

Flávia Amin Barbosa\*  
Elder Rodrigues\*  
Ildevana Poltronieri\*

## RESUMO

Este trabalho tem como objetivo inserir o método de teste A/B no contexto de equipes ágeis, visando aprimorar a experiência do usuário e garantir a entrega de produtos web adaptáveis e com valor para o usuário final. Inicialmente, foi realizada uma revisão sistemática da literatura para investigar as práticas atuais de avaliação de usabilidade. Com base nas descobertas sobre o teste A/B, desenvolvemos um guia de avaliação de usabilidade que combina o teste A/B com o Kanban, projetado para orientar equipes ágeis na condução de testes de usabilidade. Essa estrutura foi então avaliada por meio de um *survey* com profissionais da indústria de software. Os resultados e recomendações finais da avaliação buscam aprimorar o guia desenvolvido e garantir a usabilidade de sistemas web em ambientes ágeis, focando na geração de resultados significativos e adaptáveis.

**Palavras-chaves:** Avaliação de usabilidade; Teste A/B; Kanban

## ABSTRACT

This work aims to integrate A/B testing into the context of agile teams to enhance user experience and ensure the delivery of adaptable and valuable web products for the end user. Initially, a systematic literature review was conducted to investigate current practices in usability evaluation. Based on the findings about A/B testing, we developed a usability evaluation guide that combines A/B testing with Kanban, designed to guide agile teams in conducting usability tests. This framework was then assessed through a survey with software industry professionals. The final results and recommendations of the evaluation seek to improve the developed guide and ensure the usability of web systems in agile environments, focusing on generating meaningful and adaptable outcomes.

**Keywords:** Usability evaluation; A/B Testing; Kanban

## 1. INTRODUÇÃO

Quando debatido sobre melhoria contínua da usabilidade em sistemas *web*, é fundamental garantir o engajamento do usuário para o sucesso dos produtos desenvolvidos. No contexto das equipes ágeis, em que a adaptação rápida, com valor agregado é fundamental, integrar métodos eficazes de avaliação de usabilidade pode ser desafiador. Neste trabalho, propomos a integração do método de teste A/B no ambiente ágil, visando aprimorar a experiência do usuário e garantir a entrega de produtos *web* adaptáveis e com valor para o usuário final.

\*Aluna do Curso de Engenharia de Software da Universidade Federal do Pampa, Alegrete, Rio Grande do Sul, Brasil  
E-mail: [flaviabarbosa.aluno@unipampa.edu.br](mailto:flaviabarbosa.aluno@unipampa.edu.br)

\*\*Orientador, Professor do Curso de Engenharia de Software da Universidade Federal do Pampa, Alegrete, Rio Grande do Sul, Brasil, E-mail: [elderrodrigues@unipampa.edu.br](mailto:elderrodrigues@unipampa.edu.br)

\*\*Coorientadora, Professora do Curso de Engenharia de Software da Universidade Federal do Pampa, Alegrete, Rio Grande do Sul, Brasil, E-mail: [ildevanarodrigues@unipampa.edu.br](mailto:ildevanarodrigues@unipampa.edu.br)

Segundo o estudo de Dobrigkeit, Paula et al. (2017) as equipes ágeis estão em constante aprimoramento para melhorar a forma como os produtos de software são desenvolvidos. Utiliza-se uma abordagem centrada no cliente, usuário e na entrega iterativa que tem sido amplamente debatida na indústria de software. No entanto, ainda existe um "gargalo" na integração de métodos de avaliação de usabilidade eficazes dentro dessas equipes, como é abordado no trabalho de Silva e Silveira (2010), a abordagem nem sempre é direta, pois testes tradicionais e trabalhosos muitas vezes não se encaixam no ritmo acelerado das metodologias ágeis, dificultando a avaliação e aprimoramento contínuo da usabilidade dos sistemas *web*.

A condução da avaliação de usabilidade em sistemas *web* apresenta alguns desafios, eles podem ser identificados na dificuldade de coletar dados, interpretá-los e posteriormente aplicar os resultados em forma de melhorias no sistema. Outro desafio é que as equipes ágeis apresentam uma característica que é a necessidade de adaptação rápida e contínua, e por isso, necessitam de métodos flexíveis de avaliação de usabilidade. Considerando isso, o método de teste A/B surge como um tipo de avaliação promissora, pois permite a experimentação controlada e avaliações que podem gerar rápidos resultados e mudanças em um ambiente de desenvolvimento ágil. Pensando em otimizar a organização de testes A/B dentro do contexto de equipes ágeis, sugere-se uma estrutura de integração da avaliação de usabilidade utilizando teste A/B com o quadro *Kanban*, essa estrutura busca clarificar e dividir as fases e tarefas a serem realizadas na avaliação.

Este trabalho tem como objetivo principal propor e avaliar a integração da avaliação de usabilidade com teste A/B no contexto de equipes ágeis, para melhorar a usabilidade de sistemas *web*. Os objetivos específicos são: a) Desenvolver um guia prático de avaliação de usabilidade com teste A/B; b) Realizar a avaliação do guia por meio de um *survey* de profissionais da indústria de software para que possam contribuir com suas experiências profissionais empíricas e não apenas teóricas. c) Adequar o guia conforme os resultados do *survey* realizado com os profissionais de equipes ágeis que desenvolvem sistemas *web*.

Este trabalho está estruturado da seguinte forma: na Seção 1, apresentamos uma visão geral do problema, dos objetivos do trabalho e da metodologia adotada. Na Seção 2, apresentamos a fundamentação teórica do trabalho. Na Seção 3, abordamos uma breve discussão sobre as limitações que as empresas de software enfrentam para realizar avaliações focadas na usabilidade do sistema. Na Seção 4, detalhamos o desenvolvimento do guia de avaliação de usabilidade utilizando teste A/B. Na Seção 5, é realizada a avaliação por *survey* e adequações no guia. Por fim, na Seção 6, discorremos sobre as conclusões do trabalhos, contribuições e possíveis direções futuras.

## 1.1. METODOLOGIA

A metodologia deste trabalho está dividida em quatro etapas principais: Revisão da literatura, Desenvolvimento do guia de avaliação de usabilidade utilizando teste A/B no contexto de equipes ágeis, Avaliação do guia por meio de um *survey* e Adequações no guia.

A primeira etapa consistiu na realização de uma revisão sistemática da literatura (RSL) para obter informações sobre a condução atual das avaliações de usabilidade em sistemas *web*. Foram analisados trabalhos publicados que descrevem métodos, práticas e desafios associados à avaliação de usabilidade. Esta revisão permitiu identificar lacunas e oportunidades para aprimorar a integração da avaliação de usabilidade, focando nos trabalhos relacionados ao uso de testes A/B.

Na segunda etapa, com base nas descobertas da RSL, foi desenvolvido um guia de teste A/B integrado ao *Kanban*, visando capacitar equipes ágeis na execução eficaz de

testes de usabilidade em projetos de sistemas web. A integração do método A/B com a metodologia *Kanban* proporciona uma estrutura clara para a condução da avaliação, otimizando a coleta e análise de dados. O guia não apenas facilita a execução prática das avaliações, mas também aprimora a usabilidade de sistemas web em ambientes ágeis, garantindo resultados adaptáveis ao longo do ciclo de desenvolvimento.

A terceira etapa envolveu a avaliação do guia desenvolvido por meio de um *survey*. Profissionais da indústria de *software*, como *product owners*, *scrum masters*, analistas de qualidade, desenvolvedores e *designers*, receberam o material para leitura e responderam a um questionário. Este *survey* foi utilizado para captar percepções sobre a utilidade, facilidade de uso, aplicabilidade e potenciais melhorias do guia. As contribuições coletadas serviram como base para ajustes e refinamentos do material, assegurando que ele atenda às necessidades práticas da indústria de software ágil.

Por fim, a quarta etapa consistiu nas adequações no guia, baseadas no *feedback* obtido através do *survey*. Este processo de *feedback* não apenas avaliou a relevância do guia, mas também ofereceu *insights* sobre os desafios enfrentados pelos profissionais da indústria de software ao utilizarem o material proposto. As melhorias foram implementadas com o objetivo de aprimorar a facilidade de uso e a utilidade percebida do guia, garantindo sua eficácia e aplicabilidade em ambientes ágeis.

A metodologia delineada fornece um roteiro objetivo para o desenvolvimento, implementação e avaliação de uma estrutura de avaliação de usabilidade adaptável em equipes ágeis e empresas industriais com pouca experiência na área. Este processo permite identificar desafios, avaliar soluções e fornecer orientações para aprimorar a usabilidade de sistemas web.

## **2. FUNDAMENTAÇÃO**

Nesta Seção serão apresentados os principais conceitos teóricos que compõem o escopo do trabalho.

### **2.1. USABILIDADE EM INTERAÇÃO HUMANO-COMPUTADOR (IHC)**

De acordo com Barbosa et al. (2021), usabilidade é o aspecto de qualidade mais conhecido em IHC, relacionado à facilidade de aprendizado e satisfação do usuário ao usar a interface Nielsen (1994). A usabilidade envolve como o uso do sistema é afetado pelas características do usuário, incluindo cognição e capacidade de interação, abrangendo também os atos e sentimentos dos usuários.

Para a usabilidade, a qualidade de uso é fundamental, com métricas como eficácia, que avalia se o usuário interage corretamente com o sistema, e eficiência, que mede os recursos necessários para alcançar objetivos. A satisfação do usuário é igualmente importante, afetando a fidelidade ao produto Nielsen (1994). Nielsen lista cinco fatores de usabilidade: facilidade de aprendizado, facilidade de recordação, eficiência, segurança no uso e satisfação do usuário. É raro encontrar todos esses critérios em um único sistema devido às suas complexidades.

Não existe uma "receita pronta" para interação com todos os sistemas, pois cada software possui sua individualidade. A facilidade de aprendizado considera tempo e esforço despendidos para alcançar níveis de competência Boucinha e Tarouco (2013). A facilidade de recordação refere-se ao esforço cognitivo necessário para lembrar como interagir com a interface Falcão et al. (2018). Interfaces bem estruturadas ajudam a evitar confusão e facilitar a lembrança.

A segurança no uso garante que o usuário possa explorar funcionalidades sem medo de errar, envolvendo procedimentos para evitar e recuperar-se de problemas Barbosa e Silva (2010). A satisfação do usuário, um critério subjetivo, considera sentimentos e emoções durante o uso do sistema Winckler e Pimenta (2002). É crucial conhecer as necessidades do usuário para priorizar critérios de usabilidade adequadamente.

A avaliação de usabilidade é dividida em três categorias: Investigação, Inspeção e Observação Barbosa et al. (2021).

O método de avaliação através de investigação interpreta expectativas e reações dos usuários a sistemas interativos, útil no início do desenvolvimento do design Barbosa et al. (2021). As técnicas incluem entrevistas, questionários, surveys e grupos focais.

As entrevistas, estruturadas, não estruturadas ou semiestruturadas, são usadas para eliciação de requisitos, compreendendo opiniões, comportamentos e experiências dos usuários Barbosa et al. (2021). Os questionários permitem a coleta de dados de muitos indivíduos, sendo preferíveis perguntas fechadas para facilitar o preenchimento e análise. OS *Surveys* coletam informações sobre uma área específica através de questionários, entrevistas ou observações, com análise posterior usando técnicas estatísticas Wohlin et al. (2012).

A avaliação por inspeção é realizado por especialistas em IHC, identifica problemas que o usuário possa encontrar e compara designs alternativos Barbosa et al. (2021). Os métodos principais são Avaliação Heurística, Percurso Cognitivo e Inspeção Semiótica.

Barbosa et al. (2021) explica que a avaliação através da observação fornece dados sobre situações em que os usuários realizam suas atividades, com ou sem apoio de sistemas interativos. Através do registro dos dados observados, esses métodos permitem identificar problemas reais que os usuários enfrentaram durante sua experiência de uso do sistema sendo avaliado. O avaliador pode observar os usuários em contexto ou em laboratório. A observação em contexto permite coletar uma gama mais ampla de dados mais ricos sobre a atuação dos usuários em seu ambiente de atividade. Já a observação em laboratório costuma ser mais direcionada e simples, pois o avaliador tem controle sobre o ambiente. OS principais métodos de avaliação por observação são: Estudo de campo, teste de usabilidade e avaliação de comunicabilidade.

## 2.2. REVISÃO SISTEMÁTICA DE LITERATURA

No início desta pesquisa realizou-se uma revisão sistemática de literatura para analisar os métodos de avaliação de usabilidade utilizados para ferramentas *web*. O estudo foi conduzido em várias bases de dados digitais, como *IEEE Xplore*, *SCOPUS*, *SpringerLink*, *ACM Digital Library* e *ScienceDirect*, resultando em 123 estudos selecionados após um rigoroso processo de inclusão e exclusão.

A pesquisa foi estruturada em três fases principais: planejamento, execução e análise. O planejamento incluiu a definição do escopo, objetivos, questões de pesquisa e critérios de inclusão e exclusão. As questões de pesquisa focaram nos métodos de usabilidade, sua aplicação, problemas encontrados e validação empírica.

Durante a execução, uma *string* de busca genérica foi criada e adaptada para cada base de dados. Os critérios de inclusão e exclusão foram aplicados, resultando em 1.633 artigos inicialmente, reduzidos a 260 após a aplicação dos critérios. A avaliação de qualidade final resultou em 123 artigos aceitos para a análise detalhada.

Foram identificados 18 métodos de avaliação de usabilidade, com os mais comuns sendo questionários, avaliação heurística, *think aloud*, entrevistas e teste A/B. As limitações

dos métodos incluíram falta de clareza na definição dos métodos e desafios na aplicação prática. Além disso, a maioria dos estudos não relatou as métricas de usabilidade utilizadas de maneira clara, com apenas alguns focando em eficiência, eficácia e satisfação conforme a norma ISO 9241-11.

A revisão destacou que, embora existam definições robustas de experiência do usuário e usabilidade, as pessoas frequentemente encontram dificuldades em entender e utilizar corretamente os métodos de avaliação. Portanto, a criação de um guia acessível e de fácil leitura é essencial para melhorar a aplicação correta desses métodos, refletindo as diversas dimensões e técnicas de avaliação. A pesquisa também incentivou a continuidade das discussões e investigações na área para aprimorar os métodos e abordagens de avaliação.

A revisão sistemática completa pode ser consultada no repositório do projeto <sup>1</sup>.

### **2.3. TESTE A/B**

De acordo com o trabalho de Silva (2023) o teste A/B é um método de avaliação onde duas versões de um produto são comparadas para ver qual delas se sai melhor. Os usuários são divididos em dois grupos, cada um usando uma versão diferente, e os resultados são analisados para determinar qual versão oferece a melhor experiência.

Essa Seção explora a literatura específica sobre testes A/B encontrada na revisão sistemática de literatura. Dentre os artigos encontrados, dois destacam processos de teste A/B. Foi utilizado o material destes trabalhos para verificar o estado da arte que compõe esse tipo de teste e com isso se conseguiu estruturar o processo proposto nesse trabalho. Os artigos são Firmenich et al. (2019), e Speicher, Both e Gaedke (2014), discorre-se sobre eles a seguir.

No artigo de Firmenich et al. (2019) é apresentado o método de teste comparativo para resolver problemas de usabilidade através de um processo de teste de usabilidade com o método A/B para avaliar qualquer tipo de sistema web.

Observa-se que inicialmente, foi definido quais funcionalidades seriam avaliadas durante todo o teste, logo após foi executado o primeiro processo sobre a versão inicial, essa versão é chamada de Controle. Neste processo a funcionalidade foi decomposta em pequenas funções da tela, em que avaliou-se o local de botões, menus, posição dos componentes e cores da tela. Depois de categorizar as funções por tela, cada uma foi vinculada à execução de tarefas específicas, a fim de criar cenários de teste para o software.

Logo após, avaliou-se que, para que o teste seja realizado, é de grande importância escolher as métricas que serão consideradas para avaliar a funcionalidade proposta. Deve ser determinado o tempo que irá levar cada tarefa do teste e a quantificação das respostas em relação ao sistema e quais ferramentas serão utilizadas. Outro ponto importante segundo o artigo é o recrutamento de participantes para o teste. Não deve-se esquecer que é fundamental que os participantes preencham um termo de consentimento para autorizar a execução do teste e também é preciso avaliar se os participantes estão qualificados para realizar o teste sem qualquer viés.

Após o processo de extração, criou-se ações que envolvem a descrição da execução de tarefas guiadas, de acordo com as funcionalidades presentes no protótipo, como exemplo, as tarefas: (a) "Criar uma nova conta de usuário", (b) "Buscar produtos", (c) "Adicionar produtos ao carrinho" e (d) "Finalizar a compra". Quando o teste finalizado, solicita-se aos usuários que preencham um questionário de usabilidade para coletar informações e *feedbacks*. Deste modo, pode-se identificar problemas de usabilidade, como o abandono de

---

<sup>1</sup>Revisão Sistemática de Literatura sobre Métodos de Avaliação de Usabilidade em Ferramentas Web

um formulário, elementos não descritivos, com esses resultados é feita uma análise dos problemas indicados em cada funcionalidade testada.

Após a análise dos resultados do primeiro teste (Controle), são definidas alterações no software que está sendo testado, deste modo a segunda versão é gerada para avaliar se esta (Variação) é superior à primeira versão. Replica-se então o teste e, desta vez, uma nova avaliação dos resultados é realizada para comparar as duas versões e escolher a melhor versão gerada. Com a decisão da melhor versão, é elaborada uma especificação de projeto para implementar o melhor cenário, que será implantado no ambiente de produção.

No processo do teste, são envolvidos três atores específicos: 1) Especialista em design: Responsável pela elaboração e execução do teste A/B. 2) Participantes: Que preenchem informações demográficas, executam cenários de teste seguindo passos guiados e respondem ao questionário de usabilidade. 3) Especialista projetista: Encarregado de criar uma nova versão e desenvolver a solução para o ambiente de produção.

Para a condução do teste A/B, o autor cita o software chamado Visual IA, que facilita a geração de cenários de uso por meio de seus módulos. Segundo o artigo, esse software simplifica a análise dos testes, permitindo a alteração das telas do sistema e a disponibilização de versões (A/B) para avaliação dos usuários, também registra informações sobre o uso e os tempos necessários para a realização das tarefas.

No artigo de Speicher, Both e Gaedke (2014), foi desenvolvido o teste de divisão baseado em Usabilidade, uma abordagem de teste A/B que leva em consideração as interações dos usuários para derivar métricas quantitativas de usabilidade. Os autores criaram a ferramenta WaPPU, um serviço com repositório central que permite aos desenvolvedores criar e monitorar projetos de teste A/B. Na ferramenta, é possível escolher entre 27 recursos de interação predefinidos, como cliques, comprimento do cursor e quantidade de rolagem. Além disso, ela é capaz de detectar diferenças previstas na usabilidade com significância estatística para o grupo de usuários mais representativo.

O artigo tem como objetivo avaliar a usabilidade de uma interface de busca de um site, comparando a versão padrão da página de resultados de busca (SERP) com uma versão redesenhada da mesma página. O teste foi conduzido usando a ferramenta WaPPU para rastrear a interação dos usuários com ambas as versões da SERP.

Para avaliar o site, foi aplicado um questionário de avaliação de usabilidade, além do rastreamento da interação, como a barra de navegação, que foi isolada do restante da página como um componente separado.

O estudo de usuário remoto assíncrono envolveu participantes recrutados por meio de listas de e-mail da empresa e estabeleceu tarefas semi-estruturadas, simulando um cenário comum, como "Encontre um presente de aniversário para um bom amigo que não custe mais de 50 euros". Cada participante foi aleatoriamente designado para uma das duas interfaces da SERP, a original ou a redesenhada por especialistas em usabilidade.

Antes do teste A/B, dados demográficos dos participantes foram coletados para identificar o perfil do usuário. Além disso, os participantes foram instruídos a desativar bloqueadores de anúncios (AD Blocks) para padronizar o contexto em relação aos tamanhos de tela.

O teste A/B foi conduzido com o lançamento da página a ser avaliada. Durante o teste de usabilidade, o comportamento dos usuários nas interfaces foi acompanhado por meio do WaPPU, registrando a interação com os componentes da página.

Ao concluir as tarefas, os participantes avaliaram a usabilidade da página usando o

questionário disponibilizado pelo WaPPU. O tempo gasto para realizar cada tarefa foi medido para avaliar a eficiência do usuário, e um botão de encerramento controlado garantiu a coleta completa de dados.

Os resultados foram analisados usando o questionário para avaliar a usabilidade percebida pelos participantes. Além disso, os dados de interação foram coletados e analisados com o WaPPU, permitindo avaliar o desempenho dos usuários, incluindo o tempo gasto em cada tarefa, ações realizadas e taxa de sucesso.

Ao combinar as avaliações de usabilidade percebida e desempenho, os pesquisadores conseguiram determinar a eficácia das interfaces A e B, identificando pontos de melhoria para criar uma nova versão com base nas experiências coletadas, com o apoio de profissionais de *UX*.

Os resultados do artigo sugerem que essa abordagem é eficaz para detectar diferenças na usabilidade da interface e treinar modelos de qualidade, mesmo com um conjunto limitado de dados. No entanto, o artigo não fornece detalhes sobre a análise de dados nem sobre as métricas de usabilidade utilizadas para avaliar as versões da interface.

## **2.4. METODOLOGIAS ÁGEIS**

### **2.4.1. EQUIPES ÁGEIS**

Equipes ágeis representam uma abordagem de desenvolvimento de software que enfatiza a flexibilidade, colaboração e entrega contínua de valor ao cliente. O estudo de Nagai e Sbragia (2023) relata que o conceito central das equipes ágeis é derivado dos princípios estabelecidos no Manifesto Ágil, um documento fundador que destaca a importância de indivíduos e interações, software funcionando, colaboração com o cliente e resposta à mudança. Em uma equipe ágil, o trabalho é frequentemente organizado em iterações chamadas "*sprints*", geralmente de duas a quatro semanas de duração. Essas iterações visam criar incrementos funcionais do produto, permitindo que a equipe responda rapidamente às mudanças nos requisitos do cliente. A comunicação aberta, a auto-organização e a adaptação contínua são fundamentais para o sucesso das equipes ágeis.

### **2.4.2. KANBAN**

De acordo com Anderson e Carmichael (2016) o *Kanban* é uma metodologia visual para gerenciamento de processos originada no sistema de produção da Toyota e posteriormente adaptada para contextos de desenvolvimento de *software*. Ao contrário de abordagens mais prescritivas, o *Kanban* oferece uma maneira flexível de otimizar a eficiência e a produtividade, permitindo que as equipes visualizem seu trabalho em um quadro. Esse quadro é dividido em colunas representando os estágios do fluxo de trabalho, e as tarefas são representadas por cartões que se movem entre as colunas à medida que o trabalho progride. O *Kanban* enfatiza a limitação do trabalho em progresso (*WIP*) para evitar sobrecarregar a equipe, focando na entrega contínua de valor. A capacidade de adaptar rapidamente o fluxo de trabalho às mudanças nas prioridades e demandas do cliente é uma característica central do *Kanban*, tornando-o uma ferramenta valiosa para equipes que buscam otimizar a eficiência e a qualidade do processo de desenvolvimento de *software*.

## **3. O QUE AS EMPRESAS DE SOFTWARE PRECISAM PARA REALIZAR TESTE DE USABILIDADE**

De acordo com Santos (2016) o número de empresas de *software* no mundo é significativo e está em constante crescimento. Estima-se que existam milhões de empresas de

software em todo o mundo, variando de pequenas *startups* a gigantes da tecnologia. O Brasil tem um cenário de tecnologia da informação e *software* em rápido crescimento. Grandes cidades como São Paulo, Rio de Janeiro, Belo Horizonte e Florianópolis são conhecidas por abrigar muitas empresas de *software*.

A preocupação com a avaliação de usabilidade em aplicações de *software* é uma tendência crescente em todo o mundo. De acordo com Vaz et al. (2008) empresas de desenvolvimento de *software* estão reconhecendo a importância de garantir que seus produtos sejam fáceis de usar e atendam às necessidades dos usuários. Métodos de avaliação de usabilidade, como testes de usabilidade, grupos focais e análise heurística, são comuns na indústria de *software*. Muitas empresas têm equipes ou profissionais dedicados à pesquisa e ao design de experiência do usuário (UX). A importância da usabilidade é impulsionada pelo desejo de oferecer produtos de alta qualidade, melhorar a satisfação do cliente e garantir uma vantagem competitiva no mercado.

A escolha de um método de avaliação de usabilidade por parte de empresas de *software* depende de vários fatores. Primeiramente, é fundamental considerar os objetivos do projeto, já que alguns métodos são mais adequados para avaliar aspectos gerais de usabilidade, enquanto outros são melhores para identificar problemas específicos. As empresas precisam decidir qual método atende melhor às suas expectativas.

Outro fator crucial é a fase do desenvolvimento em que a avaliação de usabilidade é realizada. Métodos como a avaliação heurística são eficazes na fase de prototipação de telas do sistema, enquanto os testes de usabilidade são mais apropriados quando o produto está em uma fase mais avançada. Identificar corretamente a fase do projeto e escolher o método de avaliação mais indicado é essencial para obter resultados relevantes.

Os recursos disponíveis, como tempo e orçamento, também influenciam a escolha do método de avaliação de usabilidade. Alguns métodos exigem mais recursos do que outros, então é importante que as empresas analisem se possuem todos os recursos necessários para a execução da avaliação escolhida. Caso contrário, podem optar por métodos que demandem menos tempo, recursos financeiros e pessoas envolvidas.

A compreensão dos usuários-alvo é outro aspecto crucial. Diferentes métodos podem ser mais apropriados para coletar informações sobre as percepções e necessidades dos usuários, dependendo de fatores como nível de instrução, habilidades com o uso de tecnologia e faixa etária. Identificar o grupo-alvo e suas limitações ajuda a escolher um método de avaliação eficiente.

Além disso, o histórico de usabilidade da empresa também é importante. Empresas com mais experiência em avaliações de usabilidade podem optar por métodos mais avançados e algumas preferem adotar uma abordagem de feedback contínuo, incorporando avaliações de usabilidade ao longo do ciclo de vida do *software*.

Por fim, consultar especialistas em usabilidade ou profissionais de design de experiência do usuário pode fornecer orientação valiosa na escolha do método de avaliação mais adequado. As recomendações de especialistas ajudam a garantir que as empresas tomem decisões informadas e eficazes para melhorar a usabilidade de seus produtos.

A escolha de um método de avaliação de usabilidade deve ser baseada nas necessidades e circunstâncias específicas de cada projeto. Não existe uma abordagem única que funcione para todos os cenários, e as empresas devem adaptar sua abordagem de acordo com os requisitos e recursos disponíveis.

#### 4. INSERINDO TESTE A/B EM EQUIPES ÁGEIS: UM GUIA PARA MELHORIA DE SISTEMAS WEB

Esta Seção apresenta o desenvolvimento de um guia prático e abrangente para facilitar a incorporação de testes A/B em equipes de desenvolvimento ágil, visando a avaliação e melhoria dos produtos de *software*.

A essência desse projeto é compreender como equipes ágeis podem integrar eficazmente os testes A/B em seus processos de desenvolvimento. Isso é crucial, pois as metodologias ágeis enfatizam a entrega rápida e frequente de *software*, e a usabilidade desempenha um papel fundamental na satisfação do usuário.

O guia proposto fornece uma estrutura objetiva que inclui a integração com a metodologia *Kanban* para acompanhamento e roteiros semi-estruturados para facilitar o planejamento, condução e análise dos testes. A avaliação desse guia foi realizada por meio de um *survey* que ajudou a identificar áreas de aprimoramento, focando em aspectos como utilidade e facilidade de uso percebidas.

No guia, o quadro *Kanban* é apresentado como uma ferramenta visual para gerenciar as tarefas e fluxos de trabalho na equipe ágil durante a avaliação de usabilidade. O quadro é estruturado em três colunas: "A fazer", "Fazendo" e "Feito". A avaliação utilizando teste A/B foi dividida em quatro fases principais: 1) Iniciação; 2) Planejamento; 3) Execução; e 4) Análise e Resultados.

Na iniciação, duas etapas principais são realizadas: a definição da equipe de avaliação e a leitura do material de teste. A equipe de avaliação deve ser composta por três papéis essenciais: o responsável pela avaliação, o especialista em design e o avaliador. Após a formação da equipe, todos os envolvidos devem realizar a leitura do material de teste para garantir um entendimento completo dos procedimentos e objetivos da avaliação.

Durante a fase de planejamento, o plano de avaliação é preparado, começando pela seleção de cenários de uso, onde se entende o contexto, identificam-se cenários representativos e consideram-se as diferenças entre as versões A e B. Em seguida, define-se a intenção do usuário, estabelecendo comportamentos desejados e critérios de sucesso. As versões da funcionalidade são desenvolvidas e revisadas, e métricas, pontuações e ferramentas relevantes para a coleta de dados são escolhidas. Tarefas semi-estruturadas são criadas para os participantes realizarem, os participantes representativos são recrutados com consentimento informado, e finalmente, são distribuídos equitativamente entre as versões A e B.

Na fase de execução, diversas atividades são realizadas para garantir a eficácia do teste A/B. Primeiro, o teste é aplicado, permitindo que os participantes interajam com as versões A e B. Em seguida, os participantes são monitorados atentamente, com observações e registros de suas interações. O esforço de interação dos usuários é avaliado, coletando-se dados de usabilidade. Finalmente, as conclusões da avaliação são registradas e documentadas detalhadamente, assegurando que todos os resultados sejam devidamente reportados.

Na fase de análise e resultados, o avaliador analisa observações e métricas para identificar problemas de usabilidade, classificando e priorizando esses problemas conforme sua gravidade. Em seguida, busca padrões nos dados, analisa tendências e prepara o relatório de resultados. Finalmente, realiza uma reunião de discussão com a equipe para debater os resultados e decidir qual versão será implementada.

Os artefatos gerados durante todo o guia incluem a lista da equipe de avaliação, *check-list* de leitura, plano de avaliação, registros de interações, métricas de avaliação, relatório

de análise de resultados, e a decisão sobre a versão a ser implementada.

O guia propõe um uso eficiente do *kanban* para gerenciar e visualizar o progresso de cada etapa do processo de teste A/B, mantendo o fluxo de trabalho organizado e eficiente.

Para uma consulta completa, o guia está disponível no repositório do projeto<sup>2</sup>.

## 5. AVALIAÇÃO POR SURVEY

A avaliação do guia é um passo crucial, pois permite analisar uma versão inicial para identificar problemas, coletar *feedback* dos usuários e realizar os ajustes necessários. Esse processo envolve a interação dos usuários com o guia, a medição da eficácia das instruções fornecidas e a determinação das áreas que precisam de melhorias. Além disso, a avaliação ajuda a garantir que o guia atenda às expectativas e necessidades dos usuários, proporcionando uma experiência de uso eficiente e satisfatória. Nesse sentido, o propósito da avaliação é verificar se a estrutura do guia é de fácil uso, se apresenta aspectos úteis para ajudar na avaliação de usabilidade e se desperta o interesse dos participantes em utilizá-lo futuramente.

A metodologia empregada para a avaliação do guia envolveu a realização de um *survey* online, que reuniu dados quantitativos e qualitativos. Este *survey* seguiu um processo estruturado em quatro etapas: Planejamento, Elaboração, Execução e Análise.

Na etapa de planejamento, foram definidas as questões do estudo, a amostra e os dados a serem coletados. As questões foram discutidas e aperfeiçoadas em reunião entre os pesquisadores. A amostra foi constituída por profissionais da indústria de software. Levando em consideração os tipos de amostragem da pesquisa, optou-se pela técnica probabilística de amostragem aleatória simples, na qual todos os elementos da população, ou seja, profissionais da indústria de software que utilizam metodologias ágeis no seu cotidiano de trabalho, têm a mesma probabilidade de pertencerem à amostra Grzybovski (2005).

Na etapa de elaboração, foram desenvolvidos os instrumentos de coleta de dados, especificamente o termo de consentimento livre e esclarecido, para receber a autorização dos participantes para a utilização das respostas; e o formulário a ser utilizado no *survey*. O formulário completo, foi constituído por 31 questões, subdivididas em 3 grupos, identificados por: dados demográficos; experiência e conhecimento em usabilidade e teste A/B; e avaliação do guia. O formulário pode ser acessado no repositório do projeto<sup>3</sup>. As questões de avaliação do guia foram definidas com base no TAM Davis (1989). A Tabela 1 mostra as 21 questões objetivas que compõem o questionário, divididas em três seções: utilidade percebida (UP), facilidade de uso (FUP) e pretensão de uso (PU). As 21 questões, foram divididas em duas categorias principais de resposta: questões com escala Likert, e questões dicotômicas. As questões da escala Likert usam a avaliação de 5 pontos, que permite aos respondentes expressar seu grau de concordância ou discordância com cada afirmação: 1 - Discordo Totalmente; 2 - Discordo Parcialmente; 3 - Neutro; 4 - Concordo Parcialmente; 5 - Concordo Totalmente. A segunda categoria tem as questões em formato binário, oferecendo aos respondentes a opção de responder com “Sim” ou “Não”. Em casos onde a resposta foi “Não”, foi solicitado que os participantes fornecessem uma justificativa para a sua escolha, permitindo uma compreensão mais profunda das razões por trás da resposta negativa. A versão final das questões foram detalhadas na Tabela 1.

Na etapa de execução, foram utilizadas plataformas online como *Whatsapp*, *e-mail* e *Slack* para facilitar o alcance aos profissionais da indústria de software e assegurar uma

---

<sup>2</sup>Inserindo teste A/B em equipes ágeis: Um guia para melhoria de sistema web

<sup>3</sup>Instrumentos de coleta de dados para a avaliação do guia

Tabela 1 - Questões baseadas no TAM para avaliação do guia de usabilidade.

<b>Utilidade Percebida (UP)</b>	
<b>Q1</b>	A definição clara dos papéis da equipe de avaliação é útil na organização da avaliação com teste A/B.
<b>Q2</b>	Os documentos gerados na fase de iniciação (como a lista da equipe de avaliação) são úteis para a organização da avaliação de usabilidade.
<b>Q3</b>	No geral, a fase de iniciação apresentada no guia é uma etapa útil na minha avaliação com testes A/B.
<b>Q4</b>	As instruções do guia sobre a seleção de cenários de uso contribuem para a identificação de interações-chave dos usuários.
<b>Q5</b>	As instruções do guia sobre a definição das métricas e ferramentas de avaliação são úteis na coleta de dados para a análise da usabilidade.
<b>Q6</b>	As instruções do guia sobre a definição de tarefas semi-estruturadas contribuem na organização das tarefas que serão realizadas pelo participante durante a avaliação.
<b>Q7</b>	As instruções do guia sobre o recrutamento de participantes contribuem para a obtenção de um grupo de usuários representativo, capaz de fornecer os dados para a avaliação de usabilidade.
<b>Q8</b>	O documento gerado na fase de planejamento (Plano de Teste A/B - Roteiro Semi-Estruturado) é útil para a organização da avaliação de usabilidade.
<b>Q9</b>	As instruções do guia sobre a aplicação do teste A/B contribuem na preparação para conduzir a avaliação entre as versões A e B.
<b>Q10</b>	As instruções do guia sobre o monitoramento dos participantes durante a avaliação contribuem para a observação e registro do comportamento dos usuários.
<b>Q11</b>	As instruções do guia sobre o registro dos resultados e métricas de usabilidade ajudam na preparação para documentar os dados de forma organizada.
<b>Q12</b>	As instruções do guia sobre a documentação detalhada das interações dos participantes são úteis para entender como registrar informações relevantes para a análise posterior.
<b>Q13</b>	As orientações do guia sobre como buscar padrões e tendências nos dados coletados ajudam a identificar aspectos a serem avaliados durante a análise.
<b>Q14</b>	O guia me permite realizar avaliações de usabilidade com maior eficiência?
<b>Facilidade de Uso Percebida (FUP)</b>	
<b>Q15</b>	As instruções do guia sobre os papéis da equipe de avaliação são claras e facilitam a organização da avaliação com teste A/B.
<b>Q16</b>	O documento gerado na fase de planejamento (Plano de Teste A/B - Roteiro Semi-Estruturado) facilita a organização da avaliação de usabilidade conforme descrito no guia.
<b>Q17</b>	As instruções do guia para documentar detalhadamente as interações dos participantes durante a execução da avaliação com testes A/B são fáceis de compreender.
<b>Q18</b>	As instruções do guia para construir o relatório de análise de resultados são suficientemente detalhadas para preparar uma apresentação clara dos resultados.
<b>Q19</b>	As orientações do guia para buscar padrões e tendências nos dados coletados durante a análise dos resultados dos testes A/B são fáceis de compreender.
<b>Q20</b>	A navegação pelo guia é fácil e intuitiva?
<b>Pretensão de Uso (PU)</b>	
<b>Q21</b>	Eu utilizaria o guia proposto em meu trabalho ou na empresa que atuo?

coleta de dados eficiente. As avaliações foram coletadas no período de junho de 2024, foram fornecidas instruções aos participantes para primeiramente ler o documento do guia completo anexado ao questionário e depois preencher o questionário. Para preencher o questionário, inicialmente o participante precisou aceitar o termo de consentimento livre e esclarecido. Durante a coleta de dados, foi mantido um monitoramento contínuo para garantir a integridade dos dados coletados e para resolver quaisquer dúvidas ou problemas técnicos que os participantes pudessem encontrar.

Na etapa de análise, os dados coletados foram organizados e analisados. Em seguida, foi realizada uma análise de conteúdo para identificar padrões e temas recorrentes nas respostas. As respostas foram categorizadas, para uma análise das percepções e experiências dos participantes. Os resultados foram interpretados considerando os objetivos do guia, de ser de fácil leitura e possibilitar avaliações de usabilidade em equipes ágeis.

## 5.1. PERFIL DOS PARTICIPANTES

A amostra foi constituída por 17 profissionais da indústria de software que trabalham em equipes ágeis. Desses, 7 participantes (41,2%) possuem ensino superior completo, 5 participantes (29,4%) cursaram pós-graduação lato sensu, 4 participantes (23,5%) estão cursando ensino superior, e 1 participante (5,9%) possui mestrado. Quanto às funções dos

participantes, incluem-se 5 *product owners* (29,4%), 3 *designers* (17,6%), 3 desenvolvedores de *software* (17,6%), 2 analistas de qualidade (11,8%), 2 *scrum masters* (11,8%) e 2 analistas de requisitos (11,8%). Com relação à experiência, 7 participantes (41,2%) possuem entre 3 e 5 anos, 5 participantes (29,4%) têm entre 1 e 3 anos, 4 participantes (23,5%) têm mais de 10 anos, e 1 participante (5,9%) tem entre 5 e 10 anos."

Dentre os participantes, 12 (70,6%) já participaram de avaliações de usabilidade, utilizando métodos como testes de usabilidade, testes A/B, entrevistas e avaliações heurísticas. Com relação ao nível de conhecimento em usabilidade e testes A/B, 9 participantes (52,9%) classificam seu conhecimento como intermediário, 6 participantes (35,3%) como básico, 1 participante (5,9%) possui conhecimento avançado, e 1 participante (5,9%) não possui conhecimento. Quanto à experiência com testes A/B, 12 participantes (70,6%) afirmaram já ter participado, enquanto 5 participantes (29,4%) não participaram devido à falta de oportunidades.

## 5.2. ANÁLISE E DISCUSSÃO QUANTITATIVA

Analisamos a Utilidade Percebida (UP) com base em treze questões e 17 respondentes em duas perspectivas: coesão interna das respostas através da aplicação da análise Alfa de Cronbach Cronbach (1951) e porcentagem de respostas por questões.

Aplicamos o Alfa de Cronbach às respostas das UP para medir sua coesão interna. Portanto, chegamos a  $\alpha = 0,802$ , ou seja, os respondentes tendem a responder da mesma maneira para todas as questões de UP apresentadas na Tabela 1 (Q1 a Q13). A Figura 1 resume os resultados obtidos em relação a cada questão UP.

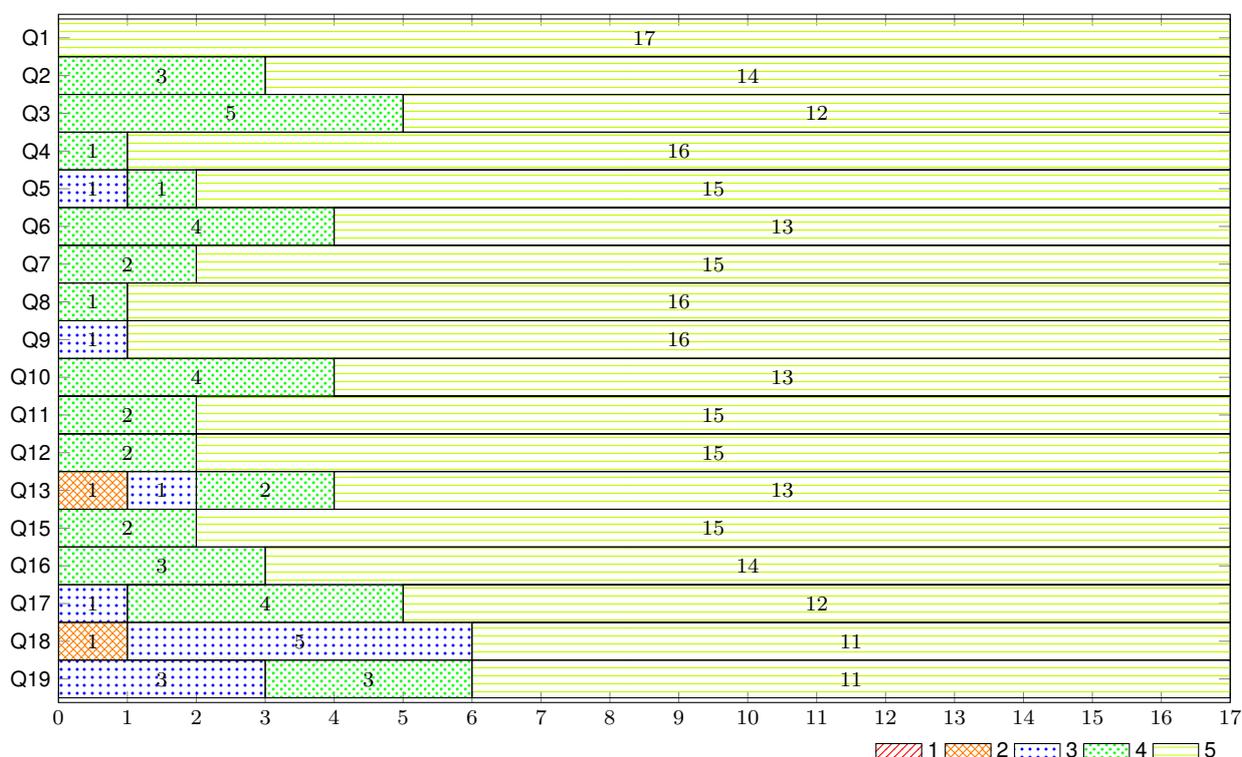


Figura 1 - Respostas para as questões de utilidade percebida (UP).

Considerando as respostas relacionadas à UP de forma individual, é possível identificar que em Q1, todos os 17 (100%) respondentes concordam que na etapa de iniciação, a definição clara dos papéis é útil. Isso sugere um consenso total sobre a importância dessa tarefa, indicando que a definição de papéis é um fator crucial para a organização e o sucesso

da avaliação com teste A/B. No que diz respeito à questão Q2, a maioria dos respondentes 14 (82,3%) concorda totalmente que os documentos gerados na fase de iniciação são úteis, enquanto 3 (17,6%) concordam parcialmente. Essa alta taxa de concordância indica que a documentação é amplamente vista como um recurso vital para a organização da avaliação de usabilidade. Em geral, a maioria dos respondentes na questão Q3, 17 (100%) consideram a fase de iniciação útil, 12 (70,6%) concordando totalmente e 5 (29,4%) concordando parcialmente. Isso nos dá indícios que a fase de iniciação é percebida como uma etapa essencial, embora haja uma pequena variação no nível de concordância.

Na questão Q4, os participantes acreditam que as instruções do guia sobre a seleção de cenários de uso contribuem para a identificação de interações-chave dos usuários, sendo que 16 (94,1%) relataram concordar totalmente e 1 (5,9%) concorda parcialmente. Referente a questão Q5, a maioria dos participantes, 16 (94,1%), acredita que as instruções sobre métricas e ferramentas de avaliação são úteis, 15 (88,2%) concordando totalmente e 1 (5,9%) concorda parcialmente, a pequena porcentagem (5,9%), ou seja, 1 respondente foi neutros indicando que, embora a maioria veja valor nessas instruções, há espaço para melhoria ou maior clareza. Na questão Q6, todos os consideram úteis as instruções sobre a definição de tarefas semi-estruturadas, 13 (76,5%) concordando totalmente e 4 (23,5%) concordaram parcialmente. Isso sugere que, apesar de um consenso geral sobre a utilidade, há uma variação na intensidade da concordância.

Em relação a questão Q7, os concordam sobre a utilidade das instruções de recrutamento, desses 15 (88,2%) concordando totalmente, que as instruções são vistas como fundamentais para garantir um grupo representativo de usuários. A variação mínima na concordância parcial de 2 (11,8%) participantes reforça a importância desse aspecto. Na questão Q8, os participantes acreditam que o documento gerado na fase de planejamento é útil, com 16 (94,1%) participantes concordando totalmente e 1 participante (5,9%) concorda parcialmente. Esse alto nível de concordância total reflete a importância da documentação para a organização da avaliação de usabilidade.

Relacionado a questão Q9, a maioria dos participantes, 16 (94,1%), acredita que as instruções sobre a aplicação do teste A/B são úteis, com apenas 1 participante (5,9%), neutro. Isso indica que, embora a maioria veja valor nessas instruções, há uma necessidade de esclarecer ou melhorar a abordagem para atender todos os participantes. Na questão Q10, quando perguntado se acreditam que as instruções do guia sobre o monitoramento dos participante contribuem para a observação e registro, 13 (76,4%) concordando totalmente e 4 (23,5%) concordam parcialmente. Isso destaca a importância do monitoramento detalhado durante a avaliação.

De acordo com as respostas obtidas na questão Q11, 15 participantes (88,2%) concordam totalmente que as instruções sobre o registro de resultados e métricas são úteis, enquanto 2 participantes (11,8%) concordam parcialmente. Isso sugere que a documentação organizada é vista como uma parte crítica do processo de avaliação de usabilidade. Na questão Q12, a maioria dos respondentes acreditam na utilidade das instruções sobre a documentação detalhada das interações, sendo 15 (88,2%) concordando totalmente, 2 (11,8%) concordam parcialmente. Isso mostra um alto nível de concordância com a importância da documentação detalhada para a análise posterior.

Referente a questão Q13, dos 17 participantes, 15 (88,2%) acreditam que as orientações sobre a busca de padrões e tendências são úteis para a análise dos dados. A presença de respostas neutras 1 (5,9%) e de discordância parcial 1 (5,9%) sugere que, embora as orientações sejam amplamente valorizadas, pode haver áreas que necessitam de maior clareza ou detalhamento.

Para analisar a Facilidade de Uso Percebida (FUP), primeiramente aplicamos o Alfa de Cronbach às respostas de cinco questões para medir sua coesão. Os resultados mostraram  $\alpha = 0,801$ , indicando que os respondentes tendem a responder de forma consistente em todas as perguntas apresentadas na Tabela 1 (Q15 a Q19). A figura 1 resume os resultados obtidos em relação a cada questão FUP.

Na fase de iniciação, a clareza sobre os papéis da equipe é essencial para garantir que todos entendam suas responsabilidades e contribuam de forma eficaz para a avaliação, na questão Q15, todos os 17 participantes (100%) consideram as instruções claras. Sendo que, 15 participantes (88,2%) responderam que concordam totalmente e 2 (11,7%) que concordam parcialmente. O que reflete uma forte aceitação da estrutura fornecida, comparado a outras fases, a alta concordância aqui sugere que a definição de papéis é uma área de destaque na percepção positiva do guia

Na fase de planejamento, o documento gerado é vital para estruturar a avaliação de usabilidade. Na questão Q16, 14 participantes (82,3%) responderam que concordam totalmente e 3 participantes (17,6%) que concordam parcialmente, que o "Plano de Teste A/B - Roteiro Semi-Estruturado" facilita a organização da avaliação. Este consenso indica que o documento ajuda a garantir um planejamento de avaliação organizada.

A compreensão clara das instruções sobre documentar interações é necessária durante a execução da avaliação. Na questão Q17, 12 participantes (70,6%) responderam que concordam totalmente, 4 participantes (23,5%) responderam que concordam parcialmente e 1 participante (5,9%) se manteve neutro. A presença de respostas parciais e neutras sugere que, embora a maioria tenha encontrado as instruções claras, pode haver áreas onde a clareza adicional ou exemplos práticos seriam benéficos para melhorar a compreensão.

Para a construção de um relatório de análise de resultados, a clareza e o detalhamento das instruções são fundamentais. Na questão Q18, 11 participantes (64,7%) concordaram totalmente que as instruções são suficientemente detalhadas, 5 participantes (29,4%) mantiveram-se neutros, e 1 participante (5,8%) discordou parcialmente. Embora a maioria tenha considerado as instruções adequadas, a neutralidade e a discordância indicam a necessidade de revisar e possivelmente expandir as diretrizes para garantir uma compreensão e aplicação efetivas para todos os usuários.

A identificação de padrões e tendências nos dados é crucial para uma análise eficaz dos resultados dos testes A/B. Na questão Q19, 11 participantes responderam (64,7%) que concordam totalmente que as orientações são fáceis de compreender, enquanto 3 participantes (17,6%) responderam que concordam parcialmente e 3 participantes (17,6%) se mantiveram neutros. A significativa quantidade de respostas neutras e parciais indica que, embora a maioria tenha encontrado clareza nas instruções, há oportunidades para tornar essas orientações ainda mais compreensíveis e acessíveis.

Considerando as respostas relacionadas à utilidade pretendida (UP), é possível identificar que, na questão Q21, todos os 17 participantes (100%) concordam que utilizariam o guia proposto em seu trabalho ou na empresa em que atuam. Essa unanimidade sugere uma aceitação total do guia e indica que ele é considerado altamente aplicável e útil no contexto profissional dos participantes. A alta taxa de aceitação (100%) é significativa, pois reflete a percepção positiva dos participantes sobre a praticidade e a relevância do guia para suas atividades diárias. Isso é relevante, pois demonstra que o guia atende às necessidades dos usuários e é visto como uma ferramenta eficaz para melhorar o processo de avaliação de usabilidade dentro das empresas.

Também foi disponibilizada uma questão aos participantes "O guia me permite realizar avaliações de usabilidade com maior eficiência?". Todos os 17 participantes (100%)

confirmaram que o guia permite realizar avaliações de usabilidade com maior eficiência. Essa unanimidade reflete uma percepção clara de que o guia é uma ferramenta eficaz para otimizar o processo de avaliação.

### 5.3. ANÁLISE E DISCUSSÃO QUALITATIVA

Nesta seção, analisamos as respostas de duas questões abertas, que foram essenciais para que os participantes pudessem expressar a sua percepção sobre o guia.

Na questão sobre a facilidade e intuitividade da navegação pelo guia, foi perguntado: "A navegação pelo guia é fácil e intuitiva?", 16 participantes (94,1%) relataram que a navegação é fácil e intuitiva, o que indica uma alta aceitação do design atual do guia. No entanto, 1 participante (5,9%) discordou, apontando que o guia pode ser melhorado em alguns aspectos específicos para facilitar ainda mais a navegação. O participante P6 sugeriu a inclusão de um sumário ou índice no início do documento para proporcionar uma visão geral do conteúdo e facilitar a navegação entre as seções. Além disso, recomendou a utilização de uma paleta de cores mais variada e com maior contraste para melhorar a legibilidade das imagens e diagramas presentes no guia. A clareza das setas no fluxograma também foi mencionada como um ponto de melhoria, destacando a necessidade de um sentido mais claro e consistente nas indicações de fluxo. Essas sugestões são valiosas, pois a navegação intuitiva e a clareza visual são elementos cruciais para a usabilidade do guia. Implementar essas melhorias pode não apenas tornar o guia mais acessível, mas também melhorar a experiência do usuário, permitindo uma compreensão mais rápida e eficaz do conteúdo.

Na avaliação da questão "Você identificou algum aspecto que pode ser melhorado no guia?", 11 participantes (64,7%) relataram que não identificaram nenhum aspecto a ser melhorado, sugerindo uma percepção positiva geral do guia. No entanto, 6 participantes (35,3%) identificaram áreas que, em sua visão, poderiam ser aprimoradas. Essas sugestões se concentram em seis aspectos principais:

- **Melhor sinalização na análise:** P1 sugeriu indicar mais claramente o momento adequado para realizar a análise, além de incluir uma estimativa de margem de acerto considerada confiável. Isso ajudaria a definir expectativas claras para os usuários e garantir uma aplicação consistente do guia.
- **Exploração de exemplos de uso:** Também sugerido por P1, a inclusão de exemplos de uso para simulações, o que pode proporcionar uma prática segura e realista para entender o processo de análise antes de aplicá-lo em situações reais.
- **Listar tipos de comportamentos de alerta para o designer:** P3 recomendou a criação de uma seção dedicada a tipos de comportamentos que devem acender alertas para o designer. Isso ajudaria os usuários a identificar tendências problemáticas durante a análise de usabilidade.
- **Adicionar um índice das etapas no início:** P4 sugeriu uma visão geral no início do guia, resumindo todas as etapas de forma concisa, permitiria aos usuários uma compreensão rápida do processo como um todo, facilitando a navegação e a consulta rápida.
- **Melhorar o layout do guia:** P6 sugeriu revisar o layout com foco em uma paleta de cores mais variada e um contraste adequado. Além disso, P7 sugeriu melhorias na hierarquia de informações e na diagramação para melhorar a legibilidade e a experiência do usuário.

- **Detalhamento da etapa final de análise e resultados:** P15 recomendou detalhar mais a etapa final, fornecendo instruções específicas e detalhadas, para equiparar essa seção às outras etapas do guia, como a de planejamento, que já é bem detalhada e facilita a abordagem da avaliação de usabilidade.

As sugestões, apontadas pelos participantes, são significativas para a melhoria contínua do guia. Implementar essas melhorias pode aumentar a eficácia e a facilidade de uso do guia, proporcionando uma ferramenta ainda mais robusta para os profissionais que o utilizam. O *feedback* positivo da maioria dos participantes indica que o guia possui uma base sólida, e essas melhorias podem fortalecer ainda mais sua aplicabilidade e relevância.

Também foi disponibilizada uma questão aberta na qual o participante pode colocar sua opinião livremente. "Outros comentários pertinentes à pesquisa", que revelou destaques importantes:

- **Elogios à estrutura:** O guia foi descrito como "show" e "bem estruturado" por P2 e P15, destacando a divisão clara em papéis e fases, o que facilita o seguimento em cada etapa. P6 demonstrou interesse na abordagem do *design* e manifestou a intenção de discutir mais sobre essa área, oferecendo contato para mais comentários.
- **Interesse em utilização:** Os participantes expressaram interesse em usar o material em suas empresas e solicitaram autorização para esse uso. P10 e P15 consideraram a proposta interessante e pediram para serem informados sobre a possibilidade de aplicar o guia em contextos profissionais. Esse interesse destaca a relevância prática do guia e sugere que ele pode ser uma ferramenta valiosa em ambientes de trabalho reais.
- **Valorização das metodologias:** A abordagem dos métodos e ferramentas de coleta de dados foi destacada como um ponto forte por P16, especialmente a inclusão de itens como a "facilidade de recordação", que os participantes não haviam considerado antes para avaliações de usabilidade. Esse reconhecimento sugere que o guia introduz metodologias que podem enriquecer as práticas de avaliação de usabilidade.

No geral, o guia foi bem aceito, com *feedbacks* positivos sobre a sua organização e utilidade prática em contextos profissionais.

#### **Adequações:**

Entre os resultados obtidos durante a execução do *survey* a sugestão de melhoria do P6 chamou bastante atenção, pois foi o único participante a relatar que a navegação pelo guia não é fácil e intuitiva, este participante relatou ter conhecimento avançado em avaliação de usabilidade e citou exemplos das avaliações que participou, inclusive avaliações com Teste A/B. O participante sugeriu a inclusão de um sumário ou índice no início do documento para proporcionar uma visão geral do conteúdo e facilitar a navegação entre as seções. Além disso, recomendou a utilização de uma paleta de cores mais variada e com maior contraste para melhorar a legibilidade das imagens e diagramas presentes no guia. A clareza das setas no fluxograma também foi mencionada como um ponto de melhoria, destacando a necessidade de um sentido mais claro e consistente nas indicações de fluxo. Essas sugestões foram consideradas e as adequações foram realizadas, pois a fácil navegação pelo guia é um dos objetivos a ser atingido com o material construído. Deste modo, o índice no início do guia foi inserido e as imagens de fluxogramas que ilustram as etapas da avaliação foram refeitas.

## 6. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

A condução da avaliação de usabilidade em sistemas web apresenta alguns desafios, como coletar dados, interpretá-los e posteriormente aplicar os resultados em forma de melhorias no sistema. Esta avaliação ainda se agrava quando utilizadas por equipes ágeis que necessitam de adaptação rápida e contínua, sendo assim necessita de métodos flexíveis de avaliação de usabilidade. Para mitigar estes desafios, neste trabalho foi desenvolvido um guia que teve como objetivo principal propor e avaliar a integração da avaliação de usabilidade com teste A/B no contexto de equipes ágeis, para melhorar a usabilidade de sistemas web.

O guia desenvolvido foi avaliado através de um *survey*, que oportunizou a coleta de opiniões essenciais para a pesquisa. Os resultados coletados indicaram que a ferramenta pode ser eficaz para otimizar avaliações de usabilidade, pois possui uma estrutura clara e inclui métricas como facilidade de aprendizado e facilidade de recordação, destacando sua aplicabilidade e relevância prática. A maioria dos participantes indicou que utilizaria o guia para conduzir avaliações de usabilidade nas empresas em que trabalham, sugerindo que o guia pode melhorar a eficiência e a qualidade dessas avaliações. A contribuição deste trabalho reside na promoção de uma abordagem estruturada e de fácil compreensão para avaliações de usabilidade.

É importante ressaltar que a avaliação conduzida identificou aspectos pertinentes para melhorar o guia, como a necessidade de uma melhor sinalização sobre quando realizar a análise dos resultados e um maior detalhamento sobre a etapa de análise e resultados. Outro ponto relevante sugerido foi a inclusão de exemplos de uso práticos. Após receber *feedback* sobre as cores e fluxogramas do guia, foram feitos ajustes nas imagens. No entanto, é recomendável considerar uma revisão geral no *layout* do guia com a ajuda de *designers* experientes no futuro. Esses pontos sugeridos são fundamentais para a criação de uma segunda versão do guia. Após essas melhorias, será importante realizar uma nova avaliação para confirmar a eficácia do guia aprimorado.

Em trabalhos futuros, serão realizados estudos de caso com o objetivo de avaliar a eficácia do guia em diferentes contextos organizacionais e empresas de desenvolvimento de *software*. A colaboração entre pesquisadores e organizações nos ajudará a avaliar o uso do guia em ambientes de trabalho com equipes ágeis na prática, ampliando seu impacto e utilidade. Além disso, uma área promissora para desenvolvimento adicional é a expansão do guia para incluir módulos de treinamento interativos com exemplos reais de aplicabilidade e a adaptação para utilizar outros métodos de avaliação de usabilidade, garantindo uma diversificação na avaliação.

## REFERÊNCIAS

ANDERSON, D. J.; CARMICHAEL, A. Kanban esencial condensado. Kaban University, 2016.

BARBOSA, S.; SILVA, B. **Interação humano-computador**. [S.I.]: Elsevier Brasil, 2010.

BARBOSA, S. D. J. et al. **Interação Humano-Computador e Experiência do Usuário**. [S.I.]: Autopublicação, 2021.

BOUCINHA, R. M.; TAROUÇO, L. M. R. Avaliação de ambiente virtual de aprendizagem com o uso do sus-system usability scale. **RENOTE**, v. 11, n. 3, 2013.

CRONBACH, L. Coefficient alpha and the internal structure of tests. **Psychometrika**, Elsevier, p. 297–334, 1951.

DAVIS, F. D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. **MIS quarterly**, JSTOR, p. 319–340, 1989.

DOBRIGKEIT, F.; PAULA, D. de et al. The best of three worlds—the creation of innodev a software development approach that integrates design thinking, scrum and lean startup. In: **DS 87-8 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 8: Human Behaviour in Design, Vancouver, Canada, 21-25.08. 2017**. [S.l.: s.n.], 2017. p. 319–328.

FALCÃO, M. et al. Avaliação de usabilidade em simuladores de autômatos. In: SBC. **Anais da XI Escola Potiguar de Computação e suas Aplicações**. [S.l.], 2018. p. 132–135.

FIRMENICH, S. et al. Usability improvement through a/b testing and refactoring. **Software Quality Journal**, Springer, v. 27, p. 203–240, 2019.

GRZYBOVSKI, D. Revisão teórica sobre pesquisa quantitativa, mensuração, amostragem e análise multivariada. **Texto para Discussão**, v. 13, p. 2–20, 2005.

NAGAI, R. A.; SBRAGIA, R. As origens da metodologia ágil: de onde saímos e onde estamos? uma revisão sistemática da literatura. **Revista de Gestão e Projetos**, v. 14, n. 1, p. 11–41, 2023.

NIELSEN, J. **Usability engineering**. [S.l.]: Morgan Kaufmann, 1994.

SANTOS, M. C. F. R. d. **O ecossistema de startups de software da cidade de São Paulo**. Tese (Doutorado) — Universidade de São Paulo, 2016.

SILVA, I. F. S. d. Desenvolvendo um processo de teste a/b para avaliação de usabilidade e experiência do usuário. Universidade Federal do Maranhão, 2023.

SILVA, T. S. da; SILVEIRA, M. S. Integrando avaliação de usabilidade e métodos ágeis. **V Mostra de Pesquisa da Pós-Graduação–PUCRS**, 2010.

SPEICHER, M.; BOTH, A.; GAEDKE, M. Wappu: usability-based a/b testing. In: SPRINGER. **Web Engineering: 14th International Conference, ICWE 2014, Toulouse, France, July 1-4, 2014. Proceedings 14**. [S.l.], 2014. p. 545–549.

VAZ, V. T. et al. Inspeção de usabilidade em organizações de desenvolvimento de software—uma experiência prática. In: SBC. **Anais do VII Simpósio Brasileiro de Qualidade de Software**. [S.l.], 2008. p. 369–377.

WINCKLER, M.; PIMENTA, M. S. Avaliação de usabilidade de sites web. **Escola de Informática da SBC Sul (ERI 2002). Porto Alegre**, v. 1, p. 85–137, 2002.

WOHLIN, C. et al. **Experimentation in Software Engineering**. [S.l.]: Springer Publishing Company, Incorporated, 2012. ISBN 3642290434.