidade de software, confiabilidade do sistema operacional e experiência do usuário e, posteriormente, trabalhos que apresentaram métodos para mensurar a experiência do usuário.

1.5.2 Material

Para coletar os dados necessários, foram aplicados dois questionários: um online e outro presencial. A vantagem de se aplicarem questionários online é o baixo custo e tempo para coletar dados de um número maior de respondentes. No entanto esses questionários não garantem que os respondentes tenham entendido as perguntas. A fim de mitigar esse viés e confirmar os dados coletados por meio do questionário online, foi aplicado um questionário presencial durante um estudo de observação.

O questionário online foi criado por meio da ferramenta LimeSurvey [LimeSurvey, 2017] e apresentou 20 questões de múltipla escolha, que auxiliaram na criação do perfil dos respondentes (gênero, profissão, nível de conhecimento em TI) e oito diferentes cenários de falha em que os respondentes assumiram estar envolvidos. Esses cenários descreveram diferentes atividades, contextos, eventos de falhas e consequências geradas por eles (ver Seção 3.2.4). Para cada conjunto de quatro cenários, foram apresentadas quatro perguntas (duas de múltipla escolha e duas discursivas) sobre o quanto a experiência dos respondentes seria negativa, caso eles estivessem enfrentando esses cenários em seu cotidiano.

Por meio dos dados coletados no questionário online, o estudo de observação foi estruturado e, assim, foi aplicado o questionário presencial. Com os resultados coletados, foi possível avaliar a tolerância e percepções dos participantes ao presenciarem determinada falha. Para auxiliar na coleta dos dados, como não foi encontrada na literatura nenhuma ferramenta que atendesse a esse propósito, foi proposta a implementação de uma ferramenta de software capaz de simular falhas no sistema operacional Microsoft Windows 7 [Assunção, 2016]. Logo, foi possível simular três falhas de SO encontradas em [Dos Santos, 2016] e uma falha de aplicação do usuário. A falha de aplicação do usuário foi simulada com o intuito de verificar se o impacto das falhas do sistema operacional na qualidade da experiência do usuário é diferente do impacto causado por falhas de software que não sejam do SO (por exemplo, falha no editor de texto).

Para analisar os dados coletados nesta pesquisa, utilizou-se o StudioR [Rstudio, 2017]. Todos os materiais citados nesta seção foram explicados detalhadamente na Seção 3.2.

1.5.3 Métodos

Não foi encontrado nenhum método que atendesse ao propósito deste trabalho, portanto foi criado um protocolo próprio (Seção 3.3.1) para avaliar a experiência do usuário quanto à sua confiança no SO. Nele, primeiramente, foi aplicado o questionário online e, posteriormente, o questionário presencial.

Os participantes foram recrutados por meio de convites enviados para diferentes listas de e-mail. Além disso, para o recrutamento dos participantes do estudo de observação, foram enviados convites para empresas e instituições de ensino, a fim de propor a seus funcionários e alunos que participassem desta pesquisa. Esse recrutamento foi planejado objetivando incluir participantes humanos de diferentes gêneros, idades, ocupações, afiliações e cidades.

Como os dados coletados por meio da aplicação desses questionários foram categóricos, realizou-se uma análise descritiva a fim de encontrar a frequência absoluta e a relativa de cada variável investigada. Todos esses dados

foram separados em categorias de acordo com o objetivo da pesquisa (Seção 1.3) e inseridos em uma tabela de contingência. Ao finalizar essa análise, foram aplicados os testes Qui-quadrado e exato de Fisher com a finalidade de verificar se existia associação entre as variáveis e os grupos estudados. Os métodos utilizados foram explicados detalhadamente na Seção 3.3. O estudo de observação foi explicado no Capítulo 4.

1.6 Estrutura do Documento

Os capítulos desta dissertação estão organizados da seguinte forma:

O **Capítulo 2** apresenta a fundamentação teórica utilizada neste trabalho, descrevendo conceitos de confiabilidade de software, confiabilidade de sistema operacional, experiência do usuário e os métodos utilizados para avaliação UX.

O **Capítulo 3** descreve a metodologia utilizada neste trabalho, explicando os materiais e métodos utilizados.

O **Capítulo 4** apresenta o estudo de observação realizado nesta pesquisa, identificando o local e o modo como cada etapa foi conduzida.

O **Capítulo 5** discute as respostas encontradas para as perguntas de pesquisa deste trabalho.

Por fim, o **Capítulo 6** expõe as conclusões, apresentando uma síntese dos principais resultados, limitações e dificuldades encontradas no decorrer da pesquisa, as ameaças à validade, contribuições para o estado da arte e trabalhos futuros.

Fundamentação Teórica

2.1 Confiabilidade de *software*

Em [Avizienis et al., 2004], a dependabilidade de um sistema computacional é definida como a área que se preocupa em fornecer um serviço em cujo funcionamento o usuário possa confiar. Nesse trabalho, o autor apresenta os principais conceitos relacionados a essa área, descrevendo suas definições básicas, incluindo os seguintes atributos:

- **Disponibilidade:** prontidão para o serviço correto.
- **Confiabilidade:** continuidade do serviço correto.
- **Segurança:** ausência de consequências catastróficas sobre o usuário e o ambiente.
- **Integridade:** ausência de alterações impróprias no sistema.
- **Manutenibilidade:** habilidade de sofrer modificações e reparos.

Levando em consideração esses diferentes atributos da dependabilidade, esta pesquisa concentra-se na confiabilidade de software, especificamente de sistemas operacionais. A avaliação da confiabilidade é uma das maiores preocupações da comunidade de engenharia de software, dada a importância dos sistemas computacionais no dia a dia da sociedade [Li, et al., 2010]. A confiabilidade impacta diretamente o funcionamento de sistemas computacionais, portanto garantir níveis elevados para esse atributo é um importante requisito para os sistemas computacionais modernos [Dos Santos, 2016]. Sendo assim, para calcular a confiabilidade de software, é necessário, principalmente, analisar as falhas que ocorreram no sistema.

Segundo [Avizienis et al., 2004], a falha ocorre quando o serviço prestado não está em conformidade com a especificação funcional, ou quando essa especificação não descreve adequadamente a

função do sistema. Considera-se que um sistema falhou quando sua saída é recebida pelo usuário do sistema e é percebida como um resultado inesperado ou errado, ou seja, o serviço não segue a especificação funcional determinada.

O trabalho descrito em [Xavier, et al., 2014] realizou um mapeamento sistemático sobre a confiabilidade de software, objetivando avaliar os tópicos mais investigados na engenharia de confiabilidade de software (SER) e, também, o que esses trabalhos abordam. Para realizar esse estudo, os autores consideraram 12 anos de publicações em 92 veículos internacionais (revistas e conferências), totalizando uma amostra de 300 artigos. No estudo, observou-se que trabalhos sobre a avaliação de confiabilidade e análise de falhas são as pesquisas mais frequentes da área de SER recentemente. Além disso, os autores perceberam que os pesquisadores em SER estão migrando de abordagens puramente teóricas para abordagens teórico-experimentais, que são aquelas em que os pesquisadores apresentam um corpo conceitual e também realizam experimentos para avaliar tal conceito.

Em [Jalote et al., 2004], foi apresentado um modelo de cálculo de confiabilidade no qual eles consideram a confiabilidade de software como um vetor composto por taxas de diferentes tipos de falha, provenientes de diferentes instâncias do sistema analisado e de diferentes usuários. Segundo os autores, as diferentes percepções de confiabilidade de vários grupos de usuários podem ser refletidas por meio da atribuição de pesos adequados para diferentes tipos de falha. Para capturar a experiência do usuário quanto à confiabilidade do sistema, a classificação da falha tem que ser realizada a partir da perspectiva desse usuário. Os autores acreditam que medir a confiabilidade dessa forma, levando em consideração a ideia das necessidades dos usuários, pode ajudar a determinar as áreas do software que precisam ser reforçadas para melhorar a experiência deles quanto à confiabilidade do sistema.

De acordo com [Xavier, et al., 2014], a maioria dos trabalhos de pesquisa nesta área tem-se concentrado apenas em confiabilidade de aplicações de software. O trabalho descrito em [Jalote et al., 2004] apresentou um método para avaliar a confiabilidade de aplicações de software levando em consideração a perspectiva do usuário. Entretanto, segundo [Antunes e Matias, 2014], mesmo que o hardware e as aplicações em nível de usuário sejam confiáveis, a confiabilidade do sistema como um todo é prejudicada se o SO não possuir a confiabilidade necessária. De acordo com o exemplo apresentado na Figura 1, é possível observar que, apesar de as aplicações em nível de usuário e do hardware possuírem alta confiabilidade (0,99999), a confiabilidade do sistema computacional é diminuída pela menor confiabilidade do SO.

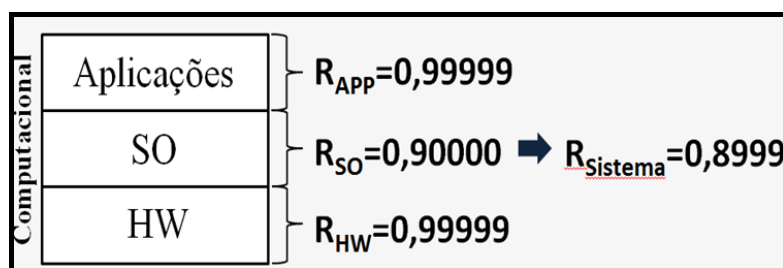


Figura 1 – Confiabilidade sistêmica

Fonte: [Antunes e Matias 2014]

Desse modo, visando a compreender as abordagens utilizadas para mensurar a confiabilidade de SO, foram estudados trabalhos relacionados a essa área. A seção seguinte descreve as definições de confiabilidade de sistemas operacionais que foram utilizadas durante o desenvolvimento desta pesquisa.

2.2 Confiabilidade de Sistemas Operacionais

Visando a compreender como a confiabilidade de sistemas operacionais é investigada, os trabalhos que retratam esse tema foram estudados.

Em [Xu et al. 1999], os autores investigaram a confiabilidade dos servidores Windows NT. Eles coletaram dados de falha no mundo real de 503 servidores ao longo de quatro meses, considerando como falhas de sistema operacional todos os eventos inesperados que levam à reinicialização/travamento/parada do sistema operacional. Eles encontraram muitos casos que exigiam mais de uma reinicialização do sistema operacional para se recuperar de uma falha, sugerindo que os modos de falha relacionados eram persistentes nessas reinicializações.

Em [Swift, 2003], os autores apresentam um subsistema de confiabilidade que melhora a confiabilidade do sistema operacional, isolando o sistema operacional das falhas dos drivers de dispositivos. Esse subsistema, chamado Nooks, impede a maioria das falhas de driver de dispositivo com pouca ou nenhuma alteração nos drivers existentes ou no código do sistema. Eles relataram que as extensões do kernel do sistema operacional (drivers de dispositivo) representavam mais de 70% do código do kernel do Linux, enquanto o Windows XP apresentava mais de 35.000 drivers de dispositivos diferentes. Eles concluíram que as extensões de kernel de terceiros eram a principal causa das falhas do Windows XP, em que os drivers eram responsáveis por 85% das falhas relatadas.

Na mesma direção, os autores em [Ganapathi et al., 2006] analisaram 2.528 eventos de falha no kernel do Windows XP de 617 computadores. Eles descobriram que os principais subsistemas do sistema operacional não eram responsáveis pela maioria das falhas. A principal causa de falhas do sistema operacional era a falta de drivers de dispositivos de terceiros. Portanto o trabalho corrobora o estudo apresentado em [Swift, 2003], enfatizando que os drivers de dispositivo contribuem para a maioria das falhas do sistema operacional.

Como pode ser observado, esses trabalhos investigaram a confiabilidade de SO considerando apenas falhas que ocorreram em nível de kernel como falhas do SO. No entanto, segundo [Matias et al., 2013], a experiência do usuário é afetada independentemente de o componente do SO ter falhado em nível de usuário ou de kernel. Desse modo, objetivando realizar um estudo exploratório sobre confiabilidade de sistemas operacionais com foco na perspectiva do usuário, esses autores utilizaram uma abordagem que considera não apenas falhas de kernel como falhas de SO, mas também falhas causadas nas aplicações e serviços de SO.

A categoria de falhas de aplicações de SO apresenta aquelas que são causadas pelo mau funcionamento das aplicações do SO. Essas aplicações geralmente estão sendo executadas em primeiro plano e realizam funções específicas, que são requisitadas de acordo com a necessidade do usuário do sistema. Um exemplo de aplicação do SO é o programa explorer.exe, presente no sistema operacional Microsoft Windows 7. Esse programa é responsável pelo gerenciamento de janelas e arquivos [Dos Santos, 2016].

A segunda categoria contém as falhas que são causadas nos serviços do SO. Esses serviços geralmente são executados em segundo plano e são usualmente iniciados durante a carga do SO. Além disso, normalmente esses serviços são utilizados por outros componentes do Microsoft Windows 7 [Dos Santos, 2016] como o svchost.exe., que é o programa responsável por carregar as bibliotecas dinâmicas utilizadas para executar cada serviço hospedado pelo próprio svchost.exe [Anson et al., 2012].

A terceira e última categoria contém as falhas que aconteceram no espaço do kernel e normalmente exigem a reinicialização do sistema quando ocorrem. Essas falhas estão relacionadas com as bibliotecas do sistema e device drivers (ex. ntdll.dll). Um exemplo de falha dessa categoria é a VIDEO_TDR_ERROR, que ocorre durante a tentativa de recarregar o driver de vídeo em computadores com o Microsoft Windows 7 [Dos Santos, 2016].

Os autores de [Matias et al., 2013] utilizaram essa abordagem e realizaram uma análise de 2.634 casos de falhas de 106 computadores reais que utilizavam o Microsoft Windows 7 como sistema operacional. As amostras de falhas foram coletadas e organizadas em dois grupos, G1 e G2. Após essa divisão, eles calcularam várias métricas de confiabilidade e, por meio dos resultados da pesquisa, os autores chegaram à conclusão de que o grupo G1 apresentava maior MTTF (tempo médio até a falha) do que o grupo G2 devido ao processo sistemático de manutenção do sistema operacional adotado no ambiente de produção de G1, especialmente no sistema de atualização. No estudo, os autores concluíram que as falhas presentes na amostra de dados do grupo G2 causaram maior impacto na experiência do usuário por ter apresentado uma percentagem maior de falhas de aplicação.

Além desse trabalho, outras pesquisas utilizaram a mesma abordagem (ex. [Matias et al., 2014], [Antunes e Matias, 2014], [Dos Santos e Matias, 2015], [Dos Santos e Matias, 2016], [Dos Santos e Matias, 2017a], [Dos Santos e Matias, 2017b]) para investigar a confiabilidade de sistemas operacionais.

Em [Matias et al., 2014], o objetivo da pesquisa foi realizar um estudo exploratório sobre confiabilidade de sistemas operacionais. Os autores analisaram 735 computadores de diferentes ambientes de trabalho que utilizavam o Microsoft Windows 7 como sistema operacional. Essa análise resultou em 30.815 falhas que foram coletadas a partir do RAC (Reliability Analysis Component) [ANSI/IEEE 1991], componente que contém todos os eventos de falha registrados desde a instalação do sistema operacional. Essas falhas foram categorizadas em aplicação de SO, serviços de SO e kernel de SO. Os autores chegaram à conclusão de que a maioria das falhas no sistema operacional é da categoria serviços de SO, sendo que a ocorrência de falhas das outras duas categorias é menos frequente e muito similar. A partir de análises estatísticas dos tempos de falha, descobriram que as distribuições de probabilidade Gamma e Weibull foram as distribuições que melhor aderiram aos dados. Também constataram que as falhas de kernel de SO ocorriam mais em ambientes corporativos do que em ambientes acadêmicos.

O trabalho [Antunes e Matias, 2014] teve como propósito apresentar um estudo experimental sobre a confiabilidade de um sistema operacional e, por essa razão, os autores analisaram uma amostra com 3.235 registros reais de falhas de SO obtidos de diferentes computadores. Eles completaram esse estudo investigando as principais causas de falha em cada uma das categorias de falhas de SO. Além disso, avaliaram quantitativamente a importância de cada uma das causas identificadas sobre confiabilidade do Microsoft Windows 7. Por meio das causas de falhas encontradas, os autores criaram modelos estocásticos de confiabilidade que permitiram avaliar a confiabilidade do SO com respeito a cada uma das causas de falhas observadas. Uma das contribuições do trabalho foi apresentar que as evidências da amostra

analisada sugerem as falhas que se encaixam na categoria de falhas de serviço de SO e a categoria de falhas de aplicação de SO que tem impacto importante na redução da confiabilidade do sistema, com ênfase nas falhas de serviços. Os autores concluíram que a análise das principais causas de falhas de SO observadas mostra que a atualização de software é o serviço de SO com maior índice de importância de confiabilidade, ou seja, o componente do sistema cuja melhoria deve ser priorizada.

Os trabalhos descritos a seguir nesta seção analisaram a confiabilidade do sistema operacional com base na amostra de mais de 7000 registros de falhas obtidos de 566 computadores de diferentes ambientes de trabalho executando o sistema operacional Microsoft Windows 7.

Em [Dos Santos e Matias, 2015], foi realizado um estudo exploratório sobre o comportamento das falhas de SO. Para isso, foram desenvolvidos cenários para investigar padrões de falhas por meio da análise das causas que ocorreram antes ou depois das falhas mais recorrentes na amostra. A partir desses dados, os autores realizaram uma análise qualitativa e quantitativa para investigar diferentes propriedades das falhas de SO analisadas. Eles concluíram que as falhas de serviço de SO, principalmente falhas de atualização de software, foram mais frequentes do que qualquer outra categoria de falha. Esse resultado confirmou que os estudos de confiabilidade de sistemas operacionais não devem se concentrar apenas em falhas de kernel. Com relação ao motivo das falhas no trabalho, foi identificado que a principal causa das falhas de atualização é o mau funcionamento do software .Net Framework. Além disso, os autores identificaram fortes evidências de correlação entre elas, uma propriedade que muitas vezes é negligenciada por estudos em confiabilidade de software.

No trabalho apresentado em [Dos Santos e Matias, 2016], foi realizada uma análise temporal das falhas de atualização de software. Como resultado, os autores apresentaram que uma falha no sistema de atualização padrão do Microsoft Windows 7 tinha sido a causa para a repetição de falhas durante tentativas sucessivas sem sucesso de instalação de atualização e, também, de falhas que tinham ocorrido em um curto período de tempo. Além disso, eles identificaram padrões de ocorrência entre diferentes falhas de SO como padrões de ocorrência durante um período muito curto de tempo e falhas que ocorreram ao mesmo tempo.

Em [Dos Santos e Matias, 2017a], foi desenvolvido um protocolo para detectar e caracterizar padrões de falha dos diferentes computadores investigados. Por meio do estudo, detectaram-se 45 padrões de falhas com 153.511 ocorrências, encontrando evidências de correlação entre falhas de componentes específicos de SO, o que ajudou a melhor compreender o comportamento dessas falhas de software. Com base nesses padrões, os autores concluíram que as falhas mais frequentes estavam relacionadas às atualizações de software dos componentes do sistema operacional. As principais causas dessas falhas envolveram infraestruturas e fatores ambientais, tais como indisponibilidade de espaço em disco e execução simultânea de serviços de SO. Além disso, com relação às falhas de aplicação de SO, o estudo identificou que o `explorer.exe`, o `mscorsvw.exe` e o `rundll32.exe` foram os programas que mais apresentaram falhas. Já em relação às falhas de kernel, os autores identificaram que os resultados estão de acordo com outros trabalhos que investigaram falhas de SO, observando que os device drivers foram as principais causas de falhas de kernel, principalmente aquelas de gerenciamento de energia e vídeo.

O trabalho descrito em [Dos Santos e Matias, 2017b] utilizou o mesmo protocolo criado em [Dos Santos e Matias, 2017a] com o intuito de analisar se as

causas de falhas de SO também seguem padrões. Eles identificaram 47 padrões entre essas causas e determinaram que a maioria delas apresentou autocorrelação, isto é, várias falhas ocorreram próximas

umas das outras em relação ao tempo com uma mesma causa raiz. Além disso, eles identificaram casos de correlação cruzada. Por exemplo, quando o .Net Framework falha, o serviço de atualização do Microsoft Windows 7 também falha, indicando uma relação de dependência entre essas falhas. Além disso, o .Net Framework foi a principal causa das falhas da amostra, apresentando que esse é o componente menos confiável do SO investigado.

Apesar de os trabalhos descritos nesta seção terem investigado a confiabilidade do sistema operacional utilizando uma abordagem que reflete a perspectiva do usuário, eles não levaram em consideração a opinião dos usuários ao vivenciarem essas falhas. Desse modo, utilizando a mesma abordagem proposta em [Matias et al., 2013], o trabalho descrito nesta dissertação buscou investigar a percepção do usuário diante das falhas mais frequentes encontradas nos trabalhos [Dos Santos e Matias, 2017], avaliando a UX em relação à confiança do usuário no funcionamento do sistema operacional.

2.3 Experiência do Usuário

O termo “experiência do usuário” foi criado por Donald Norman, enquanto era vice-presidente do Grupo de Tecnologia Avançada da Apple, entre 1995 e 1997 [Cruz, 2016]. Norman conta que criou o termo por considerar os conceitos de usabilidade e interface do usuário muito restritos e, por isso, tinha a intenção de abordar todos os aspectos da experiência de uma pessoa com um sistema [Merholz, 2007]. Sendo assim, o termo UX tem sido discutido com frequência na literatura de engenharia de software, sendo definido de diferentes maneiras.

Em [Tulis e Albert, 2013], os autores contribuíram para essa questão apresentando as diferenças entre usabilidade e UX. Segundo os autores, a usabilidade é geralmente considerada como a habilidade do usuário em executar uma tarefa com sucesso, enquanto a experiência do usuário tem um sentido mais amplo, pois avalia toda a interação do indivíduo com o sistema computacional, levando em consideração os pensamentos, sentimentos e percepções que resultam dessa interação.

O trabalho descrito em [Hassenzahl e Tractinsky, 2006] também contribuiu para essa questão, discutindo a experiência do usuário por meio de três abordagens: além do instrumental, emoção/afeto e empírico. Na primeira abordagem, os autores afirmaram que o objeto interativo tem uma função além do instrumental, pois estudos referentes à interação homem-computador concentram-se em como o usuário é capaz de realizar determinada atividade. Entretanto, não apenas de usabilidade se compõe a experiência do usuário.

A segunda abordagem mostra a importância da emoção e do afeto em alguns processos centrais do usuário. A UX deve prestar atenção ao afeto e às emoções a partir da perspectiva do usuário, pois, dessa forma, podemos entender seu papel ao utilizar o sistema.

A terceira abordagem considera que uma experiência é uma combinação de vários elementos, como o produto e o estado interno do usuário (temperamento, expectativas, objetivos, entre outros). Sendo assim, os autores definiram a UX como uma consequência do estado interno do usuário (predisposições, expectativas, necessidades, motivação e humor), as características do sistema projetado (complexidade, finalidade, usabilidade e funcionalidade) e o ambiente dentro do qual a interação ocorre (ambiente organizacional/social).

Em [Hassenzahl, 2007], assume-se que as pessoas compreendem a experiência do usuário em dois eixos independentes: hedônico e pragmático. O primeiro eixo tem como foco a pessoa, seu estado interno e

vai além do instrumental. Segundo o autor, esse eixo distingue três fatores: estimulação, identificação e evocação. A estimulação está ligada à novidade e ao crescimento pessoal, enquanto a identificação está conectada com a comunicação de própria identidade, e a evocação está relacionada à provocação de memórias e simbolismos. Ou seja, o eixo hedônico está conectado com os objetivos de “ser”: “ser divertido”, “ser relacionado a outros”, “ser especial”. O segundo eixo relaciona-se aos objetivos de “fazer”: encontrar uma imagem, escrever um texto, criar uma planilha. Esse eixo enfatiza o produto, a utilidade, a eficiência e a eficácia. A importância do trabalho deu-se em definir um modelo que apresenta como a experiência do usuário é compreendida.

Apesar desse interesse crescente pela experiência do usuário, é difícil encontrar uma definição clara e compartilhada entre os pesquisadores para o termo UX. Pensando nisso, os autores de [Law, 2011] realizaram uma pesquisa que reuniu diferentes opiniões sobre UX de 275 pesquisadores e profissionais da área acadêmica e da industrial. Como resultado, eles concluíram que a maioria dos entrevistados concorda que a UX é dinâmica, depende do contexto e é subjetiva. Sendo assim, de acordo com as questões mais controversas, os autores propuseram definir a UX como algo individual ao invés de social e que emerge da interação do usuário com o produto, o sistema, o serviço ou o objeto.

Além disso, eles concluíram que a definição apresentada pela ISO 9241 [ISO 9241, 2010] parece estar de acordo com as respostas apresentadas no questionário. Segundo esses autores, a partir das respostas obtidas por meio do questionário, foi possível recolher sistematicamente diferentes pontos de vista, científicos e práticos, sobre a natureza e escopo de UX. Como conclusão geral, os autores afirmaram que a UX é vista como algo novo que deve fazer parte do domínio da interação homem-computador e ser fundamentada por meio de práticas utilizadas no design centrado no usuário.

Desse modo, analisando os trabalhos citados, pode-se concluir que é necessário observar a experiência do usuário dentro do processo de desenvolvimento e durante a utilização do sistema para obter-se uma visão geral da UX [Bevan, 2008]. Vale destacar o quanto o papel do usuário é importante no momento de analisar e compreender a UX completamente. Sendo assim, é preciso escutar o ponto de vista do cliente ao utilizar um produto de software. Por meio dessas opiniões, é possível identificar em quais aspectos do software a melhoria deve ser priorizada. A experiência do usuário leva em consideração o tempo em que ele está envolvido com uma experiência, buscando compreender os seus fatores emocionais antes ou depois da interação. Por isso, como é o usuário que decide se irá utilizar a aplicação ou não, forte atenção deve ser dada para a sua satisfação.

Funcionalidade, confiabilidade, desempenho, usabilidade e portabilidade são as principais características consideradas pelo usuário [AL-Badareen et al., 2011]. A satisfação do usuário é fundamental para o sucesso de qualquer produto no mercado, portanto esses atributos são os principais requisitos que precisam de atenção ao se desenvolver um sistema que será aceito pelo usuário. Entretanto a qualidade de um serviço ou produto é avaliada pelo usuário não apenas investigando o desempenho técnico do sistema, mas também observando o contexto em que ele está inserido (tarefa que está sendo realizada, localização, urgência), bem como outros fatores que estão relacionados com o próprio usuário (expectativa, experiências passadas) [Schatz et al., 2013]. Desse modo, para avaliar a experiência do usuário ao utilizar determinado produto, esses fatores também precisam ser considerados.

A qualidade da experiência é definida em [Brunnström, et al., 2013] como o grau de prazer ou aborrecimento do usuário ao utilizar um aplicativo ou serviço, resultando no cumprimento de suas

expectativas em relação à utilidade e/ou prazer da aplicação em função da personalidade do usuário e estado atual. Desse modo, a definição de QoE enfatiza a importância de considerar os aspectos pragmáticos (utilidade) e hedônicos (prazer) do julgamento da qualidade do serviço [Schatz et al., 2013]. Os fatores que influenciam a qualidade da experiência não devem ser considerados de forma isolada, pois se relacionam [Brunnström et al., 2013].

Pensando nisso, vários métodos têm sido propostos para investigar a experiência do usuário com relação ao uso do software. A seção seguinte apresenta os principais métodos encontrados na literatura para realizar a mensuração da experiência do usuário.

2.4 Métodos para Avaliar a Experiência do Usuário

Ao estudar os trabalhos que abordaram a experiência do usuário apresentados na seção anterior, foi possível compreender que a UX vem sendo discutida e definida. Considerando esses conceitos, pode-se concluir que, para avaliar a experiência do usuário, é necessário identificar todos os fatores humanos relacionados ao perfil do usuário, como nível de conhecimento, emoção, expectativas, necessidades, motivações, comportamento e, também, o contexto no qual ele se encontra. É necessário, também, interpretar todos os atributos e requisitos do produto utilizado, isto é, a usabilidade, a funcionalidade, a interface, a marca, o custo, a confiabilidade, a segurança e a situação na qual ele está sendo utilizado, bem como a eficiência da interação do usuário com o produto analisado.

Entretanto, como os termos “experiência do usuário” e “qualidade da experiência” ainda não possuem um conceito definido pela academia, diferentes trabalhos apresentam definições e métodos de como mensurar essa experiência. Desse modo, surge o desafio de identificar qual é o melhor método para medir e quantificar a experiência do usuário ao interagir com determinado produto ou serviço.

Segundo [Filardi e Traina, 2008], a satisfação do usuário pode ser avaliada ou estimada por meio de medidas subjetivas ou objetivas. Medidas objetivas podem ser baseadas na observação do comportamento do usuário, como postura corporal, movimento do corpo, frequência de distração, ou por meio do monitoramento de respostas psicológicas do usuário. As medidas subjetivas de satisfação são adquiridas quantificando subjetivamente a intensidade das reações, atitudes ou opiniões expressas pelo usuário [Filardi e Traina, 2008].

Métodos subjetivos qualitativos que geralmente são utilizados para avaliar a qualidade da experiência do usuário são grupos focais, entrevistas, aplicação de questionários, entre outros [Schatz et al., 2013]. Geralmente, esses testes são conduzidos em um ambiente controlado de laboratório e existe um planejamento cuidadoso para escolher quais variáveis e fatores de influência precisam ser controlados, medidos, investigados e monitorados [Schatz et al., 2013].

O método para coletar o nível de tolerância do usuário e encontrar os eventos de falha que causam mais impacto em sua experiência foi baseado no estudo descrito em [Rivero, 2016]. Nesse trabalho, os autores apresentaram um mapeamento sistemático sobre as principais tecnologias utilizadas atualmente para avaliar a experiência do usuário, com base em 1940 artigos que foram publicados entre 2010 e 2014. Eles investigaram métodos de design UX, incluindo soluções padrões que podem ser úteis na identificação de oportunidades de melhoria nas aplicações de software. Os métodos e as ferramentas encontrados a partir dessa revisão da literatura foram classificados de acordo com a forma como foram aplicados, dividindo-os em formulários, entrevistas, escalas, entre outros. A principal contribuição da pesquisa foi apresentar diversos trabalhos relacionados à utilização e à criação de métodos para a mensuração da UX. Dessa

forma, segundo o autor, os resultados da avaliação empírica sobre as tecnologias apresentadas no trabalho podem auxiliar empresas e pesquisadores a compreender como os métodos e as ferramentas que capturam a UX são aplicados.

Ao estudar esse trabalho, foi possível compreender que as técnicas para mensurar a UX vão desde a aplicação de formulários até a utilização de dados biométricos para avaliar as emoções dos usuários. Foi observado, também, que novos métodos para avaliar a interação entre o usuário e diferentes aplicações vêm sendo propostos diariamente. A maioria dos métodos encontrados avaliam diferentes tipos de aplicações, como aplicações web e móveis. Além disso, esses métodos podem ser adaptados para capturar a experiência dos usuários em diferentes contextos, pois os artigos descrevem tecnologias que podem ser aplicadas durante o uso de um produto, porém esses métodos requerem algum tipo de observação ou monitoram a fisiologia dos usuários. Desse modo, como observado em [Rivero, 2016], há a necessidade de criar outros métodos mais fáceis de serem empregados sem que causem impacto no usuário.

Apesar de a maioria dos métodos apresentados focar a emoção do usuário no momento em que ele utilizou o sistema computacional, eles não levam em consideração outros fatores importantes que devem ser considerados ao avaliar a experiência do usuário. Quanto à percepção, muitos métodos conseguem capturá-la em relação à estética dos sistemas, principalmente aplicações web. Entretanto, não foi encontrado um método que avalie a percepção do usuário quando o sistema apresenta algo diferente do esperado. Desse modo, nota-se que existe a necessidade da criação de tecnologias capazes de capturar a experiência do usuário como um todo.

No trabalho descrito em [Bargas-Avila et al., 2011], foram identificados 96 métodos e ferramentas que foram listados em [Roto et al., 2011]. Esse site contém diferentes informações referentes à experiência do usuário, como os diferentes métodos encontrados na literatura e conceitos que definem o termo UX. Nesse estudo, os autores identificaram que os métodos de questionário para avaliar períodos curtos de experiência produziram alta confiabilidade. No entanto, utilizar esses métodos na prática pode exigir equipamentos específicos, alto conhecimento ou, até mesmo, software apropriados para coletar os dados.

Em [Schrepp et al., 2014], foi apresentado um formulário denominado questionário da experiência do usuário (UEQ - user experience questionnaire). Os autores apresentaram várias questões associadas à medição da experiência do usuário, discutindo como elas podem ser respondidas, por meio da aplicação de questionários, sem que seja necessário um esforço elevado para sua criação. Segundo os autores, diferentes grupos de usuários podem julgar o mesmo produto de forma bastante diferente em relação à sua experiência. Sendo assim, um método de baixo custo e eficiente para medir essa experiência é o uso de questionários. O objetivo principal desses questionários é permitir uma medição rápida e imediata da experiência do usuário, pois considera aspectos de qualidade pragmática e hedônica.

Os autores concluíram que o UEQ oferece a possibilidade de avaliar a experiência do usuário de forma rápida e eficiente. Além disso, observaram que a utilização do questionário permite medir a experiência do usuário utilizando diferentes versões do produto para o controle da qualidade de software. Também perceberam que é possível comparar um produto com o seu concorrente para obter informações sobre a posição comparativa do produto avaliado. Como desvantagem, porém, os autores observaram que esse questionário fornece apenas informações de alto nível sobre a resistência e fraquezas de um produto.

Embora os trabalhos descritos nesta seção tenham apresentado uma importante contribuição para o estado da arte relacionada à experiência do usuário, por meio desses estudos é possível observar que, assim

como ainda não há um consenso quanto à definição da experiência do usuário, também não existe uma técnica altamente utilizada para mensurá-la, pois a maioria desses métodos é utilizada como complemento nos testes de usabilidade, focando a eficiência e estética dos sistemas computacionais [Rivero, 2016].

Em [Niels et al., 2017], os autores investigam se as atribuições causais de usuários de computadores influenciam a percepção e a experiência do usuário ao presenciar mensagens de erro. A teoria da atribuição lida com as explicações causais que os usuários encontram para determinados eventos e como essas atribuições ocorrem. O principal objetivo da pesquisa foi investigar se existe alguma diferença entre o modo como os indivíduos atribuem falhas ao longo das dimensões atribucionais e a percepção do usuário com o sistema ou mensagens de falha.

Eles realizaram esse estudo aplicando um questionário online com 196 participantes. Nesse estudo, foi observado que as atribuições causais influenciam a percepção dos usuários com relação às mensagens de erros do sistema. Eles concluíram que os usuários com diferentes estilos de atribuição provavelmente reagirão de maneira diferente com relação à sua percepção de erro, incluindo possíveis sentimentos de frustração ou desânimo. Desse modo, os autores citam que desenvolvedores de sistemas devem prestar mais atenção ao modo como irão desenvolver as mensagens de erro para evitar efeitos negativos indesejados. Esse trabalho, entretanto, investigou apenas o impacto que as mensagens de falhas de software causam na experiência do usuário, diferentemente do trabalho descrito nesta dissertação, que avaliou a percepção do usuário diante das falhas de SO.

Desse modo, ao avaliar todos os estudos aqui citados, foi escolhida a aplicação de questionários para medir, controlar e investigar as reações, atitudes e opiniões do usuário ao presenciar diferentes falhas do sistema operacional, pois concluiu-se que esse era o método que melhor se encaixava no propósito desta pesquisa.

2.5 Aplicação de Questionários

Pesquisadores de diversas áreas parecem estar mais confiantes quanto à utilização desse método, pois o uso do questionário pode auxiliar na produção de dados de bom nível [Vieira, 2000]. Uma pesquisa feita por meio da utilização de questionários pode ser realizada de várias maneiras: pesquisas postais, pesquisas online, entrevistas por telefone, entrevistas presenciais. Em [Vieira, 2000], a autora apresenta os principais conceitos relacionados a cada uma dessas técnicas de aplicação de questionário. Esses conceitos serão descritos a seguir.

Os questionários autoaplicáveis são aqueles enviados para os respondentes para que eles mesmos os preencham. Para os respondentes, a grande vantagem desse método de aplicação é o fato de ser possível preencher o questionário no momento em que se achar possível; não há pressão de tempo ou do pesquisador. As respostas coletadas têm formato padrão, definido pelo pesquisador, facilitando a análise dos dados. Além disso, como vantagem para o pesquisador, existe a certeza de que o entrevistador não influenciou o respondente durante a coleta dos dados.

Esses questionários encaixam-se em pesquisas postais (questionários enviados por correio, por e-mail) ou pesquisas online (questionários online). Nota-se, contudo, que o levantamento dos dados por correio é demorado e de difícil obtenção de respostas para a maioria das pessoas que receberam o questionário. Por isso, recentemente, os pesquisadores têm adotado levantamento de dados por e-mail ou internet, pois os questionários são facilmente distribuídos. Além disso, a coleta de dados e o processamento das informações coletadas são rápidos.

A maior desvantagem da utilização desse método é não existirem garantias de que pessoas que receberam o questionário irão querer responder a ele. Outra desvantagem é o fato de as perguntas serem respondidas, muitas vezes, sem informações adicionais, como comentários ou justificativas.

Devido a essas desvantagens, muitos pesquisadores optam por aplicar questionários por meio de entrevistas por telefone ou presenciais. Em entrevistas por telefone, os números podem ser selecionados aleatoriamente, garantindo que vários perfis de respondentes participem da pesquisa. Contudo esse é um método caro e, muitas vezes, demorado, devido à dificuldade de encontrar os respondentes em casa.

Quando o pesquisador tem por objetivo buscar fatos específicos sobre determinado tema, os questionários autoaplicáveis são vantajosos, porém, quando o desejo é buscar pelo sentimento e reações, o melhor método é a entrevista presencial. Levando em consideração que os respondentes podem não

compreender o sentido da pergunta descrita nos questionários autoaplicáveis, a vantagem de aplicar um questionário em uma entrevista presencial é o fato de ser possível explicar detalhadamente as questões mais difíceis. Além disso, esse método garante que todas as perguntas sejam respondidas, permitindo ao pesquisador coletar informações extras como reações e opiniões dos respondentes.

Para aplicar o questionário em uma entrevista presencial, recomenda-se que o entrevistador tenha em mãos um roteiro e instruções que devem ser lidas para o participante. Por conta disso, o pesquisador e o respondente podem conversar durante a sessão de entrevista livremente.

Nesta pesquisa, foram aplicados dois questionários utilizando esses métodos. O primeiro foi aplicado por meio de uma pesquisa online, objetivando coletar fatos relacionados à qualidade da experiência do usuário quanto à sua confiança no funcionamento do sistema operacional do computador e, principalmente, resultados para planejamento do ambiente do estudo de observação e criação do questionário presencial. Em virtude dos fatos mencionados acima, com o propósito de coletar as reações dos usuários ao presenciarem um evento de falha, aplicou-se um segundo questionário durante uma entrevista presencial realizada em um ambiente do estudo de observação em laboratório.

Para criação dos questionários, levaram-se em consideração variáveis como: o perfil do usuário (nível de conhecimento em TI, em SO e os sentimentos relacionados à sua insatisfação ao presenciar a falha), o contexto no qual o usuário se encontra no momento em que a falha ocorre e o tipo de falha que ocorreu durante sua interação com o sistema operacional. Esses fatores foram escolhidos levando em consideração os conceitos apresentados neste capítulo, pois acredita-se que todas essas variáveis agrupadas formam a experiência do usuário.

Além da escolha do método de aplicação e o propósito de cada pergunta descrita nos questionários, é fundamental, durante a criação dos questionários, escolher como o pesquisador deseja que os participantes respondam às questões [Jen, 2006]. As respostas podem ser de múltipla escolha ou abertas. Nesta pesquisa, decidiu-se inserir as seguintes opções para as respostas de múltipla escolha:

- Sim/Não/Não sei responder a essa questão: essas opções foram utilizadas com o propósito de compreender se os participantes sabem definir se a falha que ocorreu é ou não consequência de uma falha de operação do sistema operacional. A Figura 2 apresenta um exemplo de aplicação dessas opções de respostas utilizadas no questionário presencial aplicado nesta pesquisa.

4. Em sua opinião, essa é uma falha específica de sistema operacional?

Sim Não

Não sei responder essa questão

Figura 2 - Opções utilizadas no questionário presencial

- Escala *Likert* de 5 pontos: Por meio dessa escala, é possível coletar níveis de satisfação, tolerância, impacto, paciência, entre outros sentimentos. Objetivando coletar o nível de confiança e de tolerância dos participantes, bem como o impacto que a falha causa em sua experiência, utilizou-se essa opção no questionário presencial. A Figura 3 apresenta um exemplo de questão descrita no questionário presencial aplicado nesta pesquisa, utilizando essa escala.

5. Marque a opção que melhor representa seu grau de tolerância para essa falha após vivenciar essa situação.

0 - Nem um pouco

1 - Um pouco

2 - Mais ou menos

3 - Muito

4 - MUITÍSSIMO

• Justifique sua resposta

Figura 3 - Exemplo de escala de 5 pontos utilizada no questionário presencial

As Figuras 2 e 3 foram retiradas do questionário aplicado durante a entrevista presencial. Observe que a questão apresentada na Figura 3 permite que o usuário apresente justificativas por meio de uma questão aberta. Ao apresentar essa opção, é possível coletar informações extras com opiniões dos participantes.

Apesar de o tempo gasto ser bastante elevado no planejamento do trabalho, na criação das questões utilizando a melhor escala de respostas e interpretação dos dados, a aplicação de questionários é uma maneira de coletar informações de um grande número de pessoas dentro de um período de tempo [Jen, 2006]. Esse método é uma das técnicas de investigação usadas para coletar opiniões, satisfações e críticas. Ele é composto por um conjunto de questões que são submetidas às pessoas com o propósito de obter informações sobre conhecimentos, sentimentos, valores, interesses, expectativas, temores, comportamento presente ou passado do respondente [Gil, 2008]. A construção de um questionário compreensível e que atenda aos requisitos da pesquisa, portanto, é de extrema importância para garantir que os dados coletados resultem em informações interpretáveis e de qualidade.

Metodologia

3.1 Introdução

Para realizar este estudo, os dados foram coletados por meio de questionários, anotações e entrevistas [Godoy, 1995], sendo caracterizado como um estudo de observação transversal. As pesquisas de observação são aplicadas para auxiliar na compreensão de comportamentos e acontecimentos no momento exato em que as informações são produzidas, sem a interferência de pessoas externas. A observação leva em consideração os detalhes e coloca o pesquisador dentro do cenário estudado de forma que ele possa compreender a complexidade do estudo, ao mesmo tempo em que lhe permite uma interlocução mais competente, auxiliando na análise de comportamentos e percepção de atitudes dos participantes [ZANELLI, 2002].

Os estudos observacionais transversais têm por objetivo coletar os dados de forma individual e mensurá-los uma única vez, focando populações bem definidas [Richardson, 2014]. O observador não é apenas um espectador do fato que está sendo estudado; ele se coloca na posição e no nível dos outros elementos que compõem o fenômeno a ser observado, empenhando-se em estudar as aspirações, interesses ou rotina de trabalho de um grupo específico a ser estudado [Richardson, 2014].

Nesta pesquisa, foi observado o impacto que as falhas de sistema operacional causam na qualidade da experiência dos usuários não especialistas (leigos e intermediários) ao utilizar o SO em cenários hipotéticos de uso do computador. Para criar e planejar esse estudo de observação, foi aplicado um questionário online e, a partir dos dados coletados por meio dele, foi possível criar os cenários utilizados e o questionário presencial que foi aplicado durante um estudo de observação em laboratório. As falhas foram simuladas em uma ferramenta de falha que auxiliou na análise de percepção dos participantes.

Todas essas informações serão explicadas de forma detalhada neste capítulo, que está organizado da seguinte forma: a Seção 3.3 explica os materiais utilizados e a Seção 3.2 esclarece quais foram os métodos e abordagens realizadas neste trabalho.

3.2 Materiais

3.2.1 Taxonomia

Para desenvolver esta pesquisa, diversos termos foram utilizados na elaboração das perguntas de pesquisa e na criação dos questionários. Com o propósito de descrever os diversos conceitos utilizados, criou-se um dicionário de termos definidos no âmbito deste trabalho.

- **Contexto:** qualquer informação que pode ser usada para caracterizar a situação da entidade [Abowd, 1999]. Como a definição é ampla, neste trabalho o contexto foi restrito à natureza (profissional e lazer) e o prazo (curto ou longo) para a conclusão da atividade. Exemplo: O chefe

do usuário informou que ele deverá preparar o relatório da reunião que irá acontecer daqui a três horas

- **Consequência:** efeito que um evento de falha provoca no usuário. Exemplo: Devido à ocorrência do evento de falha, o usuário perde as edições que estava realizando no relatório e, conseqüentemente, tem retrabalho.
- **Impacto:** o quanto a experiência do usuário é afetada negativamente pela consequência de uma falha associada ao contexto em que o usuário está inserido. Exemplo: Após a ocorrência do evento de falha, o usuário precisou reescrever o relatório, atrasando a reunião. O usuário ficou insatisfeito e avaliou sua experiência como ruim.
- **Percepção:** é a capacidade do usuário de ver, interpretar e compreender que ocorreu uma falha na operação do computador, porém ele pode ou não possuir conhecimento suficiente para defini-la. Ex.: O usuário precisa imprimir um documento, contudo, ao tentar, ele presencia uma mensagem de falha. À vista disso, ele tem a percepção (percebe) que ocorreu um problema que o impediu de realizar a ação desejada.
- **Reação:** ação do usuário diante de uma falha do sistema operacional. Ex.: Mediante a ocorrência de uma falha, ele desiste de realizar a tarefa que estava fazendo e desliga o computador.
- **Tolerar:** significa até que ponto o usuário suporta ou aceita pacientemente aquela situação antes de gerar uma reação. Ex.: Após a ocorrência da terceira falha consecutiva, o usuário fica insatisfeito e, como reação, ele desliga o computador.
- **Cenário:** caracterizado pelo contexto (natureza/prazo), atividade (interação usuário-computador), evento(s) de falha (ocorrência de uma ou mais falhas) e a consequência gerada pelo evento de falha. Um exemplo de uso desses termos é apresentado a seguir.

“Você está no seu trabalho editando um relatório que precisa ser enviado para seu chefe dentro de duas horas. Durante a edição do documento, de repente você percebe que ocorreu uma falha, pois o sistema operacional travou. Como reação, você reinicia o computador. Essa falha resulta na perda das últimas edições realizadas, ocasionando atraso na entrega do documento e deixando-o extremamente frustrado.”

Nesse cenário, o contexto é profissional com um prazo de duas horas; a consequência é a perda das últimas edições e atraso na entrega do documento; o impacto causado pelo evento de falha (travamento do sistema operacional) na experiência do usuário foi alto (usuário extremamente frustrado). Observa-se que o usuário percebeu que um problema ocorreu e, como reação, ele reiniciou o computador.

3.2.2 Dados de Falha

As falhas selecionadas para a execução do estudo de observação foram:

- Aplicação de SO: explorer.exe (Falha I);
- Serviço de SO: Windows Update (Falha II) e svchost.exe (Falha III);
- kernel de SO: Windows-StartupRepair (Falha IV) e VIDEO_TDR_ERROR (Falha V);
- Aplicação do usuário: notepad.exe (Falha VI).

Nos cenários de falha considerados neste estudo, foram utilizados os cinco principais tipos mais recorrentes de falhas de SO, por categoria, relatados em [Dos Santos, 2016].

Com exceção da falha III, mencionada anteriormente, todas as outras falhas foram descritas no questionário online. As que foram reproduzidas no simulador de falhas foram selecionadas com base no seu impacto na qualidade da experiência do usuário relatada no questionário online. Essa análise levou em consideração as consequências geradas por elas e, também, por outros fatores, como a atividade realizada pelo usuário e o contexto em que ele estava inserido. Como resultado, foi simulada uma falha de cada uma das categorias de falhas: kernel de SO, serviço de SO e aplicação de SO. Desse modo, apenas as falhas I, III, V e VI foram simuladas de forma fidedigna pela ferramenta software.

A falha de aplicação de SO (Falha I) é referente ao programa explorer.exe, responsável pelo gerenciamento de pastas e janelas do Windows. Ela normalmente ocorre quando o usuário tenta acessar algum documento clicando em algum ícone de pasta na área de trabalho do Windows 7, resultando no travamento do explorer.exe.

A falha de serviço de SO reproduzida no simulador de falhas (Falha III) está relacionada com o svchost.exe, responsável pela hospedagem de rotinas específicas de serviços do sistema operacional. Neste estudo, na falha simulada, o svchost.exe estava hospedando uma rotina responsável pela impressão de documentos.

A falha V, VIDEO_TDR_ERROR, foi a falha de kernel de SO selecionada para ser reproduzida no simulador. Ela ocorre durante a tentativa de recarregar um driver de vídeo, ocasionando uma falha do tipo BSOD (Blue Screen of Death – “Tela azul”) no Windows 7 [Dos Santos, 2016]. Essa falha foi selecionada, pois, ao analisar os resultados do questionário online, foi observado que ela estava descrita no cenário de falha que mais causou impacto para o usuário (Seção 4.2).

A fim de avaliar se o impacto causado pela falha na QoE do usuário depende do contexto em que ele se encontra ou da categoria a que a falha pertence, foi necessário utilizar, também, uma falha que se enquadra na categoria de aplicações no nível de usuário, e não do sistema operacional. Por meio dessa avaliação, é possível verificar se as falhas de sistema operacional causam mais impacto comparado com uma falha que não seja causada pelo mau funcionamento do SO. A falha escolhida ocasiona um “travamento” do programa notepad.exe (Bloco de notas). Levando em consideração que a maioria dos usuários utiliza aplicações de edição de texto, essa aplicação foi escolhida por ser um editor de texto mais simples.

3.2.3 Instrumental

Para atingir o objetivo desta pesquisa, foi necessário expor os participantes a diferentes falhas de SO, a fim de observá-los em termos de comportamento. Entretanto, após pesquisa na literatura, não foi encontrada nenhuma ferramenta que atendesse a esse propósito. Dessa forma, em trabalho de pesquisa complementar [Assunção, 2016] e integrado ao projeto de pesquisa maior em que esta pesquisa se insere, foi proposto o desenvolvimento de uma ferramenta de software com o objetivo de coletar os dados relacionados à qualidade da experiência do usuário quanto ao uso do SO.

Essa ferramenta permitiu expor o participante a eventos e consequências de falha predefinidas durante atividade realizada no ambiente de observação (ex.: edição de texto, impressão). Todos os aspectos das falhas simuladas (ex.: modo como a falha aparece, mensagem) foram simulados de forma idêntica às

respectivas falhas reais. Nas 133 sessões de estudo de observação, nenhum participante relatou estar ciente de usar um simulador em vez do software real do sistema operacional.

A ferramenta é composta por quatro programas executáveis (Figura 4), responsáveis pela simulação das falhas estudadas neste trabalho. Ao clicar no executável “Tarefa_01”, este programa prepara o ambiente do sistema operacional, criando os atalhos dos diretórios “Documentos” e também da aplicação “Bloco de notas” na área de trabalho (Figura 5).

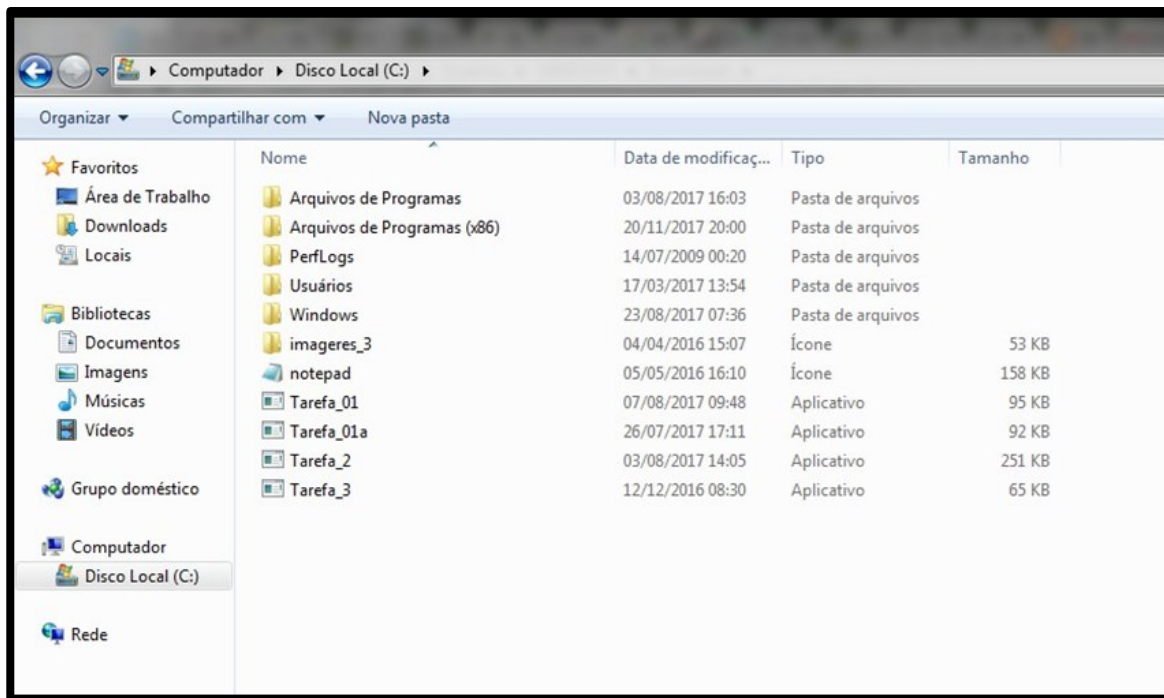


Figura 4 - Programas responsáveis pela simulação das falhas

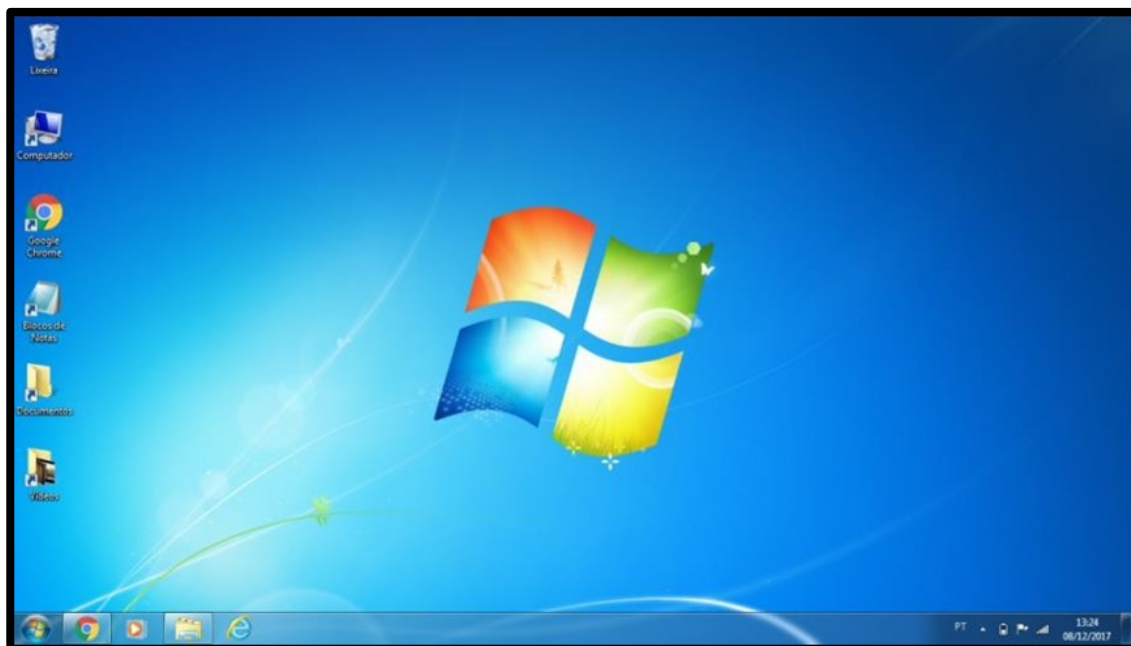


Figura 5 - Ambiente preparado após clicar no executável “Tarefa_01”

O programa “Tarefa_01” é responsável por realizar a simulação da falha relacionada ao programa *explorer.exe*. Primeiramente, esse programa prepara o ambiente do sistema operacional criando o arquivo *batch* denominado “Tarefa_01a”. Esse arquivo é criado no *drive* “C:” e é escrito no arquivo *batch* o comando “*task-kill /f /im explorer.exe*” responsável por encerrar o *explorer.exe* [Assunção, 2016]. Desse modo, ao clicar nesse executável, ele prepara o ambiente criando um atalho desse arquivo (*batch*) que possui a aparência de uma pasta, denominado “Documentos”. Durante a sessão de observação, quando o voluntário clica nesse diretório, o arquivo *batch* chama a janela de mensagem de falha (Figura 6).

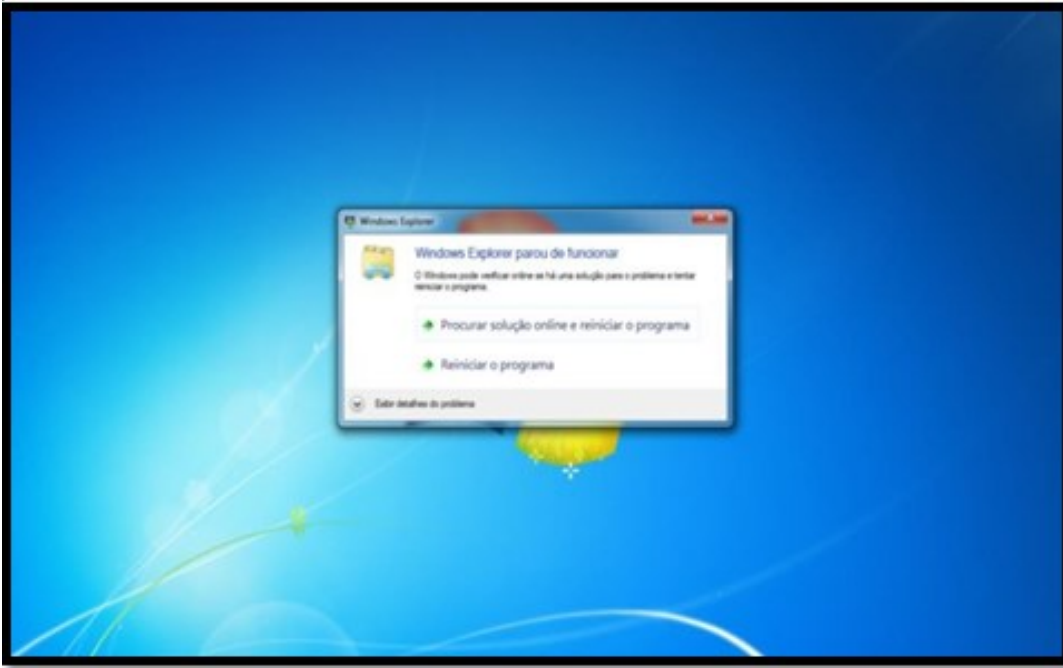


Figura 6 - Mensagem de falha da categoria aplicação de SO apresentada para o participante

O programa “Tarefa_03” é responsável pela reprodução da falha “VIDEO_TDR_ERROR”. Essa falha apresenta uma tela azul (“Blue Screen of Death” - BSOD) durante a reprodução de um vídeo. Ao clicar nesse executável, o programa cria um form invisível, que gera um objeto que procura nos processos executados no sistema operacional o *vmplayer.exe* (Windows Media Player) [Assunção, 2016]. Ao encontrar esse processo, é iniciado um *thread*, que aguarda 20 segundos e, após esse tempo, o vídeo é pausado e o cursor do mouse é ocultado [Assunção, 2016]. Nesse momento, aparece a BSOD (Figura 7).

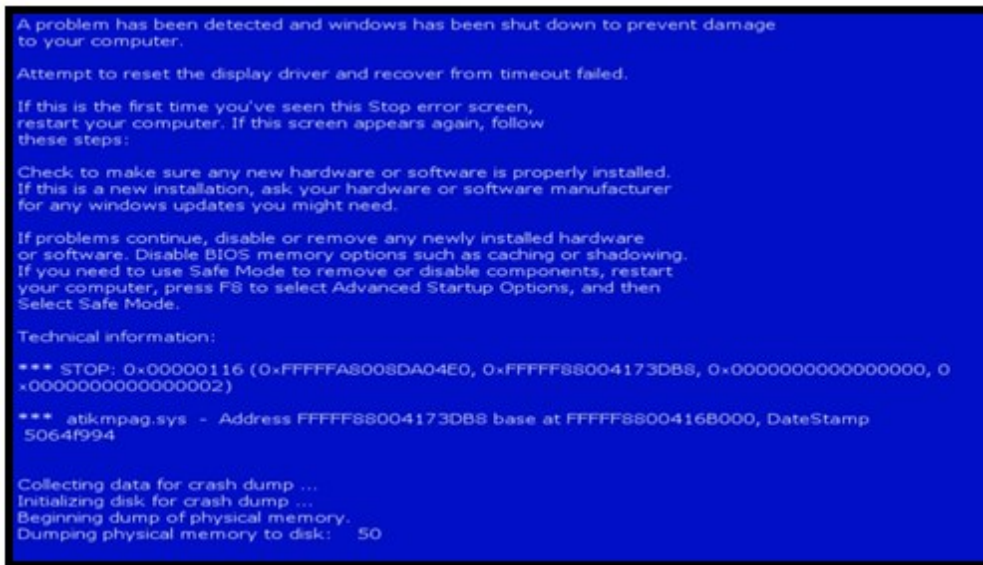


Figura 7 - Tela da falha da categoria kernel de SO

O programa “Tarefa_02” é responsável pela falha da categoria nível de usuário (notepad.exe) e serviço do SO (svchost.exe). Para simular essas falhas, foi necessário implementar um programa que fosse idêntico ao notepad.exe nativo do Microsoft Windows 7 (Figura 8). Ao clicar no programa (Tarefa_02), ele executa o aplicativo notepad.exe desenvolvido para atender ao propósito desta pesquisa [Assunção, 2016]. Por fim, é criado manualmente um atalho dessa aplicação na área de trabalho.

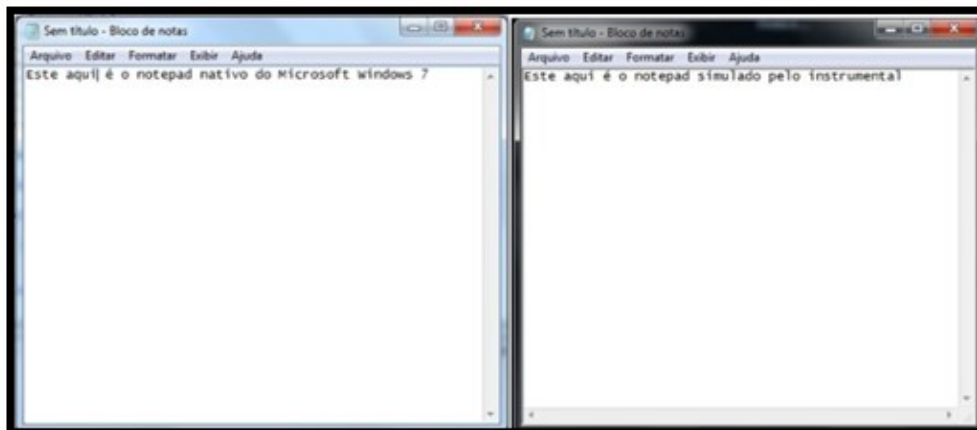


Figura 8 - Tela do bloco de notas (notepad.exe) original e o bloco de notas desenvolvido para realizar o estudo de observação

Fonte: [Assunção, 2016]

Durante o estudo de observação, o participante precisa digitar um texto e depois salvá-lo e imprimi-lo. Ao finalizar a edição do texto, quando o participante tenta salvar o documento, a falha é apresentada. A tela fica esmaecida e, logo em seguida, aparece uma janela com mensagem indicando a falha relacionada do programa bloco de notas (Figura 9).

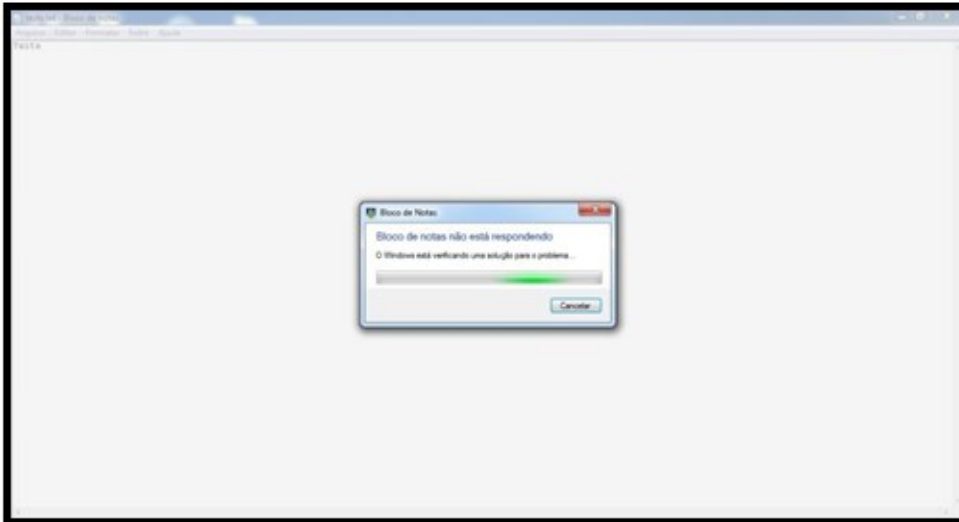


Figura 9 - Tela apresentada no momento em que o documento digitado no bloco de notas é salvo

Além disso, seguindo o cenário de falha apresentado para o usuário, ele precisa imprimir um arquivo. Ao tentar imprimir, ele depara com uma mensagem de falha que se encaixa na categoria serviços de SO. Para implementar essa falha, foi utilizado um componente de impressão do C# (Figura 10) e o botão “ok” foi programado para chamar uma janela de falha em vez de chamar o método que faz a impressão. Desse modo, no momento em que o usuário clica em “ok”, a mensagem aparece (Figura 11).



Figura 10 - Tela de impressão do Windows

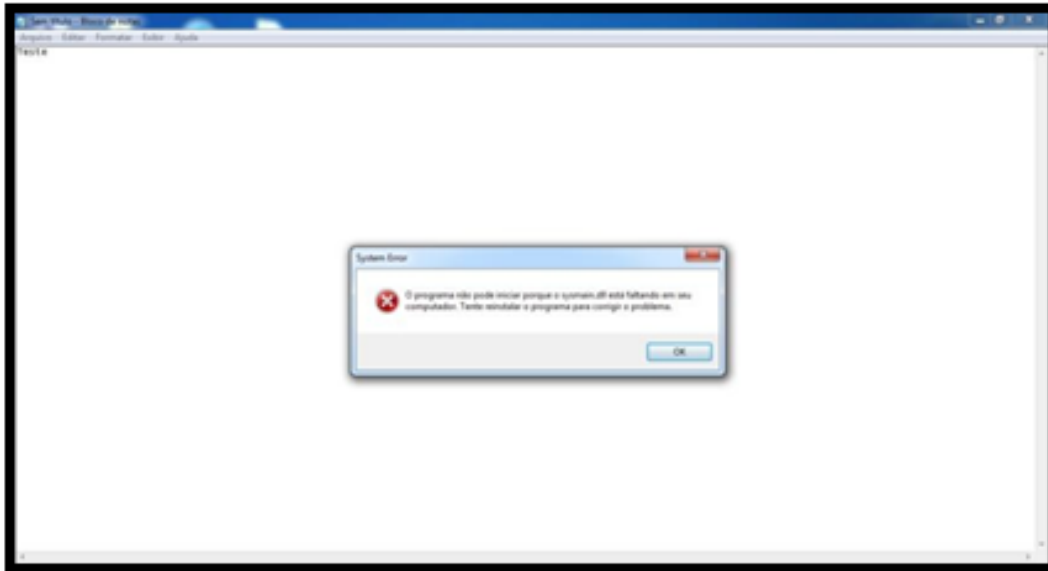


Figura 11 - Mensagem de falha apresentada para o usuário no momento em que ele clica em “ok” (Figura 10) para imprimir o documento

3.2.4 Questionário utilizados para avaliar a experiência do usuário

Para avaliar a experiência do usuário, foram utilizados dois questionários. Primeiramente, foi aplicado um questionário online, que teve como intuito coletar dados para planejamento do ambiente de estudo desta pesquisa, onde foi realizado um estudo de observação com o propósito de analisar e compreender a influência que as falhas de sistema operacional causam na experiência do usuário ao utilizar o computador.

O questionário online foi criado por meio da ferramenta LimeSurvey, software livre utilizado para auxiliar pesquisas online, pois apresenta 28 opções de perguntas, templates, gráficos, entre outros recursos, os quais ajudam o pesquisador a gerenciar e a criar seus questionários [LimeSurvey, 2017].

O questionário apresentou primeiramente 15 questões de múltipla escolha cujo objetivo era caracterizar os participantes, interrogando-os quanto ao seu gênero, área de atuação/profissão, importância do computador para o seu dia a dia e ao nível de conhecimento com respeito à área de tecnologia da informação (TI) e nível de conhecimento em SO. Por meio desses questionamentos, foi possível identificar os usuários do grupo G1 (leigos) e grupo G2 (intermediários).

Levando em consideração a definição de contexto e de cenário utilizada nesta pesquisa, o questionário online apresentou oito cenários de falha que os respondentes assumiram estar vivenciando. Esses cenários foram divididos em dois conjuntos (Set 1 e Set 2). No primeiro conjunto (Set 1), apresentaram-se quatro cenários com diferentes contextos, atividades, eventos de falha e consequências, enquanto no segundo (Set 2), optou-se por manter o mesmo contexto e atividade nos quatro cenários propostos, porém diversificando os fatores restantes (evento de falha e consequência). O contexto e a atividade realizada pelos usuários foram mantidos nesse grupo de cenários objetivando analisar o quanto a consequência e o evento de falha afetam a experiência do usuário. As Figuras 12 e 13 apresentam a descrição dos cenários I do Set 1 e do Set 2, respectivamente. As Tabelas 1 e 2 descrevem os cenários utilizados em cada conjunto.

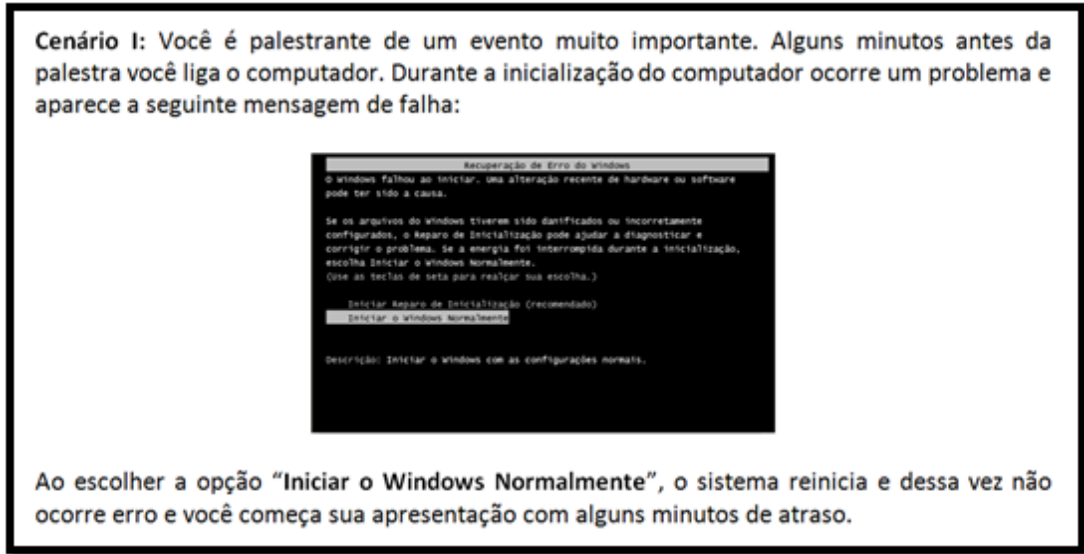


Figura 12 - Cenário I (Set 1) do questionário online (falha da categoria de falhas de kernel)

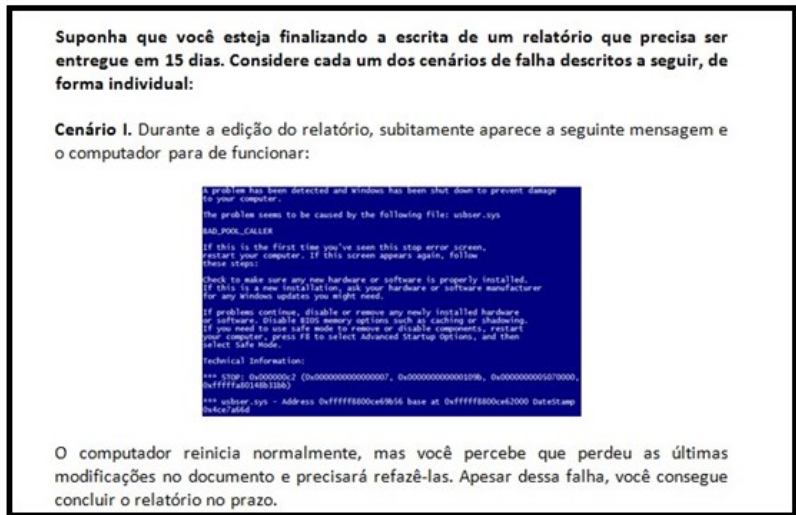


Figura 13 - Cenário II (Set 2) do questionário online (falha da categoria de falhas de kernel)

Tabela 1 - Características dos cenários descritos no Set 1

CENÁRIO I
<p>Contexto: profissional – prazo curto (Alguns minutos antes de a palestra iniciar).</p> <p>Atividade: realizando uma palestra em um evento profissional.</p> <p>Evento de falha: <i>Windows-StartupRepair</i> (falhas de <i>kernel</i> de SO).</p> <p>Consequência: iniciou a palestra com alguns minutos de atraso.</p>
CENÁRIO II
<p>Contexto: profissional – prazo curto (2 horas).</p> <p>Atividade: criando um relatório.</p> <p>Evento de falha: travamento da aplicação <i>notepad.exe</i> (falha de aplicação de usuário).</p> <p>Consequência: retrabalho e atraso no envio do relatório.</p>

CENÁRIO III
<p>Contexto: lazer – prazo longo (1 dia de folga).</p> <p>Atividade: assistindo a um filme.</p> <p>Evento de falha: <i>Windows Update</i> (falhas de serviço de SO).</p> <p>Consequência: não terminou de assistir ao filme, pois o áudio parou de funcionar.</p>
CENÁRIO IV
<p>Contexto: Profissional - prazo curto (1 hora).</p> <p>Atividade: Procurando arquivos.</p> <p>Evento de falha: <i>explorer.exe</i> (falhas de aplicação de SO).</p> <p>Consequência: enviou o arquivo com atraso.</p>

Tabela 2 - Características dos cenários descritos no Set 2

<p>Contexto único para todos os cenários: profissional – prazo longo (15 dias). Atividade única para todos os cenários: criando um relatório.</p>
CENÁRIO I
<p>Evento de falha: <i>VIDEO_TDR_ERROR</i> (falhas de <i>kernel</i> de SO).</p> <p>Consequência: perdeu todas as edições realizadas no documento, resultando em retrabalho. Contudo o relatório foi enviado dentro do prazo estipulado.</p>
CENÁRIO II
<p>Evento de falha: travamento da aplicação <i>notepad.exe</i> (falha de aplicação de usuário).</p> <p>Consequência: perdeu todas as edições realizadas no documento, resultando em retrabalho. Contudo o relatório foi enviado dentro do prazo estipulado.</p>
CENÁRIO III
<p>Evento de falha: <i>Windows Update</i> (falhas de serviço de SO).</p> <p>Consequência: o editor de texto para de funcionar, ocasionando a perda das últimas modificações. Entretanto o relatório também foi enviado dentro do prazo estipulado.</p>
CENÁRIO IV
<p>Evento de falha: <i>explorer.exe</i> (falhas de aplicação de SO); durante a edição, subitamente, a mensagem de falha aparece.</p> <p>Consequência: Não perde nenhuma edição realizada no documento, e o arquivo foi enviado dentro do prazo estipulado.</p>

Para cada conjunto, o participante respondeu a duas questões (uma de múltipla escolha e uma pergunta aberta) sobre o quanto sua experiência seria negativamente afetada, caso estivesse vivenciando esses cenários em seu cotidiano.

A vantagem de aplicar um questionário online é o menor custo e tempo para coletar os dados necessários de um número maior de respondentes. No entanto, de acordo com [Vieira, 2000], um dos principais problemas ao utilizar essa abordagem é a falta de garantia quanto ao entendimento dos respondentes às perguntas. Além disso, somente com a aplicação do questionário online não seria possível avaliar a tolerância e a percepção dos usuários ao presenciarem diferentes tipos de falha. Sendo assim, a fim de mitigar esse viés, foi aplicado um questionário presencial em um ambiente em laboratório durante o estudo de observação, objetivando avaliar as percepções e as reações dos usuários ao presenciarem determinada falha.

Primeiramente, foram extraídas informações dos resultados do questionário online para criar quatro cenários que foram utilizados durante o estudo. Com base nesses resultados (Seção 5), foi definido que as atividades realizadas pelos participantes do questionário presencial seriam de natureza profissional ou lazer. Algumas atividades narradas nos cenários descritos no questionário online não apresentaram prazos definidos com um período exato, apenas especificaram que essas atividades, em dado contexto, precisavam ser finalizadas com urgência. Para evitar que os participantes do questionário presencial ficassem confusos quanto à urgência de finalização das atividades que iriam realizar, observou-se a necessidade de especificar um prazo curto (2-3 horas) e um prazo longo (1 semana) para conclusão dessas atividades. A Tabela 3 apresenta as descrições dos contextos utilizados nos cenários propostos para as sessões de observação.

Tabela 3 - Descrição dos contextos descritos nos cenários do

CONTEXTO DO CENÁRIO I
Natureza da atividade: profissional Prazo: 7 dias (prazo longo)
CONTEXTO DO CENÁRIO II
Natureza da atividade: lazer Prazo: 3 horas (prazo curto)
CONTEXTO DO CENÁRIO III
Natureza da atividade: lazer Prazo: 7 dias (prazo longo)
CONTEXTO DO CENÁRIO IV
Natureza da atividade: profissional Prazo: 3 horas (prazo curto)

A consequência gerada pela falha foi um dos principais fatores que influenciaram o alto impacto negativo que os eventos de falha descritos nos cenários do questionário online causaram na experiência do

usuário (UX). Portanto, para não influenciar a escolha dos participantes quanto ao nível de impacto causado por determinada falha em sua experiência, a consequência gerada por ela não foi descrita de forma explícita nos cenários utilizados no questionário presencial. Todos os fatores que influenciaram o nível de impacto causado pelas falhas simuladas na ferramenta de simulação de falhas foram percebidos e informados pelos próprios participantes ao presenciarem essas falhas.

Cada participante foi envolvido em apenas um cenário e, após experimentar determinado evento de falha, era necessário responder a treze perguntas de múltipla escolha sobre o quanto sua experiência havia sido negativamente afetada. Em algumas perguntas, era solicitada uma justificativa para sua resposta. Além dessas perguntas, no final do questionário presencial foi solicitado aos participantes que respondessem a três questões relacionadas à qualidade geral da experiência em relação à confiabilidade do sistema operacional (Apêndice A).

3.2.5 Participantes da pesquisa

Os participantes (usuários de computadores) foram organizados em dois grupos (leigos e intermediários). O primeiro grupo (G1) engloba aqueles que possuem habilidades computacionais muito básicas principalmente relacionadas a sistemas operacionais. Por exemplo, esses participantes não conseguem diferenciar as aplicações que são do sistema operacional daquelas do usuário. Eles são capazes de realizar instalações simples de programas e utilizar as aplicações instaladas no sistema, porém não conhecem o sistema operacional detalhadamente.

No segundo grupo (G2), intermediários, estão aqueles participantes que se encaixam no nível médio de conhecimento em informática, sendo mais habilidosos no uso do computador, sabendo diferenciar o que faz ou não parte do sistema operacional. Eles conseguem informar que o Microsoft Word é uma aplicação de usuário, bem como que um gerenciador de impressão é uma aplicação do sistema operacional. Além disso, eles sabem definir o que é um sistema operacional e entendem seu objetivo e importância, contudo não conhecem suas especificidades. Esses participantes não são especialistas na área e não têm conhecimento sobre componentes internos do SO. Os usuários avançados foram desconsiderados, uma vez que, no estágio atual da pesquisa, não se encaixam no propósito do trabalho.

Para categorizá-los, nos dois questionários foram incluídas perguntas com o intuito de avaliar o nível de conhecimento dos participantes em SO. Essas questões foram criadas baseando-se no conceito definido para cada um dos grupos estudados nesta pesquisa, como conhecimento referente às aplicações que fazem parte ou não do sistema operacional. O capítulo 5 demonstra detalhadamente as questões utilizadas para essa finalidade.

Como se desconhece o tamanho da população, nesse caso composta de todos os usuários de computadores, foi adotada a abordagem proposta em [Barbetta, 2008], a qual é baseada na Equação 1.

$$n = 1 \div E^2, \quad (1)$$

em que n é o tamanho da amostra e E é a margem do erro amostral tolerada.

Face à dificuldade inerente de obter voluntários para participar da pesquisa, trabalhou-se com uma margem de erro de 15%, resultando em um mínimo de 45 participantes para cada grupo estudado (leigos e intermediários).

Na primeira etapa, questionário online, 150 respondentes participaram da pesquisa. Entretanto, após realizar a análise dos dados, foi possível observar que alguns respondentes haviam colocado as mesmas respostas para todas as questões ou não haviam colocado informações precisas nas justificativas, como área de estudo e motivo que os levou a escolher um cenário de falha como aquele que causou maior impacto negativo em sua experiência. Sendo assim, foram descartados 29 questionários (Seção 4.2), obtendo-se uma amostra final de 121 respondentes.

Para categorizá-los em usuários leigos ou intermediários, o questionário online apresentou 15 questões iniciais referentes ao seu nível de conhecimento em informática e sistema operacional. Ao finalizar a coleta dos dados, as respostas foram analisadas e cada respondente foi inserido no grupo referente a seu nível de conhecimento e, assim, obteve-se uma amostra com 61 usuários leigos (G1) e 60 intermediários (G2) (Tabela 4), isto é, a categorização dos usuários foi realizada após o final da coleta dos dados.

Tabela 4 - Número de participantes da etapa 1 – questionário online

GRUPO	LEIGO (G1)	INTERMEDIÁRIO (G2)	TOTAL
Nº	61	60	121
%	48,330%	51,670%	100%

Na segunda etapa, questionário presencial, 133 participantes responderam ao questionário. Entretanto, após a análise dos dados, 13 foram descartados (Seção 4.3), totalizando uma amostra final de 120 participantes. Como o questionário foi aplicado em um estudo de observação de período exato de uso do computador, os questionários daqueles participantes que se retiraram do laboratório para atender ao telefone ou realizar qualquer outra tarefa em paralelo foram descartados.

Com o intuito de categorizar os participantes, primeiramente foi aplicado um questionário com 10 questões referentes a seus níveis de conhecimento em informática e SO (Figura 14). As respostas coletadas foram analisadas antes de se realizar o sorteio referente ao contexto (profissional com prazo, lazer com prazo, profissional sem prazo, lazer sem prazo) que eles suporiam estar vivenciando.

9. O que significa software?
 Pode ser definido como os programas que comandam o funcionamento do computador. É classificado pela parte lógica do computador.
 Parte física de um computador.
 Material em estado físico necessário para fazer com que o computador funcione.
 Não sei responder essa questão.

10. Informe quais dos programas abaixo são aplicações do sistema operacional (SO), ou seja, programas que são instalados com o SO para realizarem alguma função de responsabilidade do SO.

- Gerenciador de impressão.
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- Editor de Texto.
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- Calculadora.
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- Interpretador de linha de comando (Ex.: *Prompt* de comando).
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- Tocador (player) de áudio e vídeo.
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- Editor Gráfico (Editor de imagem).
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei

Figura 14 - Exemplo de questões utilizadas para categorização dos usuários

Diferentemente do questionário online, no estudo de observação os participantes foram categorizados antes de darem início ao estudo, obtendo-se uma amostra com 60 usuários leigos e 60 intermediários (Tabela 5).

Tabela 5 - Número de participantes da etapa 2 – questionário presencial

GRUPOS	LEIGOS (G1)	INTERMEDIÁRIOS (G2)	TOTAL
Nº	60	60	120
%	50%	50%	100%

Sendo assim, após os descartes de 42 questionários, resultou uma amostra final de 241 participantes desta pesquisa.

3.3 Método

3.3.1 Protocolo de pesquisa

Não foi encontrado nenhum método que atendesse ao propósito deste trabalho, portanto foi criado um protocolo próprio para avaliar a experiência do usuário quanto à sua confiança no SO. Nele, o estudo foi dividido em duas etapas, uma online e outra presencial. A etapa online teve por objetivo principal coletar dados para planejamento e criação do estudo de observação. A Figura 15 apresenta o fluxo de trabalho realizado para dar início à aplicação do questionário online, demonstrando os passos seguidos para sua criação e aplicação.

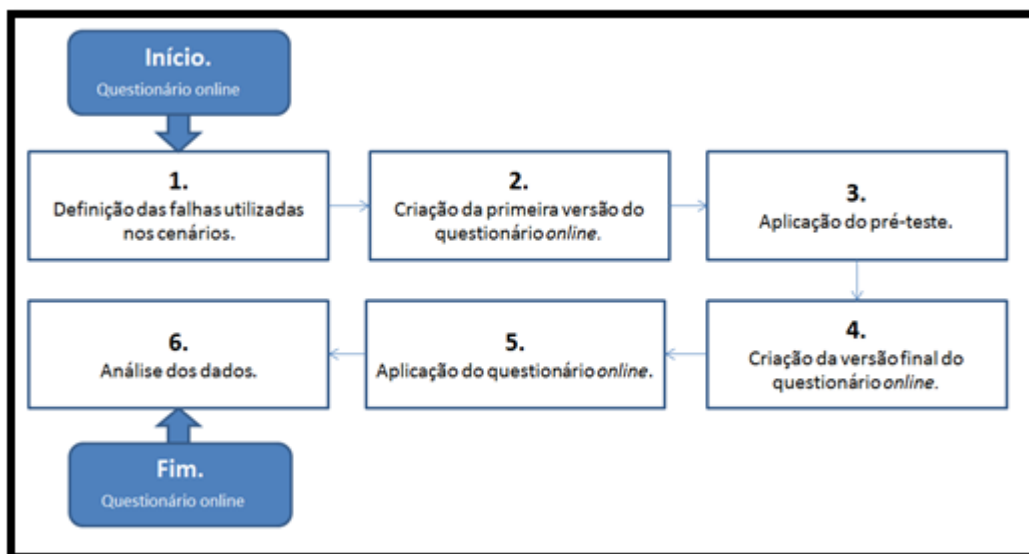


Figura 15 - Fluxo de trabalho da aplicação do questionário online

Após selecionar as falhas mais frequentes em [Dos Santos, 2016] descritas na Seção 3.2.2, o questionário online foi criado (Seção 3.2.4). Primeiramente, foi realizado um pré-teste com o intuito de verificar se as perguntas narradas no questionário atenderiam ao propósito desta pesquisa. No final do teste, a versão final do questionário online foi criada. Sendo assim, deu-se início à aplicação do questionário online para 121 participantes (Seção 3.2.5).

Ao serem finalizadas as análises dos dados coletados nesse questionário, para que se desse início ao estudo de observação (Figura 16), como mencionado na Seção 3.2.4, o questionário presencial foi criado baseando-se nos resultados do questionário online e aplicado para 120 participantes (Seção 3.2.5). Esses participantes foram, primeiramente, categorizados em usuários leigos e intermediários por meio da aplicação de 10 questões que avaliaram seus níveis de conhecimento em informática e SO. Posteriormente, eles foram sorteados aleatoriamente e encaixados em um dos quatro cenários de falha utilizados no estudo (ver Seção 3.2.4). Assim como foi realizado no questionário online, para criar perguntas que atendam ao objetivo desta pesquisa, também foi realizado um pré-teste do questionário presencial. Ao coletar os feedbacks dos participantes que fizeram parte desse teste, a versão final do questionário foi criada, dando início ao estudo. O questionário presencial foi aplicado em um estudo de observação em laboratório e, durante as sessões de avaliação, os participantes utilizaram a ferramenta de simulação de falhas. Assim, foram coletadas as percepções e reações dos participantes ao presenciarem determinado evento de falha.

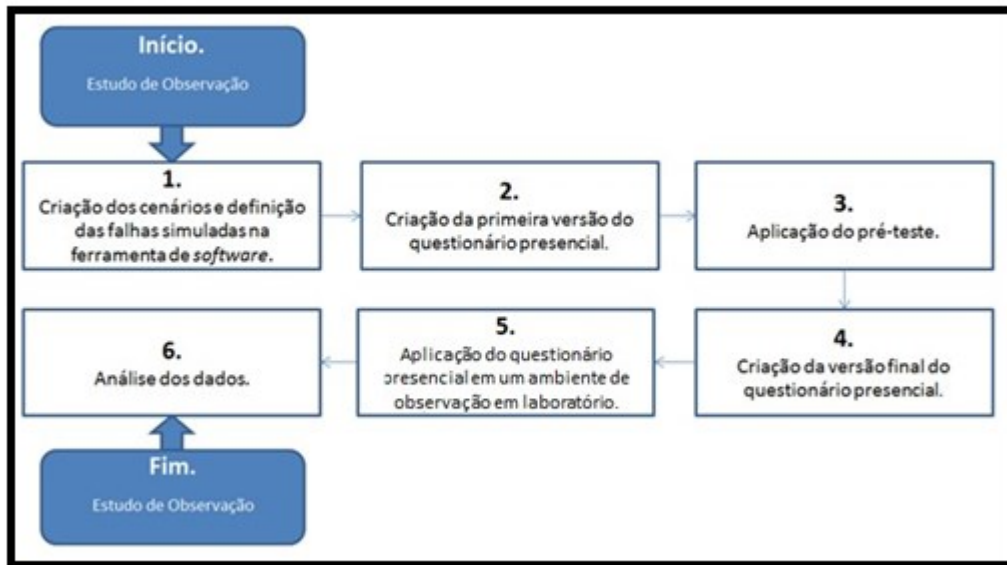


Figura 16 - Fluxo de trabalho da aplicação do questionário presencial

Após a coleta dos dados, eles foram analisados objetivando responder aos questionamentos desta pesquisa. O Capítulo 4 explica detalhadamente como o estudo foi conduzido.

3.3.2 Plano de recrutamento

Planejou-se recrutar participantes de diferentes gêneros, idades, ocupações, afiliações e cidades. Para atingir esse propósito, foi enviado um convite por e-mail com um link que os direcionava para o questionário online. O convite foi compartilhado na lista de contatos da rede de relacionamento da autora desta pesquisa (ex.: colegas de trabalho, professores da graduação e pós-graduação). Além disso, o convite foi compartilhado também na lista de contatos da Faculdade de Computação (FACOM) da Universidade Federal de Uberlândia, bem como da Sociedade Brasileira da Computação (SBC).

Após a finalização dessa etapa, foi enviado um convite por e-mail para departamentos de empresas e universidades, convidando funcionários e estudantes a responderem ao questionário presencial. O convite foi colocado também em locais frequentados pela pesquisadora, atraindo indivíduos que gostariam de se voluntariar a participar do estudo de observação.

A Escola Estadual Horizonta Lemos – Uberaba/MG, o Instituto Federal do Triângulo Mineiro (IFTM) – Uberaba/MG e a empresa Fretebras – Catalão/GO autorizaram a coleta de dados com os frequentadores desses locais, sendo eles estudantes e funcionários. Foram agendados os horários, respeitando a disponibilidade de tempo e localização de cada um (Tabela 6). Para aqueles que não podiam se locomover até a Universidade Federal de Uberlândia (UFU) – campus Santa Mônica, os pesquisadores responsáveis se deslocaram até o local de preferência do participante.

Tabela 6 - Exemplo de agendamento de horários e localização respeitando a

DATA	DIA	PERÍODO	Nº	LOCAL	CIDADE
28/ago	2ª-feira	Matutino	3	Biblioteca Municipal	Uberaba/MG
29/ago	3ª-feira	Noturno	6	Associação Caridade	Uberaba/MG
25/set	2ª-feira	Vespertino	6	Visitação residência	Catalão/GO
26/set	3ª-feira	Matutino Vespertino	4	Visitação residência	Catalão/GO
27/set	4ª-feira	Vespertino	6	UFU	Uberlândia/MG

Como este trabalho norteou seus procedimentos dentro de uma coleta de dados não invasiva com seres humanos, de acordo com a [Resolução 466, 2012], o projeto foi submetido à análise ética de um CEP (Comitê de Ética em Pesquisas com Seres Humanos) e somente se iniciou após sua aprovação. Todos os documentos necessários relacionados a este trabalho foram submetidos ao CEP da Universidade Federal de Uberlândia. Os dados foram coletados respeitando-se o parecer emitido pelo CEP (Anexo A). Além disso, vale ressaltar que a participação no questionário online foi anônima e o link de recrutamento foi enviado para diferentes grupos de pessoas de diferentes estados brasileiros. Sendo assim, a fim de aplicar o questionário presencial para pessoas que não haviam respondido ao questionário online, o convite de recrutamento foi enviado para empresas diferentes daquelas que participaram do questionário online, e o ambiente do estudo de observação foi criado em diferentes locais e cidades.

3.3.3 Análise dos dados

A Figura 12 Como todas as variáveis estudadas nesta pesquisa são categóricas, elas foram divididas em nominais e ordinais. As variáveis nominais são aquelas que não apresentam uma ordem de valores, como sim ou não, certo ou errado, gênero masculino ou feminino. As variáveis ordinais apresentam uma ordenação entre si, como nível de satisfação (ex.: baixo, médio ou alto). Desse modo, as variáveis estudadas nesta pesquisa foram classificadas da seguinte forma:

- Variáveis dicotômicas: a falha de operação do computador é ou não uma consequência de uma falha de SO; o sistema está ou não operando corretamente; de modo geral, confia ou não no SO. As respostas dessas questões entram nas variáveis sim ou não e, também, de acertos e erros.
- Variáveis ordinais: níveis de tolerância e confiança (nem um pouco, pouco, mais ou menos, muito, muitíssimo), níveis de avaliação da QoE e QoS (muito ruim, ruim, regular, bom,

muito bom, excelente). As respostas a essas questões entram na ordem de baixo, médio ou alto nível de tolerância, confiança ou satisfação.

Os dados coletados em ambos os questionários foram organizados em tabelas de contingência para comparar a distribuição da resposta variável (nível de impacto, nível de tolerância) entre os grupos de comparação (grupos de usuários, contextos).

Após realizar essa organização dos dados, aplicaram-se testes estatísticos não paramétricos visando a avaliar a significância dos dados encontrados. Esses testes são aplicados quando as variáveis estudadas são do tipo nominal ou ordinal [Lyman, 2010]. Ainda segundo esse autor, os testes não paramétricos são divididos de acordo com o nível de mensuração e número de grupos que se pretende estudar. Desse modo, neste trabalho foi utilizado o teste exato de Fisher [Lyman, 2010] e também o teste Qui-quadrado [Lyman, 2010].

Aplicou-se o teste Qui-quadrado de independência, indicado por χ^2 , que é um teste de hipótese utilizado para avaliar a associação entre as variáveis com dois ou mais grupos independentes. Esse teste realiza uma comparação, analisando se a frequência observada dessas variáveis possui uma diferença significativa ou se elas são distribuídas igualmente entre os grupos estudados. Para realizar esse teste, primeiramente é necessário calcular o χ^2 levando em consideração a frequência observada de determinada variável coletada a partir da amostra de dados. Após realizar esse cálculo, o valor encontrado é comparado com o χ^2 tabelado. Esse valor é encontrado a partir do grau de liberdade e nível de confiança adotado no teste de hipóteses [Lyman, 2010].

Quando o χ^2 calculado for maior ou igual ao χ^2 tabelado, rejeita-se a H_0 , então, não há associação entre variáveis. Quando o χ^2 calculado é menor que o χ^2 tabelado, a hipótese nula não é rejeitada e, assim, conclui-se que as variáveis estão associadas [Lyman, 2010].

Esse teste, entretanto, não deve ser utilizado em amostras pequenas (amostra menor que 40), quando um dos valores esperados é menor do que 5 ou se alguma delas for igual a 0 [Lyman, 2010]. Desse modo, para amostras menores que 40, utiliza-se o teste exato de Fisher. Esse teste também avalia se existe associação entre as variáveis em estudo. O teste exato de Fisher é interpretado avaliando se o p-valor calculado é menor ou igual ao nível de confiança estipulado. Quando esse valor é maior, aceita-se a hipótese nula e, quando ele é menor, rejeita-se a hipótese nula [Lyman, 2010].

Todas as suposições do teste χ^2 foram verificadas anteriormente para cada tabela de contingência. Para o questionário online, o χ^2 foi aplicado em todos os seis testes. Por outro lado, para o questionário presencial, oito dos doze testes planejados não atenderam aos pressupostos necessários para o teste do χ^2 . Isso aconteceu porque o número de frequências esperadas foi menor que cinco em mais de 20% das células de resposta (combinação de uma linha e coluna) em suas tabelas de contingência [Lyman, 2010]. Para esses oito casos, aplicamos o teste exato de independência de Fisher [Lyman, 2010]. Para todos os testes estatísticos deste estudo, adotou-se um nível de significância de 5% ($\alpha = 0,050$).

Todos os testes foram aplicados utilizando um software livre integrado à linguagem de desenvolvimento R denominado RStudio. Essa linguagem auxilia pesquisadores na aplicação de testes estatísticos bem como na criação de gráficos e tabelas [Studio, 2018]. Por meio dessa ferramenta, foi

possível realizar esses cálculos utilizando as funções `qchisq.test` (Teste Qui-Quadrado) e `fisher.test` (Teste exato de Fisher).

Planejamento e criação do estudo de observação

4.1 Introdução

Este capítulo descreve como os questionários foram criados e aplicados. A Seção 4.2 apresenta o objetivo do questionário online, os resultados do pré-teste, bem como um resumo da análise descritiva dos dados. Por fim, a Seção 4.3 apresenta todo o processo de realização do estudo de observação, explicando o objetivo desse estudo e um resumo da análise descritiva dos resultados encontrados.

4.2 Questionário *online*

Para criar o questionário online, primeiramente foi realizado o mapeamento das falhas mais frequentes em [Dos Santos, 2016], como pode ser visto na Seção 3.2.1. A primeira versão do questionário foi dividida em duas partes. A primeira apresentou 15 questões de múltipla escolha para a criação do perfil do respondente; a segunda parte apresentou quatro cenários de falha com seis questionamentos cada um (Figura 17).

Para verificar se as informações coletadas por meio desse questionário estavam de acordo com o propósito desta pesquisa, foi realizado um pré-teste. Um convite de participação foi enviado para 15 voluntários aleatórios da lista de contatos de e-mail da autora desta pesquisa. Desses voluntários, 12 participantes responderam ao questionário completamente, porém um precisou ser descartado, resultando em uma amostra final com 11 participantes.



Figura 17 - Exemplo de cenário de falha apresentado na primeira versão do questionário presencial

Para verificar se as informações coletadas por meio desse questionário estavam de acordo com o propósito desta pesquisa, foi realizado um pré-teste. Um convite de participação foi enviado para 15 voluntários aleatórios da lista de contatos de e-mail da autora desta pesquisa. Desses voluntários, 12 participantes responderam ao questionário completamente, porém um precisou ser descartado, resultando em uma amostra final com 11 participantes.

Com relação às questões de avaliação do conhecimento do usuário na área de TI (Figura 18 e 19), as perguntas que apresentaram as opções “concordo”, “discordo” e “não sei responder a essa questão” causaram dúvidas na hora de responder ao questionário. Desse modo, os participantes sugeriram alterar essas opções para “sim”, “não” e “não sei responder a essa questão”. Segundo os participantes, essas respostas são mais objetivas e não causam confusão no momento em que os participantes fossem responder ao questionário online oficial.

11. Se tratando de programas de computadores, você:

PROGRAMAS	Concordo	Discordo	Não sei responder essa questão
Sabe instalar programas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sabe desinstalar programas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sabe utilizar os programas que são instalados juntamente com o sistema operacional (Windows Explorer, Finder).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 18 - Exemplo de pergunta apresentada na primeira etapa do questionário online

Os participantes reclamaram que as perguntas relacionadas aos cenários de falha apresentaram muitas opções de resposta (Figura 17), causando dúvidas no momento de responder a elas. Além disso, eles acharam que os cenários de falha estavam repetitivos, tornando o questionário cansativo.

Observando os feedbacks apresentados pelos participantes, as questões para enquadrá-los no grupo de usuários leigos ou intermediários também causaram confusão. Apesar de eles terem ficado com dúvidas em algumas questões, foi possível atingir o propósito do questionário com os dados coletados. Desse modo, para evitar dúvidas no momento em que os participantes fossem responder a eles, algumas

perguntas foram alteradas. As opções apresentadas nas questões de múltipla escolha foram modificadas, respeitando as sugestões apresentadas pelos participantes (Figura 18).

11. Se tratando de programas de computador:

PROGRAMAS	Sim	Não
Você sabe instalar programas?	<input type="radio"/>	<input type="radio"/>
Você sabe desinstalar programas?	<input type="radio"/>	<input type="radio"/>
Você sabe utilizar alguns dos programas que são instalados juntamente com o sistema operacional (ex. Windows Explorer, Finder)?	<input type="radio"/>	<input type="radio"/>

Figura 19 - Exemplo de pergunta para avaliar o conhecimento do usuário em TI (após as alterações sugeridas no pré-teste)

A questão para verificar em qual grupo os participantes se encaixavam, leigo ou intermediário, foi alterada de modo que as opções de respostas não fossem específicas de nenhum software do sistema operacional estudado. As opções foram descritas de modo que os participantes que estavam familiarizados com diferentes sistemas operacionais de desktop (ex.: Linux, Windows, macOS) conseguissem responder a elas (Figura 20).

8. Informe quais dos programas abaixo são aplicações do sistema operacional (SO), ou seja, programas que são instalados com o SO para realizarem alguma função de responsabilidade do SO.

- Gerenciador de impressão
- Editor de Texto
- Calculadora
- Interpretador de linha de comando
- Tocador (player) de áudio e vídeo
- Navegador de Internet
- Formatador de disco
- Leitor de e-mail

Figura 20 - Exemplo de pergunta apresentada no questionário presencial e online para caracterização do perfil dos participantes

Além dessas alterações, para evitar que o questionário ficasse repetitivo e cansativo, as descrições dos cenários de falha também foram modificadas. Os cenários foram divididos em dois conjuntos. O primeiro apresentou quatro cenários descrevendo diferentes contextos, consequências e eventos de falha, enquanto o segundo descreveu quatro cenários que ocorreram no mesmo contexto, variando apenas os demais fatores. Após ler esses cenários de falha, o participante precisou definir quais foram aqueles que causariam mais insatisfação, bem como quais foram os que causariam menor impacto em sua experiência, caso estivessem vivenciando essas situações hipotéticas (ver Seção 3.2.4).

Após realizar essas alterações, a versão final do questionário online foi concluída e aplicada para 150 participantes. Contudo, ao analisar os resultados, 29 questionários foram descartados. Os dados foram analisados levando em consideração as variáveis qualitativas, isto é, gênero, faixa etária, escolaridade,

sistema operacional utilizado, grupo de usuários e cenários de falha que causaram maior impacto. Posteriormente, foi realizada uma análise descritiva dos dados para se conhecer a frequência absoluta e a relativa de todas as variáveis estudadas. Além disso, por fim, foi realizada uma análise estatística objetivando responder aos questionamentos desta pesquisa. Os questionários descartados não foram incluídos na amostra final, devido a diferentes fatores. O primeiro fator que levou à exclusão do questionário foi a área de estudo que não foi identificada pelo participante. Ao analisar essas respostas, não foi possível definir em qual área de estudo o participante se encaixava, portanto eles foram descartados (Tabela 7).

Tabela 7 - Amostra de respostas que os usuários colocaram na área de estudo e profissão

ID USUÁRIOS	DOCENTE DA EDUCAÇÃO BÁSICA
User202	Professora
User76	Professora
User103	Estudante
User99	Professor

Alguns participantes informaram que, naquele momento, não utilizavam nenhum sistema operacional de desktop. Como o objetivo deste trabalho é avaliar a experiência do usuário diante das diferentes falhas do sistema operacional de computadores desktop, não seria possível coletar real satisfação do usuário sem ele conseguir se imaginar naquela situação, já que, segundo o próprio participante, ele não utiliza nenhum SO de computador no momento. Além disso, aqueles participantes que colocaram a primeira resposta em todas as perguntas também foram descartados, pois eles colocaram o mesmo cenário para informar aquele que causa maior impacto bem como aquele que não causava impacto em sua experiência (Tabela 8).

Tabela 8 - Cenários que causaram maior e menor impacto de acordo com os participantes

ID USUÁRIO	MAIOR IMPACTO (SET 1)	MENOR IMPACTO (SET 1)	MAIOR IMPACTO (SET 2)	MENOR IMPACTO (SET 2)
User76	Cenário IV	Cenário IV	Cenário I	Cenário I
User99	Cenário II	Cenário II	Cenário II	Cenário II
User118	Cenário IV	Cenário IV	Cenário III	Cenário III
User159	Cenário IV	Cenário I	Cenário II	Cenário II

Esses participantes foram excluídos, porque não é possível um cenário causar maior e menor impacto negativo na experiência do usuário ao mesmo tempo. Desse modo, para não causar viés nos dados, esses questionários foram excluídos. Por fim, como seguimos as regras estipuladas pelo comitê de ética, o participante User158 foi descartado por informar que tinha menos de 18 anos.

Desse modo, obteve-se uma amostra final com 121 participantes (Seção 3.3.5). Esses dados analisados consistem de informações de 68 usuários do gênero masculino e 53 do gênero feminino. Essa distribuição pode ser observada a seguir, por meio da Figura 21. Nota-se que 56% dos usuários que participaram da pesquisa *online* são do gênero masculino.

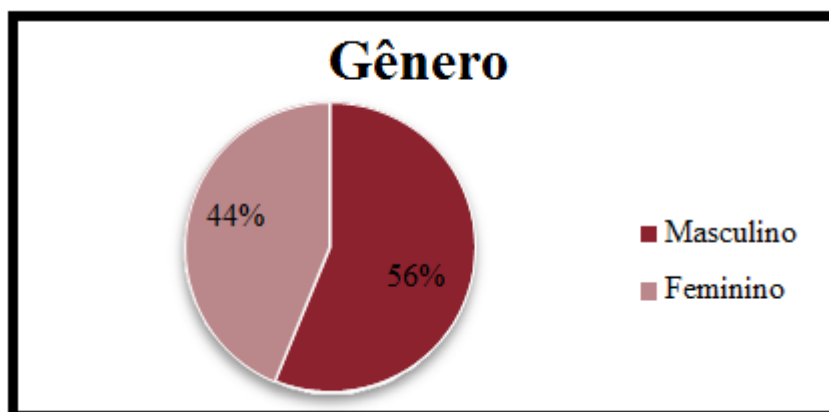


Figura 21 - Distribuição dos participantes por gênero

A Tabela 9 apresenta as análises descritivas referentes às variáveis: faixa etária e área de estudo. É possível observar que a maioria (58,600%) dos participantes está na faixa etária de 25 a 44 anos. Além disso, pode-se notar que 54,170% são da área de TI.

Tabela 9 - Distribuição dos participantes por formação profissional e faixa etária

FORMAÇÃO PROFISSIONAL	TI	OUTRAS ÁREAS	TOTAL	
Nº	66	55	121	
%	54,170%	45,840%	100%	
FAIXA ETÁRIA	18 A 24	25 A 44	45 A 65	TOTAL
Nº	25	71	25	121
%	20,700%	58,600%	20,700%	100%

As Tabelas 10 e 11 apresentam o grau de escolaridade, o nível de importância do computador para o dia a dia e o nível de conhecimento em TI informado pelos participantes, respectivamente. É possível concluir que o nível de conhecimento em TI foi alto, predominando em mais de 60% dos participantes. Além disso, pode-se observar que mais de 90% dos participantes informaram que o computador é extremamente importante para o seu dia a dia. Podemos observar também que 50,890% dos participantes possuem pós-graduação.

Tabela 10 - Distribuição dos participantes por grau de escolaridade, importância do computador e conhecimento TI

	ENSINO MÉDIO COMPLETO	SUPERIOR INCOMPLETO	SUPERIOR COMPLETO	PÓS-GRADUAÇÃO	TOTAL
Nº	1	21	35	64	121
%	0,830%	17,350%	28,920%	52,390%	100%

Tabela 11 - Distribuição dos participantes por nível de importância do computador e conhecimento em TI

IMPORTÂNCIA DO COMPUTADOR	BAIXO	MÉDIO	ALTO	TOTAL
Nº	0	1	120	121
%	0	0,830%	99,170%	100%
CONHECIMENTO EM TI	BAIXO	MÉDIO	ALTO	TOTAL
Nº	3	34	84	121
%	2,480%	28,090%	69,430%	100%

O sistema operacional mais utilizado pelos participantes desta pesquisa é o Windows 10 e, em segundo lugar, o Windows 7 e o Linux. Sendo assim, apesar de a pesquisa de mercado mostrar que o Microsoft Windows 7 é o sistema operacional mais utilizado, a amostra de participantes desta pesquisa utiliza mais o Microsoft Windows 10. Essa distribuição pode ser observada por meio da Figura 22.

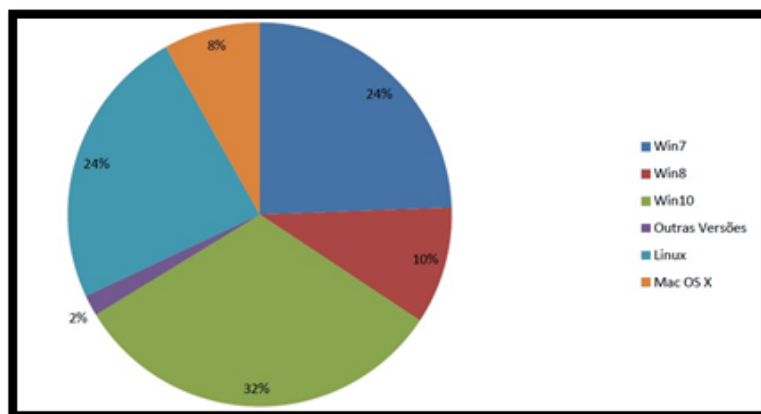


Figura 22 - Distribuição dos sistemas operacionais utilizados pelos participantes desta pesquisa

A Tabela 12 apresenta os níveis de conhecimento dos participantes definidos por meio das respostas informadas nas questões técnicas. Como podemos observar nos dados apresentados a seguir, 51,600% dos participantes são do grupo de usuários leigos e 69,170% possuem nível alto de

conhecimento em TI (Tabela 12). Esses dados foram coletados por meio de perguntas de múltipla escolha e foram analisados de acordo com o número de acertos. Quando o participante acertava acima de 70%, entrava para o grupo de usuários intermediários. Quando acertava abaixo de 69%, entrava para o grupo de usuários leigos.

Tabela 12 - Distribuição dos participantes por conhecimento básico em TI e grupo de usuários

CONHECIMENTO EM SO	LEIGOS	INTERMEDIÁRIOS	TOTAL	
Nº	61	60	121	
%	48,330%	51,670%	100%	
CONHECIMENTO EM TI	BAIXO	MÉDIO	ALTO	TOTAL
Nº	12	25	84	121
%	10%	20,830%	69,170%	100%

Após realizar a análise descritiva dos dados e categorizar cada participante de acordo com seu grupo de usuários (Leigos e Intermediários), primeiramente foram identificados os cenários que causaram maior e menor impacto de cada conjunto na qualidade da experiência dos participantes. Posteriormente, cada grupo foi analisado separadamente, objetivando verificar se o nível de conhecimento influenciou o impacto que os cenários de falha causaram na QoE dos participantes. Em seguida, foram aplicados os testes estatísticos, objetivando verificar se as frequências desses cenários possuíam ou não diferenças significativas e verificar se um tipo de falha, quando ocorrida em cenários diferentes, possuía o mesmo nível de impacto ou era percebida de forma diferente pelo usuário. Todos os resultados encontrados foram explicados na Seção 5.

Após realizar-se essa análise, iniciou-se o processo do estudo de observação, definindo as falhas simuladas na ferramenta de *software*, os cenários utilizados, bem como a criação do questionário presencial. A Seção 4.3 apresenta detalhadamente a realização de cada uma dessas etapas.

4.3 Questionário presencial

A partir dos resultados encontrados por meio do questionário online, foi possível criar quatro cenários com diferentes atividades que os participantes realizaram durante o estudo de observação. Esses cenários apresentaram diferentes contextos (profissional prazo longo, profissional prazo curto, lazer prazo longo e lazer prazo curto) para a mesma sequência de atividades computacionais e eventos de falha. Os contextos profissionais apresentaram a mesma descrição, variando apenas o prazo. Esse fato também ocorreu nos contextos de lazer. A Figura 23 apresenta a descrição dos contextos de lazer prazo curto e profissional prazo longo.

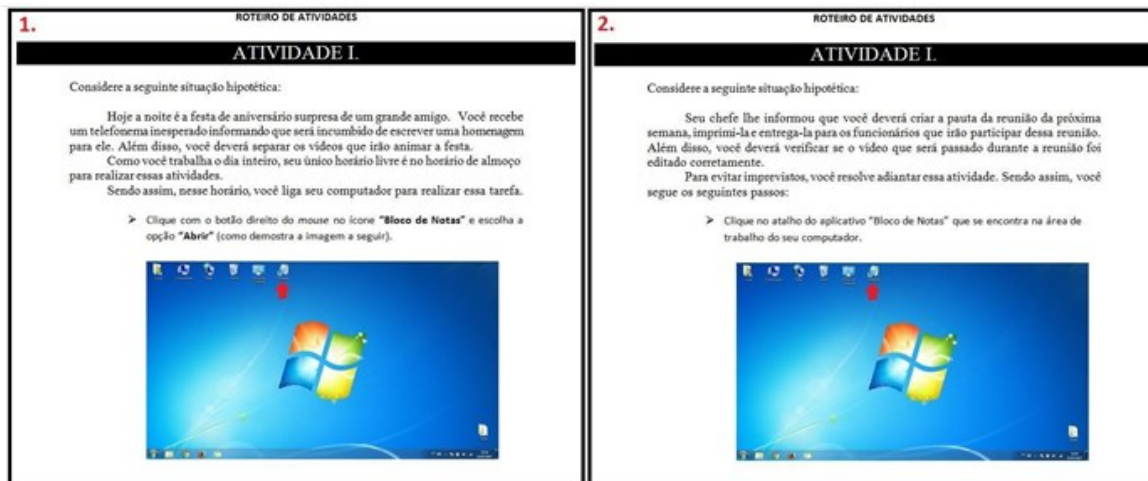


Figura 23 - Roteiros de atividades realizadas pelo participante durante a sessão de observação

No estudo de observação, os participantes passaram por duas etapas. Na primeira, eles responderam a um formulário que teve por objetivo classificá-los em usuários leigos ou intermediários. Após realizar essa investigação, eles foram sorteados aleatoriamente e encaixados em um dos quatro cenários de falha utilizados na etapa presencial. A Figura 24 apresenta o modo como o número de participantes foi dividido. Observe que eles foram alocados de acordo com o contexto descrito em cada cenário de falha utilizado.

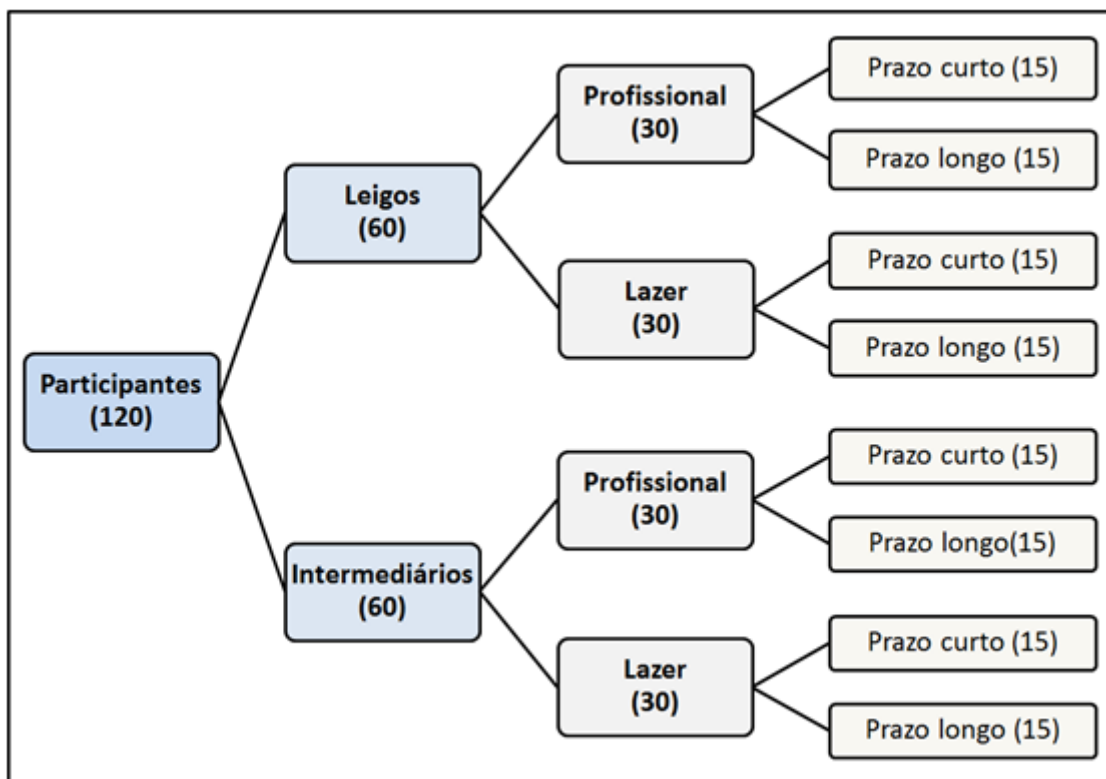


Figura 24 - Divisão dos participantes do estudo de observação

A Figura 25 apresenta o fluxo de trabalho do estudo de observação, destacando cada tipo de evento de falha simulado e a ordem em que eles foram apresentados para o participante.

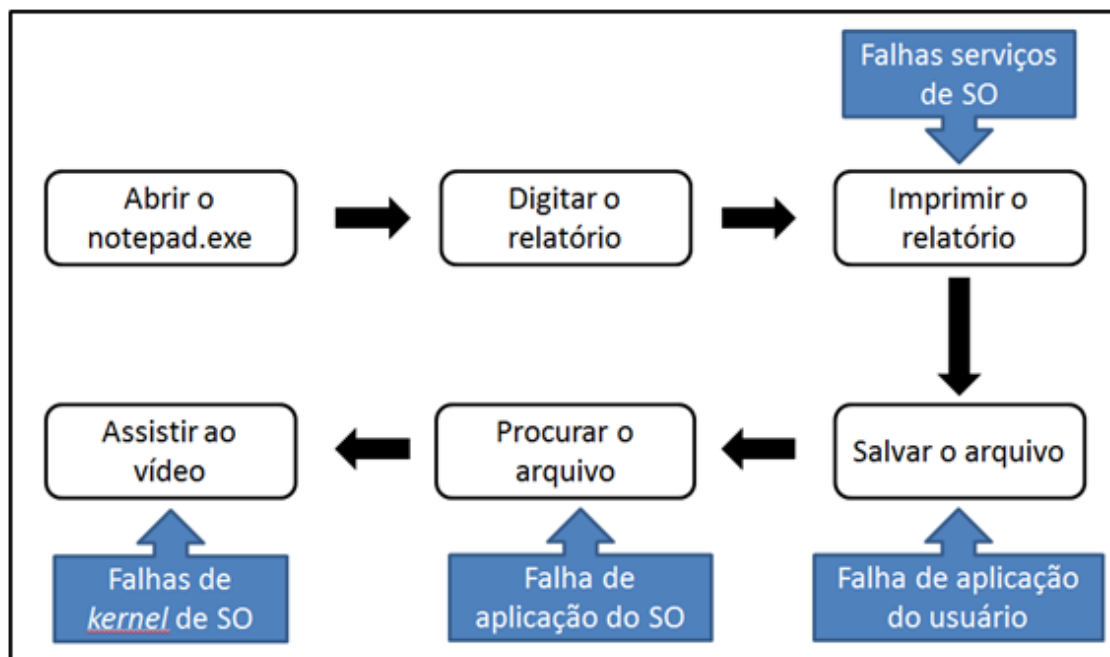


Figura 25 - Fluxo de trabalho do estudo de observação

Foram aplicados os mesmos quatro tipos de falhas para todos os cenários investigados, variando apenas as características do contexto, com o propósito de observar seus efeitos na qualidade da experiência do usuário em relação à confiabilidade do sistema operacional. No fim da ocorrência de cada evento de falha, a utilização da ferramenta era pausada para que os participantes pudessem responder às perguntas relacionadas ao evento de falha vivenciado.

Os participantes executaram essa sequência de atividades (Figura 25) no computador que possuía a ferramenta de *software* configurada e, durante a sessão de uso dessa ferramenta, após experimentarem determinado evento de falha, eles responderam a 13 questões de múltipla escolha. Essas questões avaliaram o nível de tolerância, de confiança e de percepção do participante após a ocorrência de uma falha, objetivando compreender o impacto que essas falhas causaram em sua QoE (Apêndice A).

Para verificar se as informações coletadas por meio desse questionário atendiam ao propósito desta pesquisa, foi realizado um pré-teste. Ele foi aplicado com sete participantes para verificar se as perguntas eram de fácil interpretação e se algum ajuste seria necessário. O cenário utilizado foi escolhido por meio de sorteio aleatório e, assim, o participante seguiu o passo a passo daquele que foi sorteado. Para realizar o pré-teste, foram utilizados dois procedimentos diferentes:

- **Método I:** Três participantes seguiram o roteiro e responderam ao questionário sem a interferência do pesquisador.
- **Método II:** O pesquisador que estava conduzindo o estudo leu o roteiro e informou ao participante o passo a passo que ele deveria realizar. No decorrer da sessão de observação, quando o participante finalizava cada atividade, as respectivas perguntas do questionário foram lidas.

Os participantes foram monitorados por um segundo pesquisador, e tudo foi anotado em planilhas, como o tempo que o usuário levou para finalizar a pesquisa, as dúvidas que ele teve durante a execução

das atividades, entre outros fatores. A Tabela 13 apresenta as informações que foram coletadas durante o pré-teste.

Tabela 13 - Variáveis monitoradas durante a aplicação do pré-teste

VARIÁVEIS MONITORADAS	MÉTODO I	MÉTODO II
Tempo	40/50 min	20/25 min
Questionário	Questões em branco.	Todas as perguntas foram respondidas.
Informações extras	Não	Sim
Contexto	A situação nem sempre foi lembrada.	Em todas as atividades, a situação hipotética foi lembrada.
Percepção da falha	Nem sempre a falha foi percebida.	Em todas as atividades, a falha foi percebida.
Passos do roteiro	Tiveram dúvidas.	Não tiveram dúvidas.

É possível perceber que os participantes do método I demoraram mais tempo para finalizar o questionário, comparados com aqueles que fizeram parte do método II. Os participantes do método I não levaram em consideração a situação hipotética que foi descrita no início do cenário. Acredita-se que esse fato tenha ocorrido, pois o cenário apresentou muitas informações e, no decorrer do estudo, os participantes esqueceram o que estava descrito anteriormente. No método II, o pesquisador estava lendo o roteiro e, por isso, foi possível lembrar ao participante o contexto em que ele estava inserido.

Durante o teste, observou-se que os participantes do método I deixaram várias questões em branco e, além disso, pularam algumas etapas que deveriam ser realizadas na atividade. Por exemplo, “Procure o arquivo que você salvou na pasta documentos. Para isso, siga os seguintes passos:”, ao invés de seguirem o que estava descrito, eles seguiam o passo a passo que estavam acostumados a realizar e, conseqüentemente, não presenciaram nenhuma falha.

Os participantes que fizeram parte do método I ficaram confusos, pois não sabiam quando teriam que ler a descrição do cenário e quando teriam que responder ao questionário. Em alguns momentos, a mensagem de falha foi apresentada, mas os participantes não estavam olhando para a tela, pois estavam lendo o cenário. Conseqüentemente, eles não tinham uma percepção relacionada à falha que havia ocorrido.

Por fim, no método I, como as reações dos participantes não foram monitoradas, não foi possível verificar se eles estavam respondendo a todas as questões corretamente. Diferentemente do método II, em que foi possível acompanhar o passo a passo que os participantes estavam seguindo. Sendo assim, eles realizaram exatamente o que precisava ser feito e, conseqüentemente, presenciaram a falha e não deixaram nenhuma questão do questionário em branco.

Foi possível obter mais informações na sessão de observação aplicada para os participantes que fizeram parte do método II, pois, durante a entrevista, os participantes informaram as justificativas de forma completa, apresentando informações extras, além do que era questionado. Por exemplo, citavam conseqüências que essas falhas já geraram no seu dia a dia, relatavam que sentiam falta de informações

nas mensagens de falha, entre outros. Seguindo essas observações encontradas durante o estudo, o método II foi o procedimento escolhido para coletar os dados desta pesquisa.

O questionário aplicado para a classificação dos participantes foi criado levando em consideração os conceitos definidos para os grupos de usuários leigos e intermediários, apresentados na Seção 3.2.5. Além disso, esse questionário seguiu o mesmo modelo do questionário aplicado para a categorização dos participantes na pesquisa *online*, apresentando oito questões de múltipla escolha.

Apesar de essas questões terem atingido seu propósito, algumas modificações foram realizadas. Para categorizar os participantes, eles deveriam saber diferenciar as aplicações do usuário das aplicações do sistema operacional. Desse modo, a questão responsável por coletar essas informações, antes do pré-teste, tinha oito opções de aplicações e, após realizar as modificações (Figura 26), duas opções foram inseridas, totalizando dez opções de respostas.

10. Informe quais dos programas abaixo são aplicações do sistema operacional (SO), ou seja, programas que são instalados com o SO para realizarem alguma função de responsabilidade do SO.

- **Gerenciador de impressão.**
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- **Editor de Texto.**
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- **Calculadora.**
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- **Interpretador de linha de comando (Ex.: *Prompt* de comando).**
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- **Tocador (player) de áudio e vídeo.**
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- **Editor Gráfico**
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- **Navegador de Internet.**
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- **Formatador de disco.**
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- **Gerenciador de e-mail.**
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei
- **Gerenciador de Tarefas.**
 - Faz parte do SO
 - Não faz parte do SO
 - Não sei

Figura 26- Exemplo de pergunta apresentada no questionário presencial (após alterações do pré-teste)

O questionário aplicado durante o estudo de observação apresentou uma questão na qual se perguntou aos participantes se a falha apresentada era de *software* ou de *hardware*. Para que não houvesse dúvidas com relação às informações coletadas nessa questão, foi inserida uma pergunta para investigar se os usuários sabiam diferenciar *hardware* de *software*. Por fim, com as informações coletadas no pré-teste, várias modificações foram realizadas no questionário presencial.

Um dos propósitos desse questionário foi investigar o nível de confiança do participante no *software* que ele estava utilizando após a ocorrência da falha. Ao ser questionado, um dos participantes do pré-teste informou que não confiava no *player* e editor de texto utilizado na sessão de pesquisa, pois estava acostumado a utilizar outro. Desse modo, para não gerar influência na resposta apresentada a essa pergunta, foi solicitado que o participante imaginasse que o *software* utilizado durante o estudo fosse o que ele estava acostumado a usar no seu dia a dia.

No fim do pré-teste, dois participantes informaram, sem serem questionados, que a falha de aplicação de SO (*explorer.exe*) foi a que causou maior impacto. Levando em consideração o propósito desta pesquisa, essa pergunta foi incluída no questionário. Além disso, alguns participantes classificaram

o serviço oferecido pelo sistema operacional como ruim e informaram que isso afetava muito a confiança atribuída ao sistema operacional que estavam utilizando.

Os participantes acreditam que os desenvolvedores não pensaram nos usuários leigos, pois as mensagens de falha não contêm muitas informações necessárias. Eles sentem falta de um *link* que os direcione a um passo a passo para resolver o problema, pois a falta de informação afeta seu grau de tolerância. Como o objetivo da pesquisa é avaliar a qualidade da experiência do usuário (QoE) em confiabilidade de sistemas operacionais, e a QoE é a opinião/percepção do usuário quanto à qualidade do serviço que está sendo oferecido, essa questão foi incluída no questionário final, para apresentar como o usuário classifica o serviço e a experiência que ele teve ao presenciar essas falhas.

Em vários momentos, foi possível escrever as reações de alguns participantes e também coletar algumas informações adicionais que eles falaram durante o estudo de observação (Apêndice D). Por fim, após realizar todas as alterações necessárias, o questionário utilizado para classificar os participantes ficou com dez questões, enquanto o questionário utilizado na segunda etapa apresentou treze perguntas, que foram repetidas após a ocorrência de cada falha.

Após realizar todos os ajustes necessários no questionário, deu-se início às sessões de observação. As sessões foram realizadas respeitando-se a disponibilidade dos participantes e, por isso, foram concretizadas em três cidades diferentes. Em Uberlândia-MG, os participantes utilizaram a ferramenta e responderam ao questionário na Universidade Federal de Uberlândia. Em Uberaba-MG, os participantes responderam ao questionário na Biblioteca Municipal e alguns participantes marcaram o encontro em sua própria residência. Em Catalão-GO, eles foram encontrados nos locais de sua preferência (ex. em suas residências, em locais próximos ao seu trabalho).

Nos locais estipulados pelos participantes para realizar a entrevista, foi montado um ambiente de estudo de observação com três computadores (*laptops*) e uma impressora. O participante utilizou o computador que estava com a ferramenta de simulação de falhas instalada, enquanto os outros dois computadores ficaram com os pesquisadores para que eles pudessem anotar as respostas da entrevista, bem como as reações e dados extras apresentados pelo participante. A Figura 27 mostra o ambiente montado em uma visitação para que se realizasse o estudo de observação.

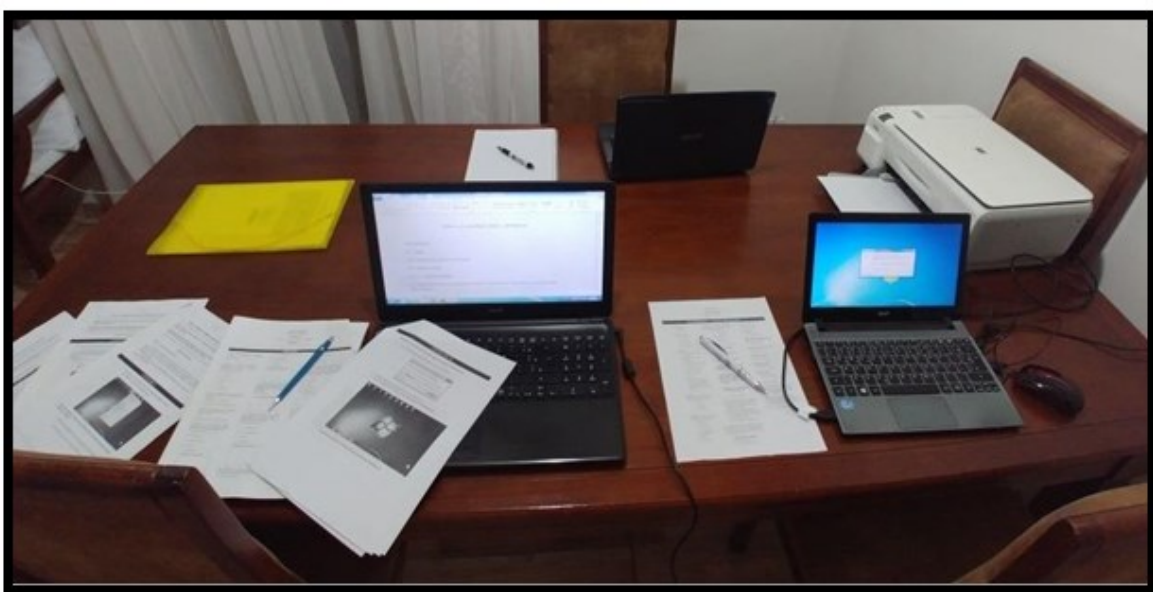


Figura 27 - Ambiente do estudo de observação montado em uma visitação residencial

Por fim, obteve-se uma amostra de 139 questionários, entretanto 19 foram descartados. Alguns participantes precisaram se retirar do local onde estava sendo realizada a observação para atender a ligações ou realizar alguma outra atividade. Quando eles se retiravam do local para voltarem algum tempo depois, o impacto causado pelas falhas consecutivas era perdido. Sendo assim, quando eles precisavam se retirar no meio da investigação, os dados coletados para esses participantes eram descartados. Além disso, por estarem utilizando o celular, eles não presenciavam a falha no momento exato em que ela ocorria. O impacto emocional (ex.: susto) causado no ato da ocorrência da falha também era perdido e, por isso, os dados desses participantes também foram descartados.

Outro motivo que levou alguns participantes a serem descartados estava relacionado com a situação hipotética, pois eles informaram que não conseguiam se imaginar naquela situação. Desse modo, eles não conseguiram informar o seu grau de tolerância à falha, o seu grau de confiança no funcionamento do SO, entre outras questões.

Dois participantes não perceberam algumas falhas, pois interpretaram a mensagem como um aviso, e não como uma falha que havia ocorrido e, por isso, não conseguiram responder às questões relacionadas à percepção da falha. Desse modo, como esse fator não está no escopo deste trabalho, esses participantes também foram descartados da amostra final desta pesquisa.

Um dos participantes, ao chegar ao ambiente de estudo, escutou outros voluntários que já haviam participado desta pesquisa, relatando as falhas que iriam aparecer e quais eram as falhas do sistema operacional. Desse modo, esse participante não foi incluído na amostra final de dados desta pesquisa.

Por conseguinte, após realizar esses descartes, obteve-se uma amostra final com 120 participantes, sendo que 60 são do grupo de usuários leigos e 60 do grupo de usuários intermediários (Seção 3.2.3). Essa amostra de dados está dividida em 74 participantes do gênero masculino e 46 do gênero feminino (Figura 28).

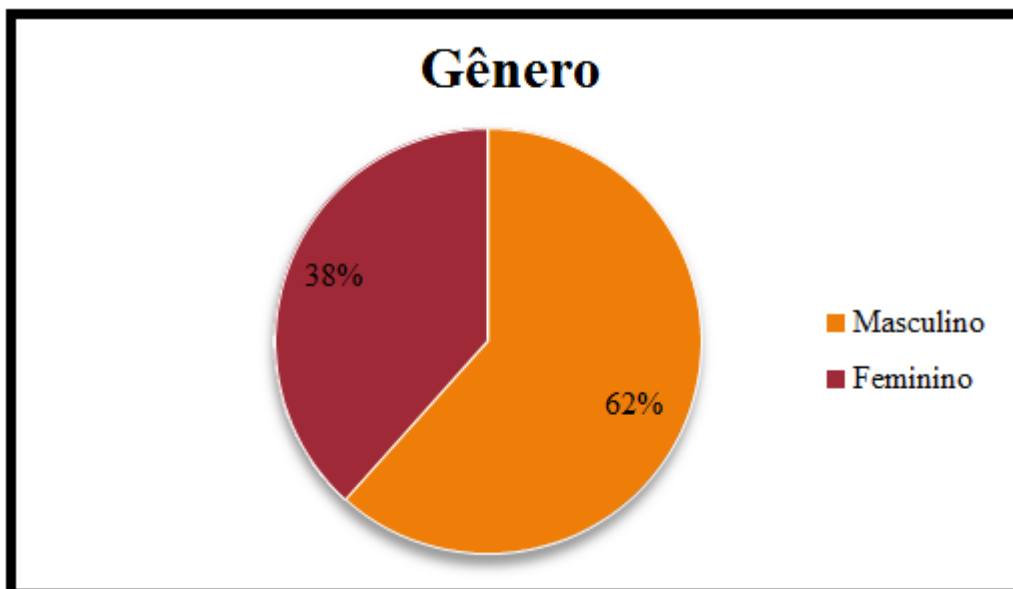


Figura 28 - Distribuição dos participantes por gênero

A maioria dos usuários participantes está em uma faixa etária entre 18 e 44 anos, 36% são de uma faixa etária entre 18 e 24 anos e 60% estão entre 25 e 44 anos. Apenas 4% dos participantes estão acima de 44 anos (Figura 29).

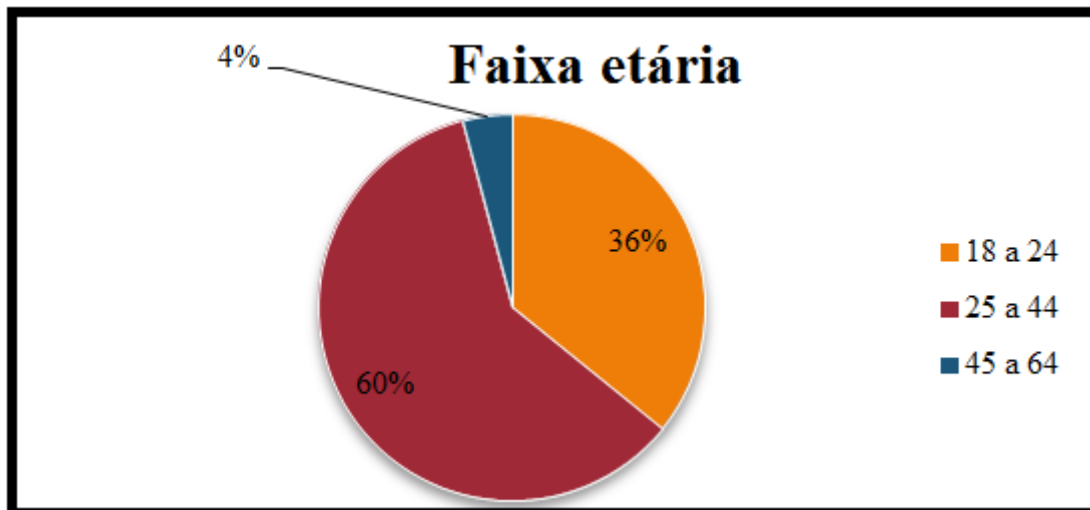


Figura 29 - Distribuição dos participantes por faixa etária

Nessa amostra de dados, 73 participantes são de outras áreas de estudo e 47 da área de TI. Os dados foram coletados em diferentes cidades e locais, por isso a maioria dos participantes desta pesquisa não faz parte da área de TI. Eles são advogados, jornalistas, profissionais da área de saúde e funcionários de departamento administrativo, estudantes de engenharia e licenciaturas. Sendo assim, pode-se concluir que 61% dos participantes pertencem a outras áreas. (Figura 30).

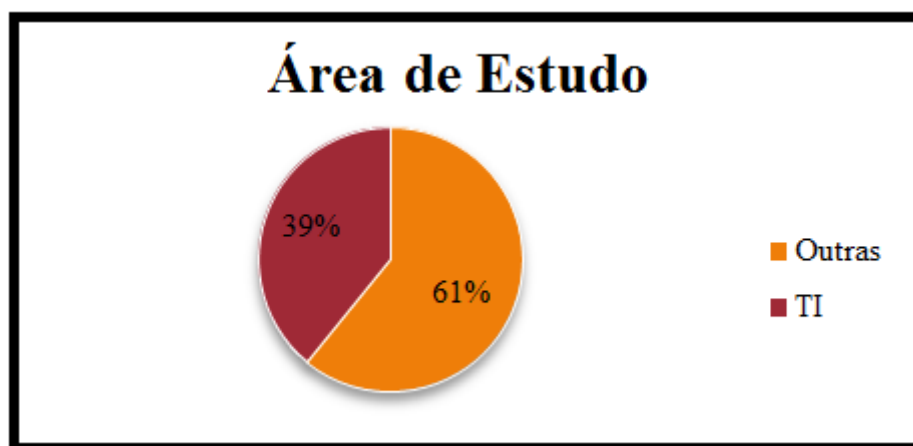


Figura 30 - Distribuição dos participantes por área de estudo

É possível concluir, analisando a Tabela 14, que o nível de conhecimento dos usuários variou entre médio e alto. Além disso, pode-se observar que os participantes estão divididos entre grau de ensino superior incompleto, superior completo e pós-graduação. A maioria, isto é, 35%, estão finalizando a graduação e 30% possuem superior completo, enquanto 27% possuem pós-graduação (Figura 31).

Tabela 14 - Distribuição dos participantes por nível de conhecimento em TI e escolaridade

NÍVEL DE CONHECIMENTO TI		
ALTO	MÉDIO	BAIXO
38	68	14

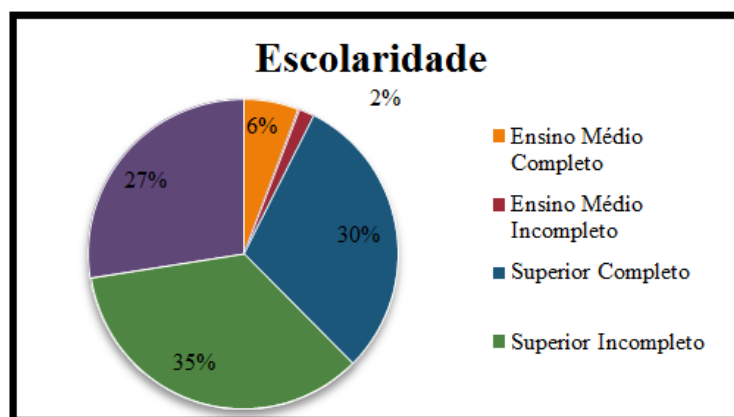


Figura 31 - Distribuição dos participantes por escolaridade

A maioria dos participantes deste estudo utiliza o Microsoft Windows 10 atualmente. Muitos marcaram mais de um sistema operacional, pois utilizam um SO no trabalho e outro em sua residência. A Tabela 15 apresenta a frequência em que cada SO foi marcado. Pode-se concluir que os participantes desta pesquisa utilizam mais o Windows 10.

Tabela 15 - Sistemas operacionais utilizados

SO UTILIZADO PELOS PARTICIPANTES				
LINUX	WIN10	WIN7	WIN8	MAC OS X
18	71	33	14	8
12%	49%	23%	10%	6%

Após realizar a categorização dos participantes, os dados de cada grupo de usuários foram analisados separadamente, objetivando avaliar, assim como foi realizado no questionário *online*, se o nível de conhecimento influenciou o impacto causado na QoE dos participantes do estudo de observação. Posteriormente, os quatro cenários utilizados foram analisados, com o intuito de verificar se os níveis de impacto, de tolerância e de confiança foram diferentes de acordo com o contexto em que os participantes estavam inseridos. Os resultados encontrados foram explicados na Seção 5.

Resultados

5.1 Q1: O usuário, ao utilizar o computador para realizar tarefas de rotina (ex.: digitar textos, procurar arquivos, assistir vídeos), consegue identificar quando uma falha de operação do computador é consequência de uma falha de SO?

Para analisar a percepção dos participantes, primeiramente eles foram questionados sobre a mensagem de falha que havia surgido. Os dados mostram que 78% dos participantes do grupo de usuários leigos (G1) e 80% dos participantes do grupo de usuários intermediários (G2) acreditam que, após a ocorrência da falha de serviço de SO, svchost.exe, o sistema operacional parou de funcionar corretamente (Tabela 16 e 17).

Tabela 16 - Falhas que afetaram o funcionamento correto do sistema operacional de acordo com os participantes (continua na próxima página)

EVENTO DE FALHA 01 – EF01(<i>svchost.exe</i>)				
AFETOU FUNCIONAMENTO DO SO	Nº (G1)	%	Nº (G2)	%
NÃO	47	78%	48	80%
SIM	9	15%	11	2%
NÃO SEI	4	7%	1	18%
EVENTO DE FALHA 02 – EF02 (<i>notepad.exe</i>)				
AFETOU FUNCIONAMENTO DO SO	Nº (G1)	%	Nº (G2)	%
NÃO	54	90%	43	72%
SIM	4	7%	15	25%
NÃO SEI	2	3%	2	3%
EVENTO DE FALHA 03 – EF03 (<i>explorer.exe</i>)				
AFETOU FUNCIONAMENTO DO SO	Nº (G1)	%	Nº (G2)	%
NÃO	54	90%	52	90%
SIM	4	7%	6	10%
NÃO SEI	2	3%	0	0%

Tabela 17 - Conclusão da tabela anterior (continuação da Tabela 16).

EVENTO DE FALHA 04 – EF04 (VIDEO_TDR_ERROR).				
AFETOU FUNCIONAMENTO DO SO	Nº (G1)	%	Nº (G2)	%
NÃO	59	98%	56	93%
SIM	1	2%	4	7%
NÃO SEI	0	0	0	0

Pode-se observar, também, que, em todas as falhas, os participantes acreditaram que o sistema operacional havia parado de funcionar. Desse modo, conclui-se que, para a maioria dos participantes, independentemente da falha que tenha ocorrido, ela afetou o funcionamento correto do sistema operacional. Após a ocorrência de um evento de falha, os participantes informaram quais falhas, em sua opinião, eram específicas do SO. Ao analisar os dados, foi encontrada pouca diferença entre as respostas dos grupos G1 e G2. Também se notou que apenas no evento de falha 02 (EF02) foi possível observar uma grande diferença entre os valores encontrados, pois apenas 3% da amostra do G1 acertaram ao informar que essa falha não era uma falha do SO. Todas as respostas informadas foram separadas em “certas”, “erradas” e em “não souberam responder a essa questão”. A Tabela 18 demonstra o número de erros e acertos encontrados na análise dos dados.

Tabela 18 - Erros e acertos dos participantes ao informarem qual das falhas era consequência de uma falha do sistema operacional

EVENTO DE FALHA 01 – EF01(svchost.exe)								
	Acertos	%	Erros	%	Não sei	%	Total	%
G 1	17	28%	15	25%	28	47%	60	100%
G 2	22	37%	21	35%	17	28%	60	100%
EVENTO DE FALHA 02 – EF02 (notepad.exe)								
	Acertos	%	Erros	%	Não sei	%	Total	%
G 1	3	3%	30	50%	27	45%	60	100%
G 2	15	25%	20	33%	25	42%	60	100%
EVENTO DE FALHA 03 – EF03 (explorer.exe)								
	Acertos	%	Erros	%	Não sei	%	Total	%
G 1	33	55%	4	7%	23	38%	60	100%
G 2	38	63%	7	12%	15	25%	60	100%
EVENTO DE FALHA 04 – EF04 (VIDEO_TDR_ERROR)								
	Acertos	%	Erros	%	Não sei	%	Total	%
G1	18	30%	12	20%	30	50%	60	100%
G2	24	40%	16	27%	20	33%	60	100%

Pode-se observar que a EF02 apresentou diferença entre os grupos estudados. Além disso, os dados foram bem distribuídos, apresentando pouca diferença entre o número de erros e de acertos, bem como entre os grupos de usuários. Para avaliar se essas diferenças encontradas foram significativas e analisar se o número de acertos e erros depende do nível de conhecimento dos grupos estudados, aplicou-se o teste exato de Fisher. Dado nível de significância ($\alpha = 0,050$), como se pode observar na Tabela 19, para as falhas apresentadas nos eventos EF01, EF03 e EF04 a H_0 não foi rejeitada, mostrando que **não há** diferença entre as respostas informadas e os grupos estudados, visto que o p-valor encontrado foi maior que o nível de confiança de 0,050 adotado. Desse modo, tem-se que as respostas estão distribuídas igualmente entre os grupos estudados, não apresentando diferença significativa. Logo, o nível de conhecimento dos participantes não foi um fator que influenciou as respostas encontradas.

Tabela 19 - Resultados encontrados no teste exato de Fisher

FALHAS	p-valor
svchost.exe (EF01)	0,120
notepad.exe (EF02)	0,005
explorer.exe (EF03)	0,260
VIDEO_TDR_ERROR (EF04)	0,370

No teste aplicado para a falha apresentada no evento EF02, obteve-se como resultado p-valor = $0,005 < 0,050$. Na Tabela 18, foi possível observar que o número de acertos da EF02 foi muito menor comparado com as outras atividades. Apenas 3 participantes (3%) do grupo G1 e 15 (25%) do grupo G2 acertaram o questionamento, informando que a falha do notepad.exe não era uma falha do sistema operacional. Observa-se que, nesse caso, o número de participantes do grupo G2 que acertaram esse questionamento foi maior que o número de participantes do grupo G1. Desse modo, apenas nessa falha foi possível observar que há diferença significativa entre as variáveis.

A falha apresentada na EF02 ocorreu devido ao funcionamento incorreto do editor de texto bloco de notas, sendo categorizada como falha de aplicação do usuário. Como foi observado nas Tabelas 16 e 17, mais de 80% dos participantes acreditaram que o sistema operacional não estava funcionando corretamente após a ocorrência das falhas. Acredita-se que essa diferença tenha ocorrido porque os participantes atribuem a causa da falha ao sistema operacional, independentemente de a falha ocorrida ser, de fato, uma falha de SO.

Como o objetivo desta questão foi apenas analisar se os usuários são capazes de diferenciar uma falha de aplicação de uma falha de SO, os participantes não foram questionados sobre o porquê de eles acreditarem que a falha apresentada era uma falha de SO. Alguns participantes, no entanto, apresentaram uma justificativa, que pode ser observada na Tabela 20.

Tabela 20 - Justificativas apresentadas pelos participantes

FALHA É UMA CONSEQUÊNCIA DO MAU FUNCIONAMENTO DO SO?	JUSTIFICATIVA
SIM (<i>sysmain.dll</i>) – G1	O programa para imprimir não está instalado.
SIM (<i>sysmain.dll</i>) – G1	A impressão faz parte do Windows.
SIM (Bloco de Notas) – G1	Devido à mensagem de falha (está informando que parou de funcionar).
SIM (Bloco de Notas) – G1	A mensagem de falha está informando que o Windows está procurando uma solução.
SIM (Bloco de Notas) – G1	Não respondeu à função que solicitei.
SIM (<i>explorer.exe</i>) – G1	A própria mensagem informou que ele parou de funcionar.
NÃO (<i>sysmain.dll</i>) – G1	É apenas algo que está faltando no computador.
SIM (Bloco de notas) – G1	Quando penso no sistema, penso no computador como um todo. Então penso que a falha seja consequência dele.
SIM (<i>sysmain.dll</i>) – G2	Dll faz parte do SO.
SIM (<i>sysmain.dll</i>) – G2	É uma Dll.
SIM (<i>sysmain.dll</i>) – G1	Por causa da Dll.
SIM (<i>sysmain.dll</i>) – G2	Deu problema na Dll, que é um arquivo de comando interno.
SIM (<i>sysmain.dll</i>) – G1	Eu acho que sim por ser uma Dll.

Desses participantes que informaram uma justificativa, apenas 5 apresentaram ter conhecimento suficiente para classificar a falha corretamente. Entretanto nenhum deles classificou a falha EF02 corretamente (Tabela 21).

Tabela 21 - Justificativas encontradas para o questionamento sobre a falha ser consequência do SO

GRUPO	svchost.exe	GRUPO	notepad.exe
G2	"Sim por causa da dll."	G2	"Vou dizer que sim, pois pode ser um problema de alocação de memória."
G2	"Sim, é uma dll, arquivo de comando interno."	G2	"Não sei."
G1	"Sim, eu acho que sim por ser dll."	G1	"Sim."
G1	"Sim, está faltando	G1	"Sim."

Desse modo, ao verificar essas justificativas, é possível observar que, apesar de a percentagem de acertos ter sido relativamente grande, não se pode afirmar que os participantes que informaram corretamente quais eram as falhas de SO possuem conhecimento suficiente para responder a essa questão. Essa resposta pode depender de outros fatores, como informações apresentadas na mensagem de falha ou a consequência gerada pela falha, isto é, “travamento” da tela, entre outros.

Neste estudo, foi possível observar que os usuários acreditam que as falhas, independentemente de serem ou não consequência do SO, afetam o funcionamento correto do sistema operacional. Sendo assim, levando em consideração todos os fatores citados nesta seção, tem-se como resultado que a maioria dos usuários estudados nesta pesquisa não foi capaz de informar quando uma falha é consequência do sistema operacional. Além disso, observou-se que ambos os grupos de usuários erraram esse questionamento igualmente, mostrando que o número de erros encontrados não está associado ao nível de conhecimento dos usuários investigados nesta pesquisa.

Por fim, para afirmar se os usuários possuem ou não conhecimento suficiente para acertar esse questionamento, primeiramente é necessário avaliar outros fatores relacionados com a falha que ocorreu, como a mensagem de falha. Como esses fatores foram encontrados neste estudo, pretende-se investigar melhor todas essas características posteriormente.

5.2 Q2: Qual o nível de tolerância do usuário para falhas de SO?

Para avaliar a confiança do usuário no funcionamento do sistema operacional, é necessário analisar o impacto que a falha causa em sua experiência. Durante o estudo de observação realizado nesta pesquisa, foram analisados os fatores que influenciaram a tolerância do usuário para falhas de SO. Para esse fim, o nível de tolerância foi investigado levando em consideração o cenário no qual o participante estava inserido, bem como o tipo de falha que havia ocorrido.

Os quatro cenários utilizados no questionário presencial foram analisados separadamente com o intuito de verificar se os níveis de tolerância dos participantes foram diferentes de acordo com o contexto em que eles estavam inseridos. Além disso, os níveis de tolerância foram avaliados de acordo com o grupo de usuários (leigos e intermediários) para avaliar se os níveis de conhecimento influenciam seus níveis de tolerância. A Figura 33 apresenta os níveis de tolerância de acordo com o contexto e grupo em que os participantes estavam inseridos.

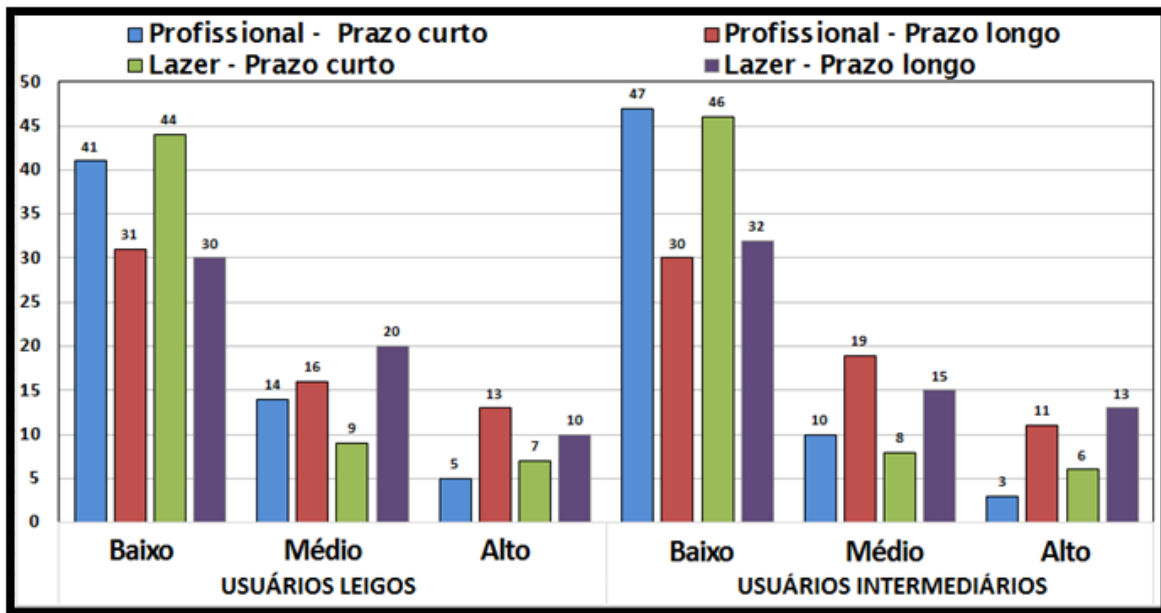


Figura 32- Níveis de tolerância separados por contexto e por grupo de participantes

O nível de tolerância dos participantes, ao presenciarem diferentes eventos de falha, foi baixo tanto para os participantes leigos quanto para os participantes intermediários. É possível observar que o nível de tolerância dos participantes quanto às falhas que ocorreram em contextos com natureza de atividade diferente, isto é, lazer ou profissional, não apresentaram uma diferença significativa. Contudo o nível baixo de tolerância com relação às falhas que ocorreram nos contextos com prazo curto de finalização das atividades foi maior quando comparado com aqueles que ocorreram em contextos que possuíam prazo longo.

Além disso, percebe-se que os participantes do grupo G1 e do grupo G2 apresentaram o mesmo nível de tolerância, podendo ser notada uma diferença mínima nas frequências observadas. Desse modo, pode-se concluir que o nível de conhecimento em sistemas computacionais não influenciou o nível de tolerância dos participantes. Pode-se inferir, também, que o prazo que os participantes tiveram para concluir determinada atividade foi a principal característica dos contextos que influenciou sua tolerância ao presenciar um evento de falha. Para testar essas hipóteses, foi investigado se as frequências encontradas para cada nível de tolerância possuíam diferença significativa entre os quatro contextos descritos nos cenários utilizados neste estudo.

Para isso, aplicaram-se dois testes estatísticos, o teste Qui-Quadrado e teste Exato de Fisher (ver Seção 3.3.2). O Teste Qui-Quadrado não deve ser utilizado em amostras pequenas (menor que 40 respondentes), quando um dos valores esperado for menor do que 5 ou igual a 0 [Lyman, 2010]. Por isso, para amostras que possuíam esses resultados, utilizamos o teste exato de Fisher. Contudo, para padronização, foi utilizado apenas o p-valor para análise dos resultados. Dado o nível de significância ($\alpha = 0,050$), como se pode observar na Tabela 22, a H0 foi rejeitada, mostrando que existe diferença entre os níveis de impacto encontrados nos contextos com prazos diferentes.

Tabela 22 - Resultado do teste Qui-Quadrado (influência dos prazos) – Usuários Leigos

LAZER LONGO VS. LAZER CURTO	
Usuários leigos	p-valor 0,006
	Resultado p-valor < 0,050 (Rejeita H0)
PROFISSIONAL LONGO VS. PROFISSIONAL CURTO	
	p-valor 0,065
	Resultado p-valor > 0,050 (Não Rejeita H0)
LAZER LONGO VS. PROFISSIONAL CURTO	
Usuários intermediários	p-valor 0,001
	Resultado p-valor < 0,050 (Rejeita H0)
PROFISSIONAL LONGO VS. PROFISSIONAL CURTO	
	p-valor 0,002
	Resultado p-valor < 0,050 (Rejeita H0)

Com base na Tabela 22, é possível observar que o prazo que os participantes, tanto leigos quanto intermediários, tiveram para concluir as atividades de lazer influenciou diretamente seus níveis de tolerância ao presenciarem um evento de falha. Contudo os resultados mostraram que, diferentemente dos participantes intermediários, os níveis de tolerância dos usuários leigos, inseridos em contexto profissional com prazo longo e prazo curto, não apresentaram uma diferença significativa.

Após realizar esses testes, foi investigado se a natureza da atividade também influenciou o nível de tolerância dos participantes ao presenciarem os eventos de falha. Dado o nível de significância ($\alpha = 0,050$), como se pode observar na Tabela 23, a H0 não foi rejeitada, mostrando que não existe diferença entre os níveis de tolerância das falhas encontradas nos contextos com natureza da atividade diferente.

Tabela 23 - Resultados do teste exato de Fisher (Influência da natureza da atividade)

Usuários leigos	PROFISSIONAL PRAZO LONGO VS. LAZER PRAZO LONGO	
	p-valor	0,648
	Resultado	p-valor > 0,050 (Não rejeita H0)
	PROFISSIONAL PRAZO CURTO VS. LAZER PRAZO CURTO	
	p-valor	0,526
	Resultado	p-valor > 0,050 (Não rejeita H0)
	PROFISSIONAL PRAZO CURTO VS. LAZER PRAZO LONGO	
	p-valor	0,813
	Resultado	p-valor > 0,050 (Não rejeita H0)
	PROFISSIONAL PRAZO LONGO VS. LAZER PRAZO CURTO	
	p-valor	0,013
	Resultado	p-valor < 0,050 (Rejeita H0)
Usuários intermediários	PROFISSIONAL PRAZO LONGO VS. LAZER PRAZO LONGO	
	p-valor	0,701
	Resultado	p-valor > 0,050 (Não rejeita H0)
	PROFISSIONAL PRAZO CURTO VS. LAZER PRAZO CURTO	
	p-valor	0,584
	Resultado	p-valor > 0,050 (Não rejeita H0)
	PROFISSIONAL PRAZO LONGO VS. LAZER PRAZO CURTO	
	p-valor	0,020
	Resultado	p-valor < 0,050 (Rejeita H0)
	PROFISSIONAL PRAZO CURTO VS LAZER PRAZO LONGO	
	p-valor	0,008
	Resultado	p-valor < 0,050 (Rejeita H0)

Os resultados (Tabela 23) mostraram que não existe diferença entre os níveis de tolerância às falhas que ocorreram nos contextos com natureza da atividade diferente, porém com o mesmo prazo para sua finalização. Os resultados mostraram também que, para os participantes leigos que fizeram parte do contexto profissional com prazo curto, os níveis de tolerância não apresentaram diferença significativa comparados com os níveis de tolerância às falhas que ocorreram nos demais contextos.

Ao analisar esses resultados (Tabela 23), conclui-se que, para avaliar a tolerância dos usuários ao presenciarem determinado evento de falha, é necessário levar em consideração os contextos, pois a percepção dos participantes variou de acordo com as características dos contextos em que estavam inseridos. Além disso, ao longo do estudo, foi possível identificar outros fatores que influenciaram os níveis de tolerância apresentados pelos participantes.

Para avaliar esses fatores, as justificativas apresentadas pelos participantes foram avaliadas. Ao analisá-las, notou-se que a consequência foi o fator que mais influenciou o baixo nível de tolerância dos participantes ao vivenciarem os eventos de falha. De acordo com eles, a consequência gerada por tais eventos, isto é, o impedimento de salvar um documento, acessar arquivos, retrabalho e assistir ao vídeo, foi o principal fator que influenciou sua baixa tolerância à falha.

A expectativa por solução dos participantes leigos foi o fator que mais influenciou o nível médio de tolerância. Esses participantes informaram que teriam paciência para corrigir o problema gerado pela falha e tentariam novamente até conseguirem finalizar a atividade proposta. Para os participantes intermediários, a familiaridade com a situação foi o fator que mais influenciou o nível médio e alto de tolerância, pois informaram que já estão acostumados com a ocorrência dos eventos de falha simulados no estudo e, por isso, não se importam tanto com a ocorrência dessas falhas.

Os contextos com prazo longo de finalização foram o principal fator que influenciou os participantes leigos a informarem que tolerariam o evento de falha vivenciado, visto que teriam tempo para finalizar a atividade com calma e buscar maneiras para retomar seu trabalho.

Esses resultados indicam que a percepção dos participantes não depende apenas do tipo de falha que ocorreu, pois o impacto causado por ela varia de acordo com outros fatores, como o contexto em que eles estão inseridos, a consequência gerada pela falha, a familiaridade com a situação e, também, a expectativa que eles têm para solucionar o problema.

Por fim, verificamos se os níveis de tolerância dos participantes leigos e intermediários possuíam diferença significativa. A tabela 24 mostra os resultados do teste Exato de Fisher para avaliar se o nível de conhecimento em sistemas computacionais dos participantes influenciou o nível de tolerância informado por eles. Dado nível de significância ($\alpha = 0,050$), como se pode observar na Tabela 24, a H_0 não foi rejeitada, mostrando que os níveis de tolerância apresentados pelos participantes leigos não são diferentes dos níveis informados pelos participantes intermediários.

É possível concluir que não existe diferença significativa entre os níveis de tolerância apresentados pelos participantes leigos e pelos intermediários. Para ambos os grupos, independentemente do contexto, os níveis de impacto não apresentaram diferença significativa. Os dois grupos de participantes, portanto, apresentaram as mesmas percepções em relação ao impacto que os eventos de falha causaram em sua experiência.

Tabela 24 - Resultados do teste exato de Fisher (Intermediários vs. Leigos)

		LAZER PRAZO LONGO	
Contexto prazo longo.	p-valor		0,694
	Resultado	p-valor > 0,050	(Não rejeita H ₀)
		PROFISSIONAL PRAZO LONGO	
	p-valor		0,785
	Resultado	p-valor > 0,050	(Não rejeita H ₀)
		LAZER PRAZO CURTO	
Contexto prazo curto.	p-valor		0,910
	Resultado	p-valor > 0,050	(Não rejeita H ₀)
		PROFISSIONAL PRAZO CURTO	
	p-valor		0,503
	Resultado	p-valor > 0,050	(Não rejeita H ₀)

Por fim, ao se analisarem todos esses resultados, conclui-se que o nível de tolerância depende mais do contexto no qual o participante está inserido do que o tipo de falha que ocorreu. Entretanto esses dois fatores possuem peso e devem ser considerados no momento em que se for avaliar a qualidade da experiência do usuário quanto à sua confiança no funcionamento correto do SO.

Além disso, apesar de a falha consecutiva ter sido encontrada nas justificativas apresentadas pelos participantes, notou-se que ela não é o principal fator a ser considerado. Para avaliar o impacto que as falhas consecutivas causam na experiência do usuário, bem como a média de falhas consecutivas que o usuário tolera, é necessário levar em consideração diferentes fatores, tais como o prazo estipulado para a finalização de uma atividade, o grau de consequência gerado pela falha, as informações apresentadas em sua mensagem e, também, a familiaridade com a situação vivenciada. Todos esses fatores em conjunto causam influência no impacto que a falha gera na experiência do usuário.

5.3 Q3: Qual categoria de falha de SO possui maior impacto negativo na experiência do usuário?

Para responder a essa questão, no questionário online foram identificados os cenários que causaram maior e menor impacto na qualidade da experiência dos participantes. Os resultados encontrados de acordo com cada grupo de usuários (leigos e intermediários) foram avaliados separadamente, com o intuito de investigar se, nesse caso, o nível de conhecimentos dos participantes também não influenciou o impacto causado pelos eventos de falha.

Como mencionado na Seção 3.2.4, o questionário online apresentou oito cenários, que foram divididos em dois conjuntos (Set 1 e Set 2). O primeiro (Set 1) apresentou diferentes contextos, atividades, eventos de falha e consequência, enquanto o segundo (Set 2) manteve o mesmo contexto e atividades para

todos os cenários, variando apenas os eventos de falha e suas consequências. A Tabela com as frequências observadas encontra-se no Apêndice B. As Tabelas 25 e 26 apresentam quais foram os cenários de maior e de menor impacto na qualidade da experiência dos participantes, respectivamente, de cada conjunto de cenários de falha e grupo de usuários.

Tabela 25 - Cenários de maior impacto na qualidade da experiência dos participantes (questionário online)

CONJUNTO DE CENÁRIOS	USUÁRIOS	
	LEIGOS	INTERMEDIÁRIOS
Set 1	Cenário 2	Cenário 2
Set 2	Cenário 1	Cenário 1

Tabela 26 - Cenários de menor impacto na qualidade da experiência dos participantes (questionário online)

CONJUNTO DE CENÁRIOS	USUÁRIOS	
	LEIGOS	INTERMEDIÁRIOS
Set 1	Cenário 1	Cenário 1
Set 2	Cenário 4	Cenário 4

Ambos os grupos de participantes definiram os mesmos cenários de falha como sendo os que mais e menos impactariam negativamente sua experiência caso estivessem vivenciando tais situações hipotéticas. Com base nesse resultado, pode-se concluir que o nível de conhecimento de sistemas computacionais não influenciou, também, a escolha dos cenários que mais impactaram a qualidade da experiência dos participantes. Para confirmar essa observação, aplicou-se o teste Qui-Quadrado [Lymann, 2010]. Dado nível de significância ($\alpha = 0,050$), como se pode observar na Tabela 27, a H_0 não foi rejeitada, mostrando que as respostas dos usuários leigos não foram diferentes das respostas dos intermediários.

Tabela 27 - Resultados do Qui-Quadrado (Leigos vs. Intermediários)

Maior impacto	SET 1	
	p-valor	0,168
	Resultado	p-valor > 0,050 (Não Rejeita H_0)
	SET 2	
	p-valor	0,485
	Resultado	p-valor > 0,050 (Não Rejeita H_0)
Menor impacto	SET 1	
	P-valor	0,365
	Resultado	p-valor < 0,050 (Rejeita H_0)
	SET 2	
	p-valor	0,187
	Resultado	p-valor < 0,050 (Rejeita H_0)

Com base nesse resultado (Tabela 27), é possível concluir que não existe diferença significativa entre as respostas dos participantes leigos e dos participantes intermediários. Portanto os dois grupos apresentaram as mesmas percepções em relação aos diferentes cenários de falha analisados (ver Seção 3.2.4).

De acordo com as respostas dos participantes, o fator que mais influenciou a escolha do Cenário 2 como de maior impacto no Set 1, tanto para os participantes leigos quanto para os intermediários, foi a consequência do evento de falha, ou seja, o retrabalho associado com o atraso na entrega do relatório impactaram diretamente a QoE dos participantes.

Os fatores que mais influenciaram a escolha do Cenário 1 como de maior impacto no Set 2 foram a mensagem de falha e o evento de falha. Para os participantes leigos, a mensagem de falha foi o principal fator que os fez escolher esse cenário como o mais impactante do Set 2, pois o evento de falha (VIDEO_TDR_ERROR) gera uma mensagem do tipo Blue Screen of Death (BSOD). Por outro lado, para os participantes intermediários o evento de falha foi o principal fator, uma vez que a falha ocorre de maneira inesperada e possui a característica de travamento do sistema. É importante destacar que todos os cenários do Set 2 apresentaram o mesmo contexto (Natureza da atividade = Profissional/Prazo= 15 dias) e atividade (Escrita de relatório). Além disso, nenhuma falha do Set 2 gerou atraso como consequência. Esse resultado indica uma possível interação entre os fatores evento de falha e mensagem de falha, cujos efeitos combinados prevalecem em cenários com um grau menor ou nenhum de consequência da falha em finalizar a atividade proposta dentro do prazo estipulado.

O fator que mais influenciou a escolha do Cenário 1 como de menor impacto no Set 1 foi a recuperação do sistema. Tanto os participantes leigos quanto os intermediários defenderam que a reparação do sistema foi curta e fácil para esse cenário de falha, ou seja, o sistema sofreu uma falha, porém minutos depois conseguiu se recuperar, visto que foi possível realizar a palestra poucos minutos depois da ocorrência do evento de falha.

A consequência descrita no Cenário 4 do Set 2, de acordo com os participantes, tanto leigos quanto intermediários, foi o fator que mais influenciou o menor impacto causado na QoE dos participantes. Para eles, a consequência do evento de falha não afetou gravemente a atividade que estava sendo realizada, visto que, nesse cenário, o texto que estava sendo escrito não seria perdido e o relatório seria entregue no prazo correto.

As Tabelas 28 e 29 apresentam rankings dos três principais fatores de maior e menor impacto, respectivamente, na qualidade da experiência dos participantes. Todos os fatores de influência encontrados nessa análise estão no Apêndice C. É possível observar que o principal fator de escolha dos cenários de maior e de menor impacto pelos participantes foi a consequência do evento de falha, ou seja, o efeito que o evento de falha causou na atividade que o participante estava realizando.

Tabela 28 - Ranking dos fatores de maior impacto

RANK	% DE RESPOSTAS	FATOR	DESCRIÇÃO
1º	48,500%	Consequência	A falha afeta a atividade realizada pelo usuário.
2º	14,100%	Evento de Falha	Como a falha ocorre e suas características.
3º	12,600%	Mensagem de Falha	Mensagem de falha exibida de repente (BSOD).

Tabela 29 - Ranking dos fatores de menor impacto

RANK	% DE RESPOSTAS	FATOR	DESCRIÇÃO
1º	42,500%	Consequência	A falha não afeta a atividade do usuário.
2º	39%	Recuperação	O sistema/aplicação se recupera rápida e facilmente.
3º	10,600%	Familiaridade	O fato de estar acostumado a vivenciar esse tipo de situação.

A falha da categoria de kernel causou maior impacto apenas no Set 2, visto que apresentou uma mensagem de falha mais “grosseira”, segundo os participantes, causando o receio de terem perdido todos os arquivos salvos em seu computador. Diferentemente da falha de kernel apresentada no Set 1, porque os participantes acreditam que essa falha possui uma resolução fácil e rápida, visto que, após a reinicialização do sistema, eles conseguem retomar sua atividade.

A falha de aplicação do usuário descrita no cenário 2 do Set 1 causou maior impacto na qualidade da experiência dos participantes por ter afetado a integridade dos dados, gerando retrabalho bem como mudança no prazo de entrega da atividade. Entretanto, no Set 2, a categoria de falhas de aplicação do usuário não afetou o prazo de entrega e, segundo os participantes, comparadas com as demais falhas, as consequências geradas foram mais graves.

Sendo assim, analisando esse resultado, conclui-se que, para definir qual categoria de falha de sistema operacional possui maior impacto negativo na experiência do usuário, é necessário analisar a consequência e o tipo de mensagem de falha apresentada. Esses fatores possuem grande peso no cálculo de impacto causado na qualidade da experiência do usuário ao vivenciar determinada falha.

Para testar essas hipóteses, foram realizados testes estatísticos para avaliar se uma categoria de falha causa o mesmo nível de impacto nos participantes quando é apresentada em cenários diferentes, ou seja, se

apenas a categoria de falha influencia a escolha feita pelos participantes. Para essa finalidade, foi aplicado o teste Qui-Quadrado. Dado o nível de significância ($\alpha = 0,050$), como se pode observar na Tabela 30, a H0 foi rejeitada, mostrando que há diferença significativa ao analisar a mesma categoria de falha em diferentes cenários (Tabela 30).

Tabela 30 - Resultados do Qui-Quadrado (influência das categorias de falha)

MAIOR IMPACTO	
p-valor	0,002
Resultado	p-valor < 0,050 (Rejeita H0)
MENOR IMPACTO	
p-valor	0,009
Resultado	p-valor < 0,050 (Rejeita H0)

Após analisar os resultados coletados por meio do questionário online, identificou-se qual falha causou maior impacto na qualidade da experiência dos participantes que fizeram parte do estudo de observação. A Tabela 31 apresenta a atividade que, segundo os participantes, causou maior impacto em sua experiência e, por isso, foi a pior falha que eles vivenciaram no estudo.

Tabela 31 - Evento de falha que causou maior impacto

GRUPOS	svchost.exe (EF01)	notepad.exe (EF02)	explorer.exe (EF03)	VIDEO_TDR_EROR (EFO4)
G1	6	0	14	40
G2	2	2	3	51

A falha simulada no EF04 encaixa-se na categoria de falha de kernel de SO. Nessa atividade, os participantes presenciaram uma BSOD (“Tela Azul”) no momento em que estavam assistindo a um vídeo. Para avaliar o motivo que levou os participantes a escolherem essa falha como aquela que causou maior impacto em sua experiência, as justificativas apresentadas foram avaliadas. A Tabela 32 apresenta os principais fatores de influência encontrados.

Tabela 32 - Ranking dos fatores que influenciaram a escolha pela falha EF04

RANK	% DE RESPOSTAS	FATOR	DESCRIÇÃO
1º	29,100%	Mensagem de falha	Mensagem de falha exibida de repente (BSOD).
2º	19,100%	Consequência	A falha não afeta a atividade do usuário.
3º	12,500%	Recuperação	O sistema não se recupera rapidamente.
3º	12,500%	Evento de falha	Como a falha ocorre e suas características.

Como se pode observar, a consequência e a mensagem de falha foram os fatores que, de modo geral, mais influenciaram o impacto negativo causado pela falha. Para o grupo de usuários intermediários, além desses fatores, esse tipo de falha afeta o funcionamento do sistema operacional completamente, impedindo que o usuário continue realizando seu trabalho. Outro fator que influenciou o grupo G2 foi a recuperação, pois, segundo esses participantes, os procedimentos que devem ser seguidos para resolver esse problema geralmente são demorados e difíceis de serem seguidos.

Entretanto, como se nota, o evento de falha foi o terceiro fator que causou maior influência no impacto gerado pela falha. Apesar de essa falha ter sido a que causou maior impacto negativo na experiência dos participantes, o tipo de falha não foi levado em consideração, mas sim o modo como a falha ocorreu e as suas características. Desse modo, avaliando todos esses resultados, conclui-se que determinada categoria de falha pode apresentar diferentes pesos no impacto causado na QoE do usuário, pois em determinada situação pode apresentar um grande impacto negativo; em outra, pode nem ser considerada.

Por fim, apesar de as falhas que ocorrem no núcleo (kernel) do SO geralmente interromperem a operação do sistema computacional como um todo, o impacto gerado por elas depende mais de outros fatores do que do tipo de falha que ocorreu. Logo, as falhas de SO que ocorrem fora do núcleo podem ter um impacto tão elevado quanto aquelas que ocorrem no kernel, visto que o impacto é influenciado por outros fatores, como a consequência, o contexto e a mensagem de falha apresentada.

5.4 Qualidade da experiência do usuário

Para quantificar a avaliação do usuário quanto à sua confiança no funcionamento correto do SO frente a diferentes tipos de falhas, é necessário levar em consideração os principais fatores que causam impacto em sua experiência, bem como avaliar os pesos atribuídos a esses fatores.

Este estudo, primeiramente, teve como propósito identificar quais contextos apresentaram influência na experiência do usuário. Foi possível observar que os participantes que estavam inseridos em contextos com prazo longo, independentemente da natureza da atividade (lazer ou profissional), possuíam baixo nível de tolerância. Entretanto a percepção de confiabilidade dos participantes atribuída ao sistema

operacional não depende apenas desses fatores, precisando ser avaliada em conjunto com a consequência gerada pela falha.

Desse modo, para identificar a avaliação dos participantes quanto ao serviço oferecido pelo sistema operacional após a ocorrência da falha, bem como sua experiência ao presenciar diferentes falhas de SO, eles foram questionados quanto à sua confiança e avaliação do serviço oferecido.

A Figura 34 mostra a avaliação dos participantes dos cenários profissional com prazo longo e lazer com prazo curto. É possível observar que a avaliação da qualidade do serviço informada pelos participantes do G1 variou de ruim para regular. Entretanto o grupo de usuários G2 apresentou boa avaliação de qualidade atribuída ao serviço oferecido pelo SO.

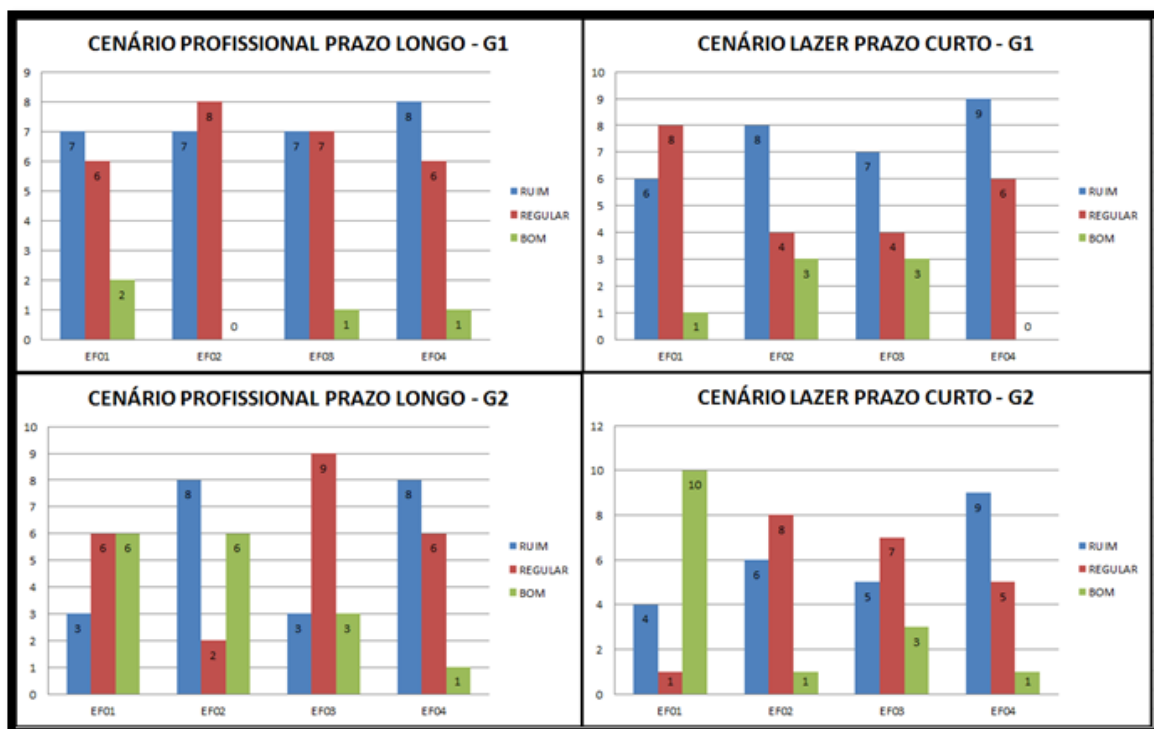


Figura 33 - Distribuição da avaliação dos serviços do SO de acordo com cada falha apresentada

Ao avaliar as justificativas apresentadas pelos participantes, foi possível observar que o principal fator avaliado é a mensagem de falha. Como se pode observar na Figura 34, a avaliação de serviço atribuída pelos participantes do grupo G2 variou entre bom e regular. Essa variação ocorreu, pois eles avaliaram o serviço de acordo com a eficiência do sistema em apresentar mensagens técnicas capazes de auxiliá-los a procurar uma solução.

Segundo os participantes desta pesquisa, quando a mensagem de falha apresenta informações que auxiliam a busca por uma solução, esse fator causa um impacto positivo na experiência do usuário ao utilizar o computador. Eles acreditam, porém, que, em algumas situações, as informações apresentadas nas mensagens são muito genéricas para uma pessoa que não possui um alto conhecimento em TI.

Ao analisar os dados, observou-se que, para os participantes do grupo de usuários leigos, as informações apresentadas na mensagem não auxiliam o usuário a procurar uma solução para o problema. Eles citaram “A mensagem deveria apresentar outro tipo de solução. Esse tipo de mensagem apresentada pelo SO é um descaso com o usuário e me deixa apreensivo”. Desse modo, é possível notar que a

mensagem de falha é um fator avaliado pelo usuário levando em consideração as informações apresentadas, bem como a sua facilidade de compreensão.

Além de analisar a qualidade do serviço, no final do estudo os participantes avaliaram sua experiência após vivenciarem diferentes cenários de falha (Figura 34).

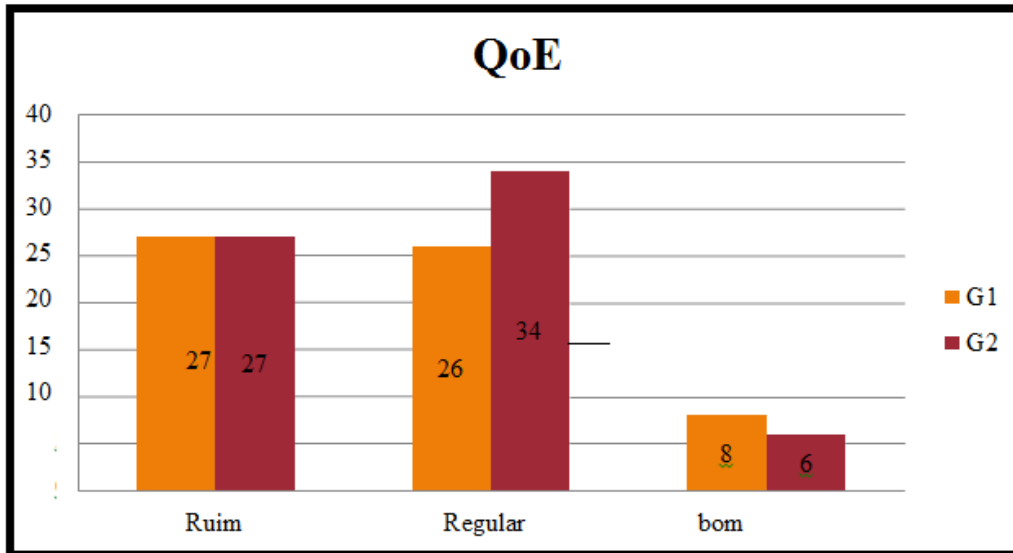


Figura 34 - Avaliação da qualidade da experiência do usuário

A Figura 34 apresenta a distribuição das respostas informadas pelos participantes quanto à avaliação de sua experiência ao presenciarem diferentes falhas do SO. É possível observar que as respostas informadas pelos usuários do grupo G1 e G2 possuem pouca diferença e estão bem distribuídas entre si. Os principais fatores de influência estão apresentados na Tabela 33. Como se pode observar, os fatores de influência se mantiveram, mostrando que o contexto, a consequência e a

expectativa de solução foram os fatores que mais influenciaram a avaliação dos usuários. Além disso, foi possível observar que muitos participantes avaliaram a QoE levando em consideração sua familiaridade com as situações vivenciadas. Esses participantes informaram que, independentemente de a falha ter ocorrido, na maioria das vezes o sistema auxilia-os em suas atividades do dia a dia. Então, de modo geral, ele é útil.

Tabela 33 - Ranking dos fatores que impactaram a QoE dos participantes

RANK	% DE RESPOSTAS	FATOR	DESCRIÇÃO
1º	25%	Consequência	A falha afeta a atividade do usuário.
1º	25%	Expectativa de solução	O usuário possui expectativa para solucionar o problema.
2º	10,600%	Contexto	Natureza da atividade e prazo para conclusão da atividade.
2º	10,600%	Falhas consecutivas	Quantidade de vezes em que a falha ocorre.
3º	8,300%	Familiaridade	Os usuários estão acostumados a vivenciar

Nota-se que as falhas consecutivas foram um fator que influenciou a avaliação dos participantes quanto à sua QoE ao vivenciarem os cenários apresentados no estudo. Contudo, comparadas com os demais fatores, as falhas consecutivas tiveram menor impacto na experiência dos usuários. Desse modo, conclui-se que, para identificar a quantidade média de falhas consecutivas que o usuário tolera, é necessário levar em consideração os demais fatores citados neste capítulo, tais como contexto, consequência, evento de falha e mensagem de falha.

Ao questionar os participantes sobre sua confiança no sistema operacional após a ocorrência das falhas, as respostas também foram bem distribuídas, não apresentando uma diferença significativa (Figura 35). Foi possível observar, no entanto, que os participantes que informaram confiar no sistema, mesmo após a ocorrência de falhas, não levaram em consideração a sua confiança ou segurança em continuar utilizando o SO, mas o fato de já estarem acostumados com a ocorrência da falha.

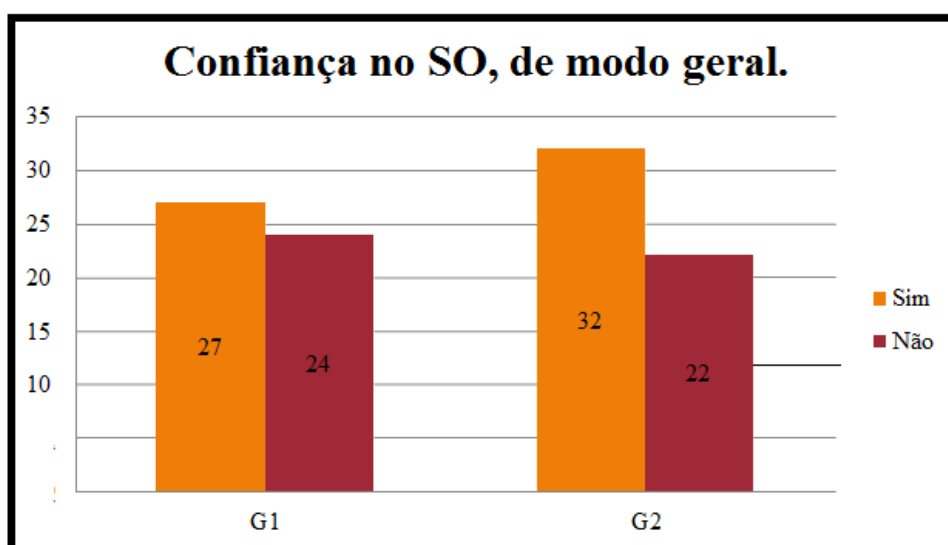


Figura 35 - Avaliação da confiança do usuário

Sendo assim, conclui-se que os participantes atribuem a sua confiança de acordo com a importância do sistema operacional para o seu dia a dia, bem como da sua familiaridade com as situações vivenciadas.

Por fim, tem-se como resultado que o contexto e a consequência foram os principais fatores que influenciaram os níveis de tolerância à falha e confiança dos participantes atribuída ao SO. Desse modo, pode-se concluir que esses são os principais pontos que devem ser considerados para quantificar a qualidade da experiência do usuário quanto à sua confiança no funcionamento do sistema operacional.

Além disso, as habilidades dos usuários não são um fator que influencia o impacto que a falha causa em sua QoE. Fatores humanos, tais como expectativa de solução e familiaridade com os eventos de falha que ocorreram, afetam positivamente a QoE dos usuários.

Resultados

6.1 Síntese dos resultados

Esta pesquisa teve como objetivo principal avaliar a qualidade da experiência do usuário quanto à sua confiança no funcionamento do sistema operacional. Para tanto, foi realizado um estudo de observação. Primeiramente, foi aplicado um questionário online, que permitiu identificar os principais eventos de falha que causam impacto negativo na experiência do usuário e, assim, criou-se o questionário presencial aplicado durante a realização de um estudo de observação em laboratório. Os dados coletados durante essas sessões permitiram avaliar os principais fatores que influenciaram o nível de tolerância, de confiança e o impacto causado pela ocorrência dos eventos de falha na QoE dos usuários.

Os resultados mostram que os usuários, ao presenciarem determinada falha, atribuem a causa ao sistema operacional e, por isso, acreditam que essa falha afete o funcionamento correto do SO, independentemente de ele ser ou não a principal fonte da falha ocorrida. Essas informações validaram a hipótese de que a maioria dos usuários não é capaz de diferenciar uma falha da aplicação de uma falha do SO.

Ao se analisarem os resultados, observou-se, também, que o tipo de falha não teve influência significativa na QoE percebida pelo usuário. Outros fatores foram mais influentes, tais como consequência, contexto (prazo), mensagem de falha, recuperação, expectativa de solução, o evento de falha e a sua frequência. Os três principais fatores que impactaram a QoE dos participantes foram: consequência, contexto e evento de falha, nessa ordem de importância. Desse modo, conclui-se que o nível de tolerância do usuário depende mais do contexto em que o usuário se encontra do que do tipo de falha que ocorreu.

Dada a influência do contexto na experiência do usuário, percebeu-se que o prazo (curto ou longo) foi mais influente do que a natureza da atividade (profissional ou lazer); os efeitos dessa última característica não mostraram significância estatística na QoE do usuário para os cenários estudados.

Além desses fatores, já planejados para serem investigados, os resultados revelaram que outros fatores de influência técnica e não técnica eram importantes. Por exemplo, a expectativa é um fator de influência humana cujos efeitos impactaram a QoE do usuário neste estudo, pois os dados mostraram que os participantes tinham expectativa de solução para corrigir a falha vivenciada. Da mesma forma, o tempo e o esforço necessários para se recuperar de um evento de falha também influenciaram a QoE do usuário.

Foi possível observar que, uma vez que o evento de falha foi vivenciado, seus efeitos na QoE do usuário podem ser atenuados se a recuperação for rápida e fácil de executar sem maiores consequências. Por exemplo, uma falha no kernel do

SO geralmente faz com que o sistema pare. Se, após a reinicialização do SO, o sistema estiver pronto para ser utilizado rapidamente e sem grandes consequências (ex.: perda de dados), o nível de impacto dessa falha na QoE poderá ser considerado baixo. Contudo, para diferenciar os eventos de falha, é necessário considerar outros fatores além do tipo de falha, a fim de fornecer uma estimativa de confiabilidade mais precisa do ponto de vista do usuário.

Além disso, os resultados mostraram que o nível de conhecimento dos participantes não é um fator que tenha influenciado o impacto causado pela falha, visto que eles obtiveram um nível de tolerância equivalente ao vivenciarem diferentes falhas de SO. Devido à dificuldade de encontrar

voluntários para participar da pesquisa, esse resultado ajudará trabalhos futuros, uma vez que os participantes poderão ser recrutados independentemente de suas habilidades.

Uma categoria de falhas de SO pode possuir falhas com diferentes pesos de impacto na experiência do usuário, pois esse impacto é diretamente influenciado pelo contexto e, principalmente, pela consequência gerada pela falha. Desse modo, os dados mostraram que as falhas de SO que ocorrem fora do núcleo (em aplicações e serviços do SO) têm impacto tão elevado ou até mais que as falhas em nível de kernel, pois o tipo de falha que ocorreu tem menos influência do que outros fatores acima mencionados.

Dado que as falhas de SO afetam diretamente a qualidade da execução das aplicações de software, compreender como elas são percebidas pelos usuários com respeito à qualidade da sua experiência frente a tais eventos vem contribuir para um melhor entendimento de como o usuário lida com a confiança no funcionamento dos sistemas computacionais atuais. O estudo em questão apresentou os principais fatores que causam impacto na QoE dos usuários após a ocorrência de eventos de falha. Esses dados mostram a importância de realizar uma análise extensiva da confiabilidade, levando em consideração tanto fatores quantitativos de confiabilidade (cálculo matemático atual) quanto aspectos qualitativos relacionados à qualidade da experiência do usuário. Todos esses resultados auxiliam no processo de desenvolvimento de sistemas mais confiáveis, contribuindo ativamente para a área de engenharia de software.

6.2 Contribuição para literatura

Os resultados deste trabalho auxiliaram na escrita do artigo intitulado “Influence Factors on the Quality of User Experience in OS Reliability: A Qualitative Experimental Study”, que foi submetido para a 13a International Conference on Availability, Reliability and Security (ARES 2018). Este artigo apresentou a influência de diferentes tipos de falha e outros fatores, tais como contexto e consequência, na percepção do usuário sobre a confiabilidade do sistema operacional (SO).

6.3 Ameaças à validade

Como nesta pesquisa foram utilizados questionários para a coleta dos dados, existem algumas limitações que precisam ser consideradas no momento em que se forem analisar os resultados.

Uma das principais restrições ao uso de questionários online é não ter garantias de que o participante tenha entendido as perguntas corretamente. Para atenuar essa ameaça, foi aplicado um questionário presencial respondido sob a supervisão dos pesquisadores.

Outra preocupação desta pesquisa foi a de diminuir fatores que influenciavam as respostas informadas pelos participantes do estudo de observação. Desse modo, as entrevistas foram realizadas em um ambiente no qual o participante não ficava em contato com outros voluntários, evitando-se outros fatores que poderiam influenciá-lo, como barulho e telefone. Além disso, existe o viés de seleção, especialmente em relação aos participantes do questionário presencial. Dada a necessidade de tê-los no ambiente de estudo, a logística de recrutamento limitou-se a três cidades próximas. Embora tenham participado indivíduos de diferentes gêneros, idades, ocupações, afiliações e cidades, sabe-se que fatores regionais ou culturais podem influenciar a QoE percebida pelo usuário.

Uma quarta ameaça é a ordem em que os eventos de falha apareceram durante o estudo de observação (ver Seção 4.3), o que pode influenciar as respostas dos participantes. Essa ameaça foi diminuída, não apresentando, no início da sessão de estudo, os eventos de falha considerados de maior impacto pelos respondentes do questionário online; esses eventos de falha foram apresentados no final do estudo.

Além das ameaças citadas anteriormente, estudos mostram que a cultura é um fator de influência que causa impacto na qualidade da experiência do usuário. Desse modo, a fim de mitigar essa ameaça, os dados foram coletados recrutando participantes de diferentes estados, empresas e religiões. Contudo, pretende-se futuramente aplicar esse estudo a usuários estrangeiros (ex.: americanos) para verificar se há diferença nos dados coletados.

A pesquisa realizada neste estudo baseou-se apenas em um sistema operacional para simulação das falhas, entretanto as questões descritas nos questionários foram genéricas. Desse modo, o estudo pode ser replicado para qualquer pesquisa que investigue a confiabilidade de sistemas operacionais, independentemente de qual SO utilize.

O software utilizado para a simulação das falhas de aplicação do usuário foi o notepad.exe, editor de texto padrão da Microsoft. Ele foi escolhido devido à facilidade de replicação do estudo, visto que a interface do software é simples. Sendo assim, outra ameaça à validade é a mensagem apresentada por essa aplicação, pois esse fator pode causar influência nas respostas apresentadas pelos participantes. Foi possível observar, contudo, por meio dos resultados, que os participantes não apresentaram conhecimento específico para informar, de maneira correta, a causa das falhas simuladas nesta pesquisa. Além disso, a mensagem de falha, apesar de ser um fator frequente de influência, não foi o fator que causou maior impacto negativo na experiência do usuário.

6.4 Dificuldades encontradas

Como não foi encontrado na literatura nenhum trabalho que tenha investigado a influência que as falhas de sistemas operacionais causam na qualidade da experiência do usuário, uma das maiores dificuldades enfrentadas no início desta pesquisa foi encontrar informações necessárias para auxiliar a coleta de dados deste estudo. Desse modo, foi necessário criar um protocolo próprio para realizar a coleta de dados.

Outra dificuldade foi encontrar voluntários para participar da etapa presencial. Para realizar a pesquisa, foi necessário conciliar os horários dos participantes de modo que eles não ficassem aguardando e, conseqüentemente, não desistissem de participar da pesquisa. Para contornar essa dificuldade, foi realizado um agendamento, respeitando o horário de cada participante e, quando necessário, os pesquisadores se locomoveram até eles.

6.5 Trabalhos futuros

Como trabalho futuro, pretende-se realizar outra coleta de dados acrescentando usuários especialistas no estudo com o propósito de verificar se o nível de tolerância e confiança desses participantes difere dos níveis encontrados nesta pesquisa. Pretende-se, também, analisar se os usuários especialistas, diferentemente dos grupos estudados nesta pesquisa, sabem diferenciar uma falha de aplicação de uma falha de SO.

Além disso, deseja-se alterar a ordem das falhas simuladas na ferramenta com o propósito de avaliar se os resultados observados neste estudo serão mantidos. Foi possível observar nos dados apresentados neste trabalho que fatores como personalidade (participante paciente, nervoso, curioso) são variáveis que influenciam o nível de tolerância bem como de confiança do usuário no SO.

Durante a coleta de dados, foi possível notar que as mensagens de falha apresentadas foram um fator que causou impacto na avaliação da experiência do usuário. Sendo assim, além de levar em consideração a personalidade dos participantes, pretende-se incluir as mensagens de falha como variável ao realizar novas sessões de estudo, com o propósito de verificar o peso que as mensagens de falha possuem na QoE.

Como foram apresentadas no Capítulo 5, as falhas consecutivas são fatores que causam influência na tolerância do usuário, independentemente do grupo em que ele se encontra. Desse modo, pretende-se investigar a quantidade média de falhas consecutivas que ele tolera e, também, como esse fator impacta a qualidade de sua experiência.

Os resultados mostraram que o tempo gasto pelo sistema operacional para se recuperar de um evento de falha pode atenuar o impacto causado pela falha na QoE do usuário. Sendo assim, pretende-se investigar o impacto causado na experiência do usuário levando em consideração fatores como o tempo em que um sistema demora para se recuperar da falha de acordo com o contexto em que ele está inserido.

Como foi observado na Seção 5.2, a maioria dos participantes não soube informar se as falhas apresentadas eram consequência do funcionamento incorreto do sistema operacional. Apesar de uma grande parcela dos participantes ter respondido de forma correta a esse questionamento, não se pode afirmar que eles acertaram por possuírem um conhecimento suficiente para diferenciar uma falha de aplicação e uma falha de SO. Desse modo, pretende-se investigar se outros fatores, além do nível de conhecimento, influenciam o participante a colocar o SO como a causa da falha apresentada.

Referências

[Abowd, 1999] Abowd, G. D., Dey, A. K., Brown, P. J., Davies, N., Smith, M., & Steggles, P. (1999, September). Towards a better understanding of context and context-awareness. In International symposium on handheld and ubiquitous computing (pp. 304-307). Springer, Berlin, Heidelberg.

https://doi.org/10.1007/3-540-48157-5_29

[AL-Badareen *et al.*, 2011] AL-BADAREEN, A. B., Selamat, M. H., Jabar, M. A., Din, J., Turaev, S., & Malaysia, S (2011). "Users' Perspective of *Software Quality*". In: The 10th WSEAS International Conference on *Software Engineering, Parallel And Distributed Systems (SEPADS 2011)*. p. 84-89.

[Andrade, 2002] ANDRADE, Maria Margarida de. "Como preparar trabalhos para cursos de pós-graduação: noções práticas". 5. ed. São Paulo: Atlas, 2002.

[ANSI/IEEE, 1991] ANSI/IEEE (1991). "Standard Glossary of *Software Engineering Terminology*". STD-729-1991, ANSI/IEEE.

[Antunes e Matinas, 2014] Antunes, M. P. e Matias, R. (2014). "Confiabilidade de Sistemas Operacionais de Propósito Geral: Um Estudo de Caso". Proc. 4th Brazilian Symposium on Computing Systems Engineering, pp. 2(1)-2(6).

[Assunção, 2016] ASSUNÇÃO, Lucas Miranda (2016). "Instrumental para Coleta de Dados sobre a Experiência do Usuário em Confiabilidade de Sistemas Operacionais". Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia – UFU.

[Avizienis *et al.*, 2004] Avizienis, Algirdas, Laprie, J. C., Randell, B., & Landwehr (2004). "Basic concepts and taxonomy of dependable and secure computing". IEEE transactions on dependable and secure computing 1.1: 11-33.]

<https://doi.org/10.1109/TDSC.2004.2>

[Barbetta, 2008] Barbetta, Pedro Alberto (2008). "Estatística aplicada às ciências sociais". Ed. UFSC.

[Bevan, 2008] Bevan, Nigel. "Classifying and selecting UX and usability measures." International Workshop on Meaningful Measures: Valid Useful User Experience Measurement. 2008

[Brunnström *et al.*, 2013] Brunnström, Kjell, Beker, S. A., De Moor, K., Dooms, A., Egger, S., Garcia, M. N. & Lawlor, B. (2013). "Qualinet white paper on definitions of quality of experience."

[Cruz, 2016] Cruz, Renato Rojas da. "Medida de experiência do usuário: elaboração de questionário com foco para websites de notícias." (2016).

[Dos Santos e Matias, 2015] Dos Santos, C.A.R. e Matias, R. (2015). An Empirical Study on Failure Causes in a Commercial Off-the-Shelf Operating System", Proc. 5th Brazilian Symp. on Computing Systems Engineering, pp. 1-6.

<https://doi.org/10.1109/SBESC.2015.8>

[Dos Santos e Matias, 2016] Dos Santos, C.A.R. e Matias, R. (2016). Exploratory Analysis on Failure Causes in a Mass-Market Operating System. ACM SIGOPS Operating Systems Review - Special Topics, vol. 50, no. 1, pp. 18-30.

<https://doi.org/10.1145/2903267.2903274>

[Dos Santos e Matias, 2017a] Dos Santos, C.A.R. e Matias, R. (2017a). Failure Patterns in Operating Systems: An Exploratory and Observational Study. Elsevier Journal of Systems and *Software*.

[Dos Santos e Matias, 2017b] Dos Santos, C.A.R. e Matias, R. (2017b). An Empirical Study on Patterns of Failure Causes in a Mass-Market Operating System. Proc. 32th ACM Symposium on Applied Computing, pp. 1542-1547.

<https://doi.org/10.1145/3019612.3019740>

[Dos Santos, 2016] SANTOS, C. A. R. dos. "Um Estudo Exploratório Sobre Padrões de Falhas de *Software* de Sistemas Operacionais". Dissertação (Mestrado), 2016.

[Filardi e Traina, 2008] FILARDI, Ana Lúcia; TRAINA, Agma Juci Machado. "Montando questionários para medir a satisfação do usuário: avaliação de interface de um sistema que utiliza técnicas de recuperação de imagens por conteúdo". In: Proceedings of the VIII

[Ganapathi *et al.*, 2006] Ganapathi, A., Ganapathi, V., e Patterson, D. (2006). "Windows XP Kernel Crash Analysis". Proc. 20th Conference on Large Installation System Administration, pp. 149-159.

[Gil, 2008] GIL, Antonio Carlos (2008). "Métodos e técnicas de pesquisa social". In: Métodos e técnicas de pesquisa social. Ed. Atlas.

[Godoy, 1995] GÓDÓY, Arlida Schmidt. "Introdução à pesquisa qualitativa e suas possibilidades". Revista de administração de empresas, 1995, 35.2: 57-63.

[Hassenzahl e Tractinsky, 2006] Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. Behaviour & information technology, 25(2), 91-97.

<https://doi.org/10.1080/01449290500330331>

[Hassenzahl, 2007] "The hedonic/pragmatic model of user experience. In: Towards a UX manifesto". BRITISH HCI GRÓUP ANNUAL CÓNFERENCE ÓN PEOPLE AND COMPUTERS: HCI... but not as we know it, 2007, Lancaster. Proceedings... British Computer Society, 2007. p. 10-14.

[ISO 9126, 1991] International Organization for Standardization, ISO DIS 9241-210 (2010). "Ergonomics of human system interaction - Part 210: Human-centered design for interactive systems" (formerly known as 13407). International Standardization Organization (ISO)

[ISO 9241, 2010] International Organization for Standardization, ISO DIS 9241-210 (2010). "Ergonomics of human system interaction - Part 210: Human-centered design for interactive systems" (formerly known as 13407). International Standardization Organization (ISO).

[Jalote *et al.*, 2004] JALOTE, Pankaj, Murphy, B., Garzia, M., Errez, B., & Way, Ó. R.. "Measuring reliability of *software* products". In: ISSRE 2004 Conference, Saint-Malo, Bretagne, France. 2004.

[Jen, 2006] Jenn, N. C. (2006). Designing a questionnaire. Malaysian family physician: the official journal of the Academy of Family Physicians of Malaysia, 1(1), 32.

[Law, 2011] Law, Effie Lai-Chong, *et al.* "Understanding, scoping and defining user experience: a survey approach." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2009

[LimeSurvey, 2017] LimeSurvey's manual, "LimeSurvey's Manual", Dec. 2017. Disponível em <<https://manual.limesurvey.org/>> Acesso em 11/01/2018.

[Lyman, 2010] R. Lyman Ott, Michael Longnecker (Ed.). 2010. An Introduction to Statistical Methods and Data Analysis (6th. ed.). Brooks/Cole, Belmont.

[Lyu, 2007] Lyu, M. (2007). "Software Reliability Engineering: A Roadmap," Proc. 29th Int. Conf. Softw. Eng., Future Softw. Eng., pp.153 -170.

<https://doi.org/10.1109/FOSE.2007.24>

[Matias *et al.*, 2013] Matias R., Oliveira G. e Araújo L. (2013). "Operating System Reliability from the Quality of Experience Viewpoint: An Exploratory Study". Proc. 28th ACM Symposium on Applied Computing, pp.1644-1649.

<https://doi.org/10.1145/2480362.2480669>

[Matias *et al.*, 2014] Matias, R., Prince, M., Borges, L., Sousa, C. e Henrique, L. (2014). "An Empirical Exploratory Study on Operating System Reliability". Proc. 29th ACM Symposium on Applied Computing, pp. 1523-1528.

<https://doi.org/10.1145/2554850.2555021>

[Merholz, 2007] MERHOLZ, P., 2007. "Peter in Conversation with Don Norman About UX & Innovation", <http://www.adaptivepath.com/ideas/e000862/>

[Murphy, 2008] Murphy, B. (2008). "Reliability Estimates for the Windows Operating System". Microsoft Research Cambridge. Disponível em: <<http://www.dcl.hpi.uni-potsdam.de/meetings/mshpsummit/slides/brendan.murphy.pdf>>. Acessado: 11/07/2018.

[NetMarketShare, 2017] NetMarketShare, "Desktop Operating System Market Share", Dec. 2017, Disponível em: <<https://www.netmarketshare.com/operating-system-market-share.aspx?id=platformsDesktopVersions>> Acesso em 11/01/2018.

[Niels *et al.*, 2017] Niels, Adelka, Tobias Lesser, and Torben Krüger. "The Impact of Causal Attributions on the User Experience of Error Messages". Advances in The Human Side of Service Engineering. Springer International Publishing, 2017. 173-184.

https://doi.org/10.1007/978-3-319-41947-3_17

[Resolução 466, 2012] Padilha, R. S. Alexandre, “Resolução Nº 466, de 12 de dezembro de 2016”, Dec. 2016, disponível em: <<http://conselho.saude.gov.br/resolucoes/2012/Reso466.pdf>>

[Richardson, 2014] RICHARDSÓN, Roberto Jarry, “Pesquisa social: métodos e técnicas”. São Paulo: Atlas, 2014.

[Rivero, 2016] Rivero, Luis Jorge Enrique. “An Approach for the Evaluation and Improvement of User Experience in Mobile Applications.” Qualificação (Doutorado), 2016.

[Roto *et al.*, 2011] Roto, Virpi; Lee, Ming; Castro, Brenda; Vermeeren, Arnold; Law, Effie; Väänänen-Vainio-Mattila, Kaisa; Hoonhout, Jettie; Obrist, Marianna. “LadderUX, Laddering made easy and reliable”. Disponível em: <www.ladderux.org.> Acesso em 11/07/2018.

[RStudio, 2017] RStudio (2017). “Why RStudio?”. Dec. 2017. Disponível em: <<https://www.rstudio.com/about/>> Acesso em 11/07/2018

[Schatz *et al.*, 2013] Schatz, R., Hoßfeld, T., Janowski, L., & Egger, S. (2013). From packets to people: quality of experience as a new measurement challenge. In Data traffic monitoring and analysis (pp. 219-263). Springer, Berlin, Heidelberg.

https://doi.org/10.1007/978-3-642-36784-7_10

[Schrepp *et al.*, 2014] Schrepp, Martin, Andreas Hinderks, and Jörg Thomaschewski. "Applying the User Experience Questionnaire (UEQ) in different evaluation scenarios." International Conference of Design, User Experience, and Usability. Springer International Publishing, 2014.

https://doi.org/10.1007/978-3-319-07668-3_37

[StatCounterGlobalStats, 2017] StatCounter Global Stats, “Óperating System Market Share Worldwide”, Dec. 2017. Disponível em <<http://gs.statcounter.com/os-version-market-share/windows/desktop/worldwide>> Acesso em 11/01/2018>

[Swift, 2003] Swift, M. M., Bershad, B. N., & Levy, H. M. (2003, October). Improving the reliability of commodity operating systems. In ACM SIGOPS operating systems review (Vol. 37, No. 5, pp. 207-222). ACM.

<https://doi.org/10.1145/1165389.945466>

[Tulio e Albert, 2013] Albert, William, and Thomas Tullis. “Measuring the user experience: collecting, analyzing, and presenting usability metrics”. 2. ed. Newnes: 2013.

[Vieira, 2000] Vieira, S. (2000). Como elaborar questionários. Editora Atlas SA.

[Xavier *et al.*, 2014] Xavier, J., Macêdo, A., Matias, R., & Borges, L. (2014, March). A survey on research in *software* reliability engineering in the last decade. In Proceedings of the 29th Annual ACM Symposium on Applied Computing (pp. 1190-1191). ACM.

<https://doi.org/10.1145/2554850.2555161>

[Xu *et al.*, 1999] Xu, J., Kalbarczyk, Z. e Iyer, R. (1999). “Networked Windows NT System Field Failure Data Analysis”. Proc. Pacific Rim International Symposium on Dependable Computing, pp. 178-185.

[Zanelli, 2002]. Zanelli, J. C. Pesquisa qualitativa em estudos da gestão de pessoas. Estudos de Psicologia, v. 7, p. 79 - 88, 2002.

6.1 ANEXO A - PARECER DO COMITÊ DE ÉTICA DA UNIVERSIDADE FEDERAL DE UBERLÂNDIA (UFU) - UBERLÂNDIA/MG



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Análise qualitativa da experiência do usuário em confiabilidade de sistemas operacionais

Pesquisador: RIVALINO MATIAS JUNIOR

Área Temática:

Versão: 2

CAAE: 64081616.4.0000.5152

Instituição Proponente: Faculdade de Computação

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 2.001.561

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

APÊNDICE A – QUESTIONÁRIO PRESENCIAL

ATIVIDADE I.

1. Em sua opinião, até o momento o sistema operacional operou corretamente?

Sim Não

2. Marque a opção que melhor representa seu grau de confiança no funcionamento desse sistema operacional até o momento.

0 - Nem um pouco

1 - Um pouco

2 - Mais ou menos

3 - Muito

4 - MUITÍSSIMO

3. Em sua opinião, essa falha ocorreu devido à um problema do software ou do hardware?

Software Hardware

Não sei responder essa questão

4. Em sua opinião, essa é uma falha específica de sistema operacional?

Sim Não

Não sei responder essa questão

5. Marque a opção que melhor representa seu grau de tolerância para essa falha após vivenciar essa situação.

0 - Nem um pouco

1 - Um pouco

2 - Mais ou menos

3 - Muito

4 - MUITÍSSIMO

• Justifique sua resposta

6. Suponha que após fechar a mensagem de falha você consiga imprimir o documento. Você acredita que esse fator influenciaria seu grau de tolerância a essa falha?

Sim Não

• Justifique a sua resposta.

7. Marque a opção que representa seu grau de confiança no funcionamento do sistema operacional nesse momento.

0 - Nem um pouco

1 - Um pouco

2 - Mais ou menos

3 - Muito

4 - MUITÍSSIMO

• Justifique sua resposta

8. Marque a opção que melhor representa seu grau de confiança no funcionamento do software Bloco de Notas que acabou de utilizar.

0 - Nem um pouco

1 - Um pouco

2 - Mais ou menos

3 - Muito

4 - MUITÍSSIMO

• Justifique sua resposta

9. Avalie o sentimento que melhor corresponde a sua experiência emocional caso vivenciasse essa situação hipotética.

• **Insatisfação:**

0 - Nem um pouco insatisfeito

1 - Pouco insatisfeito

2 - Mais ou menos insatisfeito

3 - Muito insatisfeito

4 - MUITÍSSIMO insatisfeito

• **Preocupação:**

0 - Nem um pouco preocupado

1 - Um pouco preocupado

2 - Mais ou menos preocupado

3 - Muito preocupado

4 - MUITÍSSIMO preocupado

• **Segurança:**

0 - Nem um pouco seguro

1 - Um pouco seguro

2 - Mais ou menos seguro

3 - Muito seguro

4 - MUITÍSSIMO seguro

10. Gostariamos que pensasse nessa falha que acabou de experienciar, você já vivenciou esse tipo de falha no seu dia-a-dia?

Sim Não

(CASO SUA RESPOSTA SEJA SIM, RESPONDA AS QUESTÕES 11 E 12).

11. Em quais situações ela ocorreu? Marque uma ou mais opções.

Trabalho/Ambiente profissional

Lazer/Entretenimento

Escola/Ambiente acadêmico

12. Marque a opção que melhor representa o grau de consequências que essa falha gerou.

0 - Nem um pouco

1 - Um pouco

2 - Mais ou menos

3 - Muito

4 - MUITÍSSIMO

APÊNDICE B – CENÁRIOS DE MAIOR E MENOR IMPACTO NA QoE DOS PARTICIPANTES DO QUESTIONÁRIO *ONLINE*

Tabela 34 – Cenários que causaram maior e menor impacto – Questionário *online*

MAIS INSATISFEITO – EVENTOS DE FALHA I					
Grupos	Cenário I	Cenário II	Cenário III	Cenário IV	Total
Leigos	4	32	9	16	61
Intermediários	5	37	9	9	60
Geral	9	69	18	25	121
MENOS INSATISFEITO – EVENTOS DE FALHA I					
Grupos	Cenário I	Cenário II	Cenário III	Cenário IV	Total
Leigos	37	0	23	1	61
Intermediários	36	2	16	6	60
Geral	73	2	39	7	121
MAIOR IMPACTO – SET 2					
Grupos	Cenário I	Cenário II	Cenário III	Cenário IV	Total
Leigos	37	8	14	1	60
Intermediários	28	9	13	11	61
Geral	65	17	27	12	121
MENOR IMPACTO – SET 2					
Grupos	Cenário I	Cenário II	Cenário III	Cenário IV	Total
Leigos	1	11	13	35	60
Intermediários	11	10	6	34	61
Geral	12	21	19	69	121

APÊNDICE C – FATORES DE INFLUÊNCIA

Tabela 35 – Fatores de alto impacto na QoE dos usuários

GERAL - ALTO IMPACTO		
COLOCAÇÃO	FATORES	QT D.
1º	CONSEQUÊNCIA	135
2º	EVENTO DE FALHA	68
3º	CONTEXTO	50
4º	NÃO FAMILIARIDADE	49
5º	CONTEXTO + CONSEQUÊNCIA	48
6º	MENSAGEM DE FALHA	37
7º	FREQUÊNCIA DA FALHA	30
8º	EXPECTATIVA DE SOLUÇÃO	12
9º	RECUPERAÇÃO	5
10º	OUTROS	2

Tabela 36 – Fatores de médio impacto na QoE dos usuários

GERAL - MÉDIO IMPACTO (Esse nível foi avaliado apenas no presencial)		
COLOCAÇÃO	FATORES	QTD
1º	EXPECTATIVA DE SOLUÇÃO	20
2º	CONTEXTO	17
3º	FAMILIARIDADE	13
4º	NÃO FAMILIARIDADE	12
5º	EVENTO DE FALHA	11
6º	CONSEQUÊNCIA	9
7º	MENSAGEM DE FALHA	8
8º	CONTEXTO + CONSEQUÊNCIA	7
9º	FREQUÊNCIA DA FALHA	6
10º	OUTROS	6
11º	RECUPERAÇÃO	2

Tabela 37 – Fatores de baixo impacto na QoE dos usuários

GERAL - BAIXO IMPACTO		
COLOCAÇÃO	FATORES	QT D.
1º	CONSEQUÊNCIA	60
2º	RECUPERAÇÃO	56
3º	FAMILIARIDADE	29
4º	CONTEXTO	20
5º	EXPECTATIVA DE SOLUÇÃO	17
6º	NÃO FAMILIARIDADE	7
7º	OUTROS	7
8º	ATRIBUIÇÃO EXTERNA	3
9º	FREQUÊNCIA DA FALHA	3
10º	EVENTO DE FALHA	3
11º	CONTEXTO + CONSEQUÊNCIA	2
12º	MENSAGEM DE FALHA	1

APÊNDICE D – ANOTAÇÕES REGISTRADAS DE ALGUNS PARTICIPANTES DURANTE O ESTUDO DE OBSERVAÇÃO

Tabela 38 – Anotações registradas de alguns participantes durante a aplicação do questionário presencial

PARTICIPANTE	ANOTAÇÕES
USER01	Tolerância não se alterou, mesmo após ter vivenciado algumas falhas no decorrer do estudo de observação. Precisou pensar para responder as perguntas.
USER02	Nunca tinha vivenciado os tipos de falhas do experimento.
USER03	Usuário informou que se sente afetado justamente por não ter conhecimento técnico. Ele foi extremamente lento ao digitar o texto sugerido no roteiro e achou que o SO não estava operando corretamente após primeira falha. Começou estralar os dedos e balançar a perna direita.
USER04	Usuário utiliza o Win8, mas elogiou várias vezes durante o experimento o Win7. Disse que o fato de terem tirado o botão iniciar do Win8 fez ele parar de gostar do Windows.
USER05	Demonstrou muita calma durante todo o experimento e teve um pouco de dificuldade em responder as perguntas (segundo ele).
USER06	Perdeu a confiança no sistema logo após a primeira falha.
USER14	Em vários momentos ficou colocando a mão na cabeça e mostrou estar nervoso a partir da terceira falha.
USER50	Ficou com a mão apoiada no queixo, muito pensativo.
USER43	Atendeu o telefone e, por isso, foi descartado.
USER14	Precisou sair do local do experimento e, por isso, foi descartado.
USER80	Não percebeu a primeira falha, pensou que era uma mensagem de aviso e, por isso, não soube responder as perguntas. Descartado!
USER88	Atendeu o telefone e, por isso, foi descartado.
USER105	Escutou participantes do experimento conversando sobre as perguntas no corredor e, por isso, foi descartado.
USER122	Em vários momentos ficou usando o celular e, por isso, foi descartado.
USER126	Participou do experimento dentro do seu trabalho e, na última falha, precisou se retirar da sala. Descartado!
USER158	Menor de idade. Descartado!