

# Avaliação de Métodos para Elicitação e Especificação de Requisitos de Usabilidade com Histórias de Usuário: Um Experimento Controlado

Clara Lima Fonseca  
claralima@alu.ufc.br  
Universidade Federal do Ceará (UFC)  
Russas, Ceará, Brasil

Luciana A.M. Zaina  
lzaina@ufscar.br  
Universidade Federal de São Carlos  
(UFSCar)  
Sorocaba, São Paulo, Brasil

Anna Beatriz Marques  
beatriz.marques@ufc.br  
Universidade Federal do Ceará (UFC)  
Russas, Ceará, Brasil

## RESUMO

Considerar aspectos de qualidade desde os primeiros estágios do processo de desenvolvimento de software é fundamental para aumentar seu sucesso, evitar retrabalho e garantir a satisfação do usuário final. Entretanto, no contexto do desenvolvimento ágil, esses aspectos são vistos como secundários em relação aos requisitos do software. Para resolver esse problema, pesquisas sugerem abordagens para incorporar requisitos não funcionais, como usabilidade e experiência do usuário, nas histórias de usuários, que são um artefato valioso para as equipes ágeis. Este artigo tem como objetivo comparar duas abordagens que apoiam equipes ágeis na elicitação e especificação de artefatos que ajudam a melhorar a qualidade do software. Duas abordagens propostas recentemente, USARP (USability Requirements with Personas and user stories) e ACUX (Acceptance Criteria of User eXperience) foram selecionadas para serem comparadas em um experimento controlado com a participação de 32 indivíduos. Foram coletados dados quantitativos e qualitativos para avaliar a eficácia, a eficiência, a facilidade de uso e a utilidade das duas abordagens. Os resultados revelaram que o USARP obteve um desempenho melhor em relação ao ACUX em termos de facilidade de uso. Por outro lado, o ACUX demonstrou maior capacidade de especificar as interações do usuário e as funcionalidades do sistema. Os participantes também elogiaram a praticidade, o fornecimento de exemplos e os padrões para a escrita de requisitos de qualidade. Espera-se que esses resultados sirvam para identificar as melhores práticas do USARP e do ACUX, orientar o desenvolvimento futuro de métodos nessa área e contribuir para o aprimoramento das abordagens estudadas.

## KEYWORDS

Requisitos de Software, Usabilidade, Experiência do usuário, Experimento controlado

## 1 INTRODUÇÃO

Compreender as necessidades e expectativas dos usuários fornece insumos para criar experiências mais satisfatórias e eficazes, que resultam em uma maior qualidade do produto de software [10]. Essas informações estão alinhadas aos aspectos de usabilidade e experiência do usuário (UX, do inglês *User eXperience*) e para que sejam consideradas como parte do escopo de um projeto é ideal que sejam descritas como requisitos de software [19].

Um dos principais desafios enfrentados na Engenharia de Requisitos Ágil é a negligência dos requisitos de qualidade ou não-funcionais (RNFs) [8, 40]. Três estudos retornados na revisão sistemática de Curcio et al. sobre usabilidade no desenvolvimento ágil apontaram que a falta de documentação dos requisitos de usabilidade é um obstáculo [9]. Na documentação de RNFs, os desafios incluem manter a validade dos requisitos atualizados conforme novas expectativas são comunicadas e, garantir a expressão clara e mensurável destes, evitando detalhes vagos ou irracionais.

A dificuldade em descrever requisitos de UX e usabilidade aponta para a necessidade de técnicas que estimulem muito mais do que a reflexão sobre quais requisitos devem ser descritos. É fundamental dar ferramentas que direcionem os times ao que deve ser descrito de forma que esses requisitos sejam aderentes ao seu propósito, UX e usabilidade, ao mesmo tempo que informem de forma clara qual requisito será trabalhado [47].

Diversas pesquisas propõem métodos e técnicas para levantar e documentar requisitos de qualidade em Histórias de Usuários (USs, do inglês *User Stories*) [5, 24, 28, 29]. As USs são principalmente utilizadas para representar as intenções do usuário em relação ao sistema e são usualmente complementadas por critérios de aceitação (AC, do inglês *Acceptance Criteria*) [6]. Os AC descrevem o que são requisitos a serem atendidos para que uma dada US atinja seus objetivos [6].

Embora as USs sejam artefatos valiosos para o trabalho do time sua utilização para documentar requisitos não funcionais é pouco explorada [20]. Nesse sentido, torna-se fundamental explorar técnicas e métodos que deem suporte à escrita de USs para que essas especifiquem os requisitos de UX e usabilidade relacionados ao software que será desenvolvido. Considerando essa lacuna de investigação, foi planejado um estudo que buscou analisar dois métodos propostos recentemente que possuem propósitos similares de especificar requisitos de qualidade, mas com abordagens diferentes. O método USARP (USability Requirements with Personas and user stories) [24, 34] consiste no uso de diretrizes para a elicitação de requisitos funcionais de usabilidade, utilizando personas e USs. Já o método ACUX (Acceptance Criteria of User eXperience) [28] busca descrever os aspectos de UX por meio de AC das USs. Para as duas abordagens, as USs atuam como um meio de representar as necessidades e intenções dos usuários em relação ao sistema.

Tendo como objetivo de estudo abordagens para auxiliar a especificação de requisitos de UX e usabilidade, um experimento controlado [46] foi conduzido. Foram critérios usados para avaliar as abordagens a eficácia, eficiência e aceitação dos métodos do

ponto de vista de futuros desenvolvedores de software (N=32 participantes) Uma análise quantitativa foi realizada sobre os dados para investigar a existência de diferença significativa entre os métodos em relação aos critérios definidos acima. Procedimentos de análise qualitativa foram adotados para analisar os aspectos de usabilidade e UX que os métodos permitiram especificar. Os resultados demonstraram que a USARP foi percebida com um nível maior de facilidade de uso em relação à ACUX. Não foi possível identificar diferenças significativas entre os métodos em relação à utilidade percebida, eficácia e eficiência para a elicitación e especificação de requisitos de UX e usabilidade. Estes resultados contribuem ao demonstrar as potencialidades de cada método estudado para especificação de requisitos de UX e de usabilidade. Isto torna-se importante para que futuras propostas possam focar nessas potencialidades para estudar outros métodos, ou então, propor outras características não presentes nos métodos estudados.

O restante deste artigo está organizado da seguinte forma: a Seção 2 apresenta o background sobre usabilidade, UX, histórias de usuários e critérios de aceitação. A Seção 3 apresenta os trabalhos relacionados. A Seção 4 descreve os métodos que foram objetos de estudo desta pesquisa. A Seção 5 detalha a metodologia de experimentação adotada. A Seção 6 apresenta a análise dos resultados quantitativos e qualitativos. A Seção 7 apresenta a discussão, enquanto a Seção 8 apresenta as conclusões e trabalhos futuros.

## 2 BACKGROUND

A ISO 25010:2011 destaca a usabilidade como a capacidade de um produto ser usado para atingir objetivos específicos com eficácia, eficiência e satisfação [1]. Norman e Deiró [32] e Nielsen [30] enfatizam a facilidade de aprendizado, eficiência de uso, facilidade de memorização, baixa taxa de erros e satisfação subjetiva do usuário como critérios de usabilidade. Esses critérios qualificam a interação do usuário com o sistema.

A UX vai além da usabilidade, envolvendo fatores emocionais, estéticos e de valor percebido [2, 31]. Autores como Garrett [11] e Krug [22] a definem como o resultado de todos os aspectos de um sistema, incluindo design visual, interação e arquitetura de informação. Destacam-se aspectos cognitivos, emocionais e sociais moldando a UX, influenciada pelo contexto de uso e expectativas do usuário [38].

Os dois conceitos podem estar relacionados tanto a requisitos funcionais como não funcionais. Os funcionais descrevem o comportamento do sistema, enquanto os não funcionais especificam propriedades ou restrições. Formalmente, também existem as regras de negócio que são declarações externas de política da empresa que impulsionam as funcionalidades do sistema.

As Histórias de Usuário [do inglês, User Stories (US)] surgiram como uma técnica no contexto do desenvolvimento ágil de software, especialmente no Extreme Programming (XP) proposto por Beck and Andres [4]. Elas foram concebidas para capturar requisitos de maneira colaborativa e centrada no usuário, enfocando o valor que cada funcionalidade agrega. No contexto da Engenharia de Software, os requisitos do software são levantados logo na primeira fase de Especificação e são fundamentais pois, guiam o desenvolvimento do software [16, 43]. Amplamente adotadas em projetos ágeis, as USs destacam-se por promoverem uma compreensão compartilhada dos

requisitos, centrando-se no valor para o usuário. Além disso, sua linguagem simples e acessível facilita a comunicação, permitindo estimativas mais precisas ao dividir as USs em tarefas menores [6, 40].

Jeffries [17] destaca a importância da simplicidade e comunicação efetiva no desenvolvimento de software, identificando três aspectos críticos das USs, conhecidos como "Card, Conversation, Confirmation" (3Cs). O Cartão representa uma breve descrição tangível do requisito do usuário, visualizando a funcionalidade. A etapa da Conversa envolve detalhes específicos registrados por meio de conversas colaborativas entre a equipe e *stakeholders*. A Confirmação estabelece critérios claros de aceitação, definindo as condições para considerar a US completa e correta, frequentemente seguindo o padrão "Dado [um contexto] Quando [ação] Então [resultado]" [33]. A Figura 1 exemplifica essa estrutura.

US:	01
<b>CARTÃO</b> <uma descrição da US para estimativa e planejamento>	<b>Como</b> <persona> <b>Eu quero</b> <ação> <b>para</b> <benefício>
<b>CONVERSA</b> <informações adicionais para promover uma maior compreensão da US>	Aspectos de usabilidade/UX, Protótipos, etc
<b>CONFIRMAÇÃO</b> <cenários de teste para confirmar que a US foi completada>	<b>Dado</b> <contexto> <b>Quando</b> <evento> <b>Então</b> <resultado>

**Figura 1: Modelo de uma US no modelo 3C criado por Ron Jeffries**

## 3 TRABALHOS RELACIONADOS

Os trabalhos apresentados nesta seção estão relacionados aos métodos, encontrados após uma busca sobre Elicitación e Especificação de Requisitos de Usabilidade e UX.

Keshk et al. [21] abordam as limitações das técnicas de prototipação e histórias de usuário enriquecidas, propondo uma abordagem de oito passos que combina essas técnicas com etapas extras de validação. No entanto, embora os passos propostos sejam compatíveis com o levantamento e especificação de requisitos, a abordagem não é específica para requisitos de usabilidade ou experiência do usuário.

Pereira et al. [35] apresentam cinco padrões para incorporar requisitos de experiência do usuário em Histórias de Usuário, visando facilitar a recuperação e localização dessas informações pela equipe ágil. No entanto, o estudo não detalhou as atividades de elicitación e especificação dos requisitos.

Sousa et al. [44] apresentaram uma solução móvel de apoio no levantamento de requisitos por meio de entrevistas semiestruturadas, que captura as respostas do stakeholder por voz e gera automaticamente histórias de usuário. No entanto, esse sistema não se concentra especificamente na elicitación e especificação de requisitos de usabilidade ou experiência do usuário.

Embora três métodos [21, 27, 34] para elicitação e especificação de requisitos tenham sido identificados, apenas os métodos USARP e ACUX são especificamente direcionados para requisitos de Usabilidade e UX. Enquanto o método USARP trata de requisitos funcionais de usabilidade em um campo adicional da US, o método ACUX aborda aspectos de UX como critérios de aceite na seção "Confirmação", seguindo o formato proposto por Jeffries [17]. Ambos os métodos visam melhorar a interação do usuário com a aplicação, embora em seções diferentes da US. Devido à semelhança entre eles, um experimento foi conduzido para avaliar esses dois métodos em um contexto justo para ambos.

Seis trabalhos sobre o uso dos métodos USARP e ACUX, separadamente, foram encontrados. Há investigações [39] [25] e estudos de caso documentados entre cada um, com participantes em ambientes acadêmicos [24], *startups* [28] e indústria [23] [27]. Como diferencial deste trabalho, foi realizado uma comparação entre os métodos USARP e ACUX através de um experimento controlado.

## 4 MÉTODOS SELECIONADOS

### 4.1 USARP

O método USARP (USAbility Requirements with Personas and user stories) proposto por Oliveira et al. [34] é um método para identificar e descrever requisitos de usabilidade, utilizando uma combinação de personas, histórias de usuário e diretrizes de usabilidade denominadas Functional Usability Features (FUFs) [19]. Esse método permite compreender as necessidades e expectativas de cada usuário, descrever aspectos de usabilidade específicos para cada perfil de usuário do sistema e auxiliar a equipe de desenvolvimento na concepção de uma experiência do usuário aprimorada.

As diretrizes utilizadas no método, propostas por Juristo et al. [19] são decompostas em mecanismos de usabilidade e transformadas em requisitos funcionais para uso no método USARP. Essa nomenclatura foi adotada devido à constatação de que esses requisitos não podiam mais ser considerados exclusivamente como requisitos não-funcionais, uma vez que afetavam tanto a interface quanto as funcionalidades do sistema. Além disso, o método sugere a utilização de personas para levantar os requisitos iniciais do sistema, uma vez que pode-se obter informações como necessidades, preferências e desejos dos possíveis usuários [24].

Em sua mais recente versão<sup>1</sup>, é possível realizar o *download* das Cartas, ter acesso ao Board no Figma para ter um maior controle dos registros e obter o *Checklist* da USARP que serve como um guia para identificar as cartas que se adequam à US. As cartas categorizadas em três tipos facilitam o levantamento e especificação de requisitos de usabilidade. Um conjunto de cartas como exemplo é mostrado na Figura 2.

- **Mecanismo de usabilidade:** Essa carta fornece uma descrição para cada mecanismo adotado pelo método;
- **Requisito de Usabilidade:** Contém questões a serem discutidas pela equipe para definir se um determinado mecanismo de usabilidade é aplicável e relevante para um conjunto de histórias de usuário;
- **Prototipação:** Contém questões sobre como o mecanismo de usabilidade será fornecido pela interface de usuário.

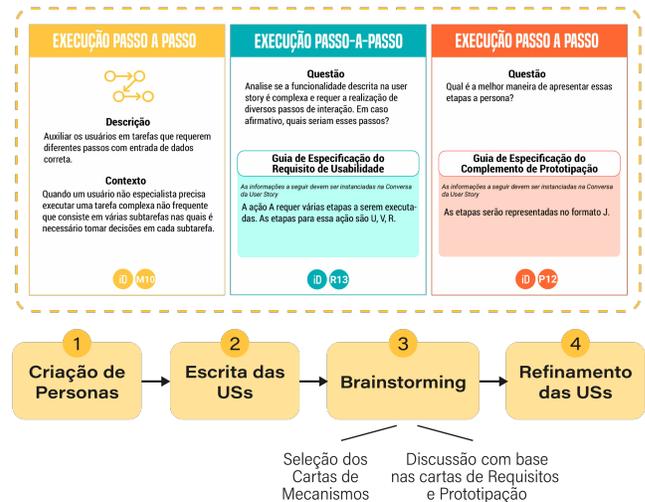


Figura 2: Cartas e Fluxograma da USARP

### 4.2 ACUX

O método ACUX (Acceptance Criteria of User eXperience) idealizado por Martinelli et al. [28], propõe um conjunto de diretrizes com o objetivo de auxiliar a redação de Critérios de Aceitação (ACs, do inglês *Acceptance Criteria*) de UX em USs, por equipes de desenvolvimento de software.

As diretrizes utilizadas no método foram resultantes de um Mapeamento Sistemático da Literatura, Estudo Exploratório e do *framework* proposto por Jesse [18]. Com o objetivo de auxiliar no desenvolvimento de produtos com uma UX aprimorada, Jesse divide a estrutura do UX em cinco camadas, que de baixo pra cima são: Estratégia, Escopo, Estrutura, Esqueleto e Interface, de modo que as camadas inferiores tratam das etapas mais abstratas da construção do software e as camadas superiores das etapas mais concretas [18]. No ACUX Martinelli et al., as diretrizes focam nas etapas mais concretas e são organizadas em dois grupos: Design de Interação e Organização da Informação (DI) e Elementos Visuais (EV). Os grupos DI e EV possuem 6 e 9 diretrizes, respectivamente. Esse agrupamento foi realizado levando em consideração o impacto direto que essas áreas têm na interação do usuário com o produto, com o objetivo de melhorar a comunicação dos significados das diretrizes [27]. A Figura 3 mostra uma diretriz de Design da Interação e Organização da Informação do guia interativo disponibilizado na plataforma Figma.

O conjunto de diretrizes ACUX é apresentado por meio de um guia interativo<sup>2</sup> que contém: as diretrizes de UX, um exemplo prático de redação em cada uma, artefatos de UX recomendados para coletar informações sobre o usuário e aborda os erros comumente cometidos na definição dos ACs. Todas as diretrizes são elaboradas seguindo o modelo proposto por Cohn [7] para as USs "Como um [tipo de usuário], eu quero [algum objetivo] para que [tenha algum benefício]" e o modelo de North [33] para a escrita dos ACs "Dado [um contexto] Quando [realizo uma ou mais ações] Então [resultado esperado]".

<sup>1</sup><https://usarp.github.io/>

<sup>2</sup>Guia interativo do ACUX no Figma



Figura 3: Guia interativo do ACUX

Por fim, ambos os métodos, USARP e ACUX, abrangem duas etapas da ER: A etapa de Elicitação e Análise de Requisitos, que envolve a compreensão das necessidades dos stakeholders, utilizando técnicas como questionários, entrevistas e workshops [38]; E a etapa de Especificação de Requisitos, que consiste na documentação detalhada dos requisitos, utilizando linguagem natural, linguagem estruturada ou notações gráficas [36, 42].

Os métodos USARP e ACUX não apresentam uma similaridade explícita. O método USARP nomeia suas cartas com base nas FUFs (do inglês, Functional Usability Features), que são funcionalidades que um sistema pode implementar. Por outro lado, o ACUX classifica suas diretrizes em dois grupos: Design de Interação e Organização da Informação e Elementos Visuais. Ambas as metodologias, USARP e ACUX, demonstram contribuições significativas para a especificação e refinamento de diversas funcionalidades. Por exemplo, ambos os métodos colaboram na especificação da parte visual do sistema, seja por meio da carta de Prototipação da USARP ou das diretrizes da categoria Elementos Visuais do ACUX. Por exemplo, um usuário pode especificar na US "Realizar uma compra na loja online" a necessidade da função Cancelar (USARP) e, ao mesmo tempo, detalhar aspectos visuais como paleta de cores, elementos

apropriados e organização na tela (ACUX). Em resumo, estes métodos apoiam diferentes aspectos e podem se complementar quando utilizados em conjunto.

## 5 EXPERIMENTAÇÃO

Os experimentos representam ferramentas cruciais para os engenheiros de software que buscam avaliar e selecionar entre diferentes métodos, técnicas, linguagens e ferramentas. Nesta seção, será delineado o procedimento adotado para conduzir uma abordagem experimental com os métodos USARP e ACUX, embasado em Wohlin et al. [46].

### 5.1 Planejamento

**Escopo:** O primeiro passo foi definir o objetivo do experimento seguindo a abordagem GQM (Goal-Question-Metrics) [3]. O objetivo do estudo é analisar os métodos USARP e ACUX com o intuito de comparar em relação à eficácia, eficiência, utilidade percebida e facilidade de uso percebida do ponto de vista de pesquisadoras em Engenharia de Software no contexto de uma experiência prática de elicitación e especificação de requisitos de usabilidade e UX com a participação de estudantes de Engenharia de Software e Ciência da Computação da Universidade Federal do Ceará (UFC).

**Contexto:** O experimento foi conduzido em uma turma da disciplina de Engenharia de Software na Universidade Federal do Ceará (UFC). O contexto é considerado específico por se tratar de um ambiente educacional. Esta turma adota as metodologias de sala de aula invertida e gamificação. Durante o decorrer desta disciplina, os estudantes obtiveram conhecimento sobre os temas Requisitos de Software, Especificação de Requisitos Ágeis com Us e Técnicas de Levantamento de Requisitos.

**Formulação de Hipóteses:** As hipóteses formuladas neste estudo serão a base para a análise estatística dos resultados obtidos para cada um dos critérios de avaliação: eficácia, eficiência, percepção de facilidade de uso e percepção de utilidade. As hipóteses nulas ( $H^0$ ) representam a ausência de diferença significativa entre os métodos, enquanto as hipóteses alternativas ( $H^1$ ) indicam a existência de diferenças substanciais. Dessa forma, foram formuladas duas hipóteses para cada um dos critérios avaliados, onde CRI (de Critério) é substituído por: Percepção de Facilidade de Uso (PFU), Percepção de Utilidade (PU), Eficácia e Eficiência, conforme a seguir:

$$H^0 : CRI(USARP) = CRI(ACUX) \quad (1)$$

$$H^1 : CRI(USARP) \neq CRI(ACUX) \quad (2)$$

Ressalta-se que as variáveis de Eficácia e Eficiência referem-se à equipe que adota o método, enquanto os fatores de Facilidade de Uso e Utilidade são atribuídos individualmente aos participantes.

**Participantes:** A turma do experimento era composta por 50 estudantes de graduação cursando a disciplina de Engenharia de Software. Esta disciplina é obrigatória para o curso de Ciência da Computação e optativa para o curso de Engenharia de Software, sendo a turma diversificada em relação aos cursos dos estudantes. No tocante à randomização das equipes, esta se mostrou semi-aleatória, em virtude da presença de participantes com distintos níveis de experiência no mercado de trabalho na área de TI. Para

equilibrar isso, uma estratégia de equilíbrio foi implementada, garantindo que o número de participantes com experiência de trabalho em cada equipe fosse aproximadamente igual.

**Design do Experimento:** No estudo em questão, as métricas definidas na Tabela ?? são consideradas variáveis dependentes, sendo observadas, medidas e registradas para analisar os efeitos das manipulações nas variáveis independentes. Essas variáveis independentes incluem o método de elicitação e especificação, o nível de experiência de trabalho na área de TI, o gênero e a experiência anterior com os métodos.

O experimento adotou um design totalmente aleatório, garantindo que cada participante avaliasse apenas um dos tratamentos, representados pelos métodos USARP e ACUX. As variáveis dependentes de PFU e PU foram medidas em uma escala de Likert de 5 pontos e o teste paramétrico foi selecionado. Neste estudo, o Teste t student foi utilizado.

**Balaceamento:** O balaceamento das equipes foi realizado com base nas informações coletadas no questionário de caracterização dos participantes, visando garantir uma distribuição justa. As características consideradas incluíram a experiência no mercado, com 9 participantes indicando trabalho ou experiência anterior em TI, distribuídos de maneira equitativa entre as equipes. Além disso, 2 participantes integrantes do projeto de pesquisa que trata da evolução do método USARP foram designados para equipes que utilizariam o método ACUX. Também buscou-se uma distribuição equitativa de gênero entre as equipes, visando formar grupos igualitários em termos de gênero. A Figura 4 ilustra os métodos atribuídos a cada equipe e o perfil dos participantes.

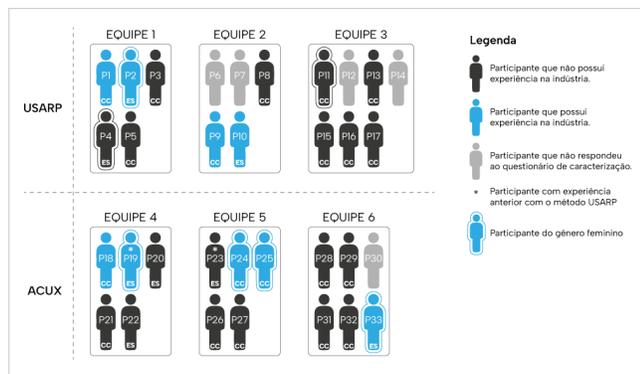


Figura 4: Distribuição dos participantes

**Instrumentação:** Para realizar este experimento, foram elaborados os objetos e instrumentos descritos na Tabela 1:

A Tabela 2 apresenta as afirmativas do modelo TAM [45] adaptadas para cada método e inseridas no questionário de feedback. Os cards presentes no guia online do ACUX, que contêm as instruções para a especificação de cada diretriz, foram impressos e transformados em cartões. Essa abordagem garantiu que a ACUX pudesse ser usada sem o uso de computador, proporcionando condições de execução similares para todas as equipes.

O cenário descrevia um problema de software relatando um aplicativo mobile de uma rede social para pets. Este cenário foi pensado na escolha de um assunto comum, a fim de reduzir as ameaças de

Tabela 1: Objetos do experimento

Questionário de Caracterização	Um questionário para identificar o perfil dos participantes.
Questionário de Feedback	Afirmações de Facilidade de Uso e Utilidade do modelo TAM [45], com escala de Likert variando entre "Discordo totalmente" (1), "Discordo" (2), "Neutro" (3), "Concordo" (4) e "Concordo totalmente" (5).
Cenário	Um problema de software simulando um desafio real da indústria, acompanhado de USs, protótipo e storyboards
Cartas da USARP	Um conjunto de 9 cartas de Mecanismos, 20 de Requisitos de Usabilidade e 17 de Prototipação.
Cartões do ACUX	Um conjunto de 6 cartões com diretrizes sobre Design de Interação e Organização da Informação e 9 cartões sobre Elementos Visuais.

não familiaridade com a aplicação por parte dos participantes. Para a atividade foram desenvolvidas um cenário, duas USs, duas telas de protótipo e duas storyboards. A Figura 5 mostra uma visão geral do material que foi recebido pelos participantes.



Figura 5: Artefatos de apoio: Cenário, USs, Protótipo e Storyboards.

Como apoio para a análise dos artefatos elaborados pelos participantes, também foi documentado um oráculo para cada história de usuário com os requisitos possíveis de serem especificados utilizando os dois métodos, USARP e ACUX. O oráculo foi construído por uma pesquisadora e revisado por duas pesquisadoras especializadas nos métodos USARP e ACUX.

**Preparação:** Foi percebido que instruir os estudantes previamente sobre os métodos poderia mitigar as ameaças relacionadas à aprendizagem e, por conseguinte, às conclusões do experimento.

**Tabela 2: Escalas utilizadas para USARP/ACUX**

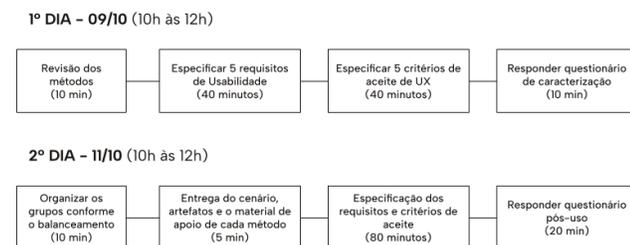
Fator	Rótulo	Afirmação	Fonte
Percepção de Facilidade de Uso	PFU1	Minha interação com a/o <método> é clara e compreensível	Adaptado de Venkatesh and Bala [45]
	PFU2	Interagir com a/o <método> não requer muito do meu esforço mental	
	PFU3	Considero o a/o <método> fácil de usar	
	PFU4	Considero fácil usar a/o <método> para <elicitar requisitos de usabilidade/escrever critérios de aceite de UX>	
Percepção de Utilidade	PU1	Utilizar a/o <método> melhora meu desempenho no refinamento de histórias de usuário	Adaptado de Venkatesh and Bala [45]
	PU2	Utilizar a/o <método> no refinamento de histórias de usuário aumenta minha produtividade	
	PU3	Utilizar a/o <método> aumenta minha eficácia no refinamento de histórias de usuário	
	PU4	Considero a/o <método> útil no refinamento de histórias de usuário	

Para isso, foram incluídas duas vídeo-aulas sobre os métodos USARP e ACUX, seguindo a metodologia de sala invertida e gamificação adotada na disciplina. As vídeo-aulas apresentaram os objetivos, estrutura, forma de adoção e exemplos práticos de ambos os métodos, além de fornecerem material de apoio, como links para o site do USARP e o guia do ACUX. Apesar dos esforços, apenas seis visualizações foram registradas nos vídeos até a véspera do experimento, o que levou à necessidade de apresentar os métodos em sala de aula, seguindo o mesmo formato das vídeo-aulas, no início do primeiro dia do experimento.

Os participantes não tinham conhecimento de quais aspectos estavam sendo estudados. Eles foram informados de que aprenderiam sobre dois métodos de eliciação e especificação de requisitos, como abordagem prática do conteúdo de requisitos ágeis incluso no plano de aula. Todos os participantes tiveram o anonimato preservado.

## 5.2 Execução

O experimento foi conduzido em dois dias. O primeiro dia foi destinado à prática de exercícios com os dois métodos USARP e ACUX, visando minimizar as ameaças de aprendizado. Neste dia, também foi realizado a caracterização do perfil dos estudantes por meio de um questionário, no qual foi questionado o nível de conhecimento sobre os assuntos relacionados ao experimento e sobre a experiência no mercado de trabalho na área de TI. A partir destas respostas, as equipes seriam balanceadas para o segundo dia de experimento. A Figura 6 apresenta o roteiro dos dois dias de execução do experimento em sala de aula.

**Figura 6: Roteiro de execução do experimento**

No segundo dia de experimento foi entregue um cenário de problema de software, acompanhado dos artefatos destinados a facilitar o processo de eliciação. As equipes receberam ou as cartas

da USARP ou os cartões do ACUX impressos, conforme designação prévia.

Os questionários de Caracterização do perfil e Coleta de Feedback foram disponibilizados de forma *online* e elaborados na plataforma Google Formulários.

## 6 ANÁLISE DOS DADOS E RESULTADOS

Nesta seção, serão descritas as técnicas utilizadas para a análise dos dados obtidos do experimento, bem como os resultados. As abordagens principais incluíram a Análise Quantitativa e a Análise Qualitativa dos artefatos especificados e percepções obtidas dos participantes. As ferramentas utilizadas foram JASP<sup>3</sup> para análise estatística, ATLAS.Ti<sup>4</sup> para análise qualitativa, Flourish<sup>5</sup> para visualização de dados e Google Planilhas<sup>6</sup> para tabulação dos dados.

### 6.1 Análise quantitativa

No total, 32 de 37 respostas registradas no questionário foram consideradas para realizar os testes estatísticos. Os dados removidos incluem: 1 participante que declarou não desejar participar da pesquisa, conforme opção do Termo de Consentimento Livre e Esclarecido (TCLE); e 4 respostas de participantes da equipe que seria descartada da análise por não participarem do treinamento.

**6.1.1 Eficácia.** A abordagem definida para avaliar a eficácia dos métodos consistiu em calcular a taxa de acerto da equipe. Esse cálculo foi realizado dividindo o número de Requisitos ou ACs corretamente especificados pelo total previsto no oráculo de cada US. Durante a análise dos artefatos, foi identificado em ambos os métodos, a presença de Requisitos e ACs adequados no contexto do cenário de software, mas que não foram contemplados no oráculo. Em resposta, estes foram incorporados ao oráculo como parte do processo de correção e aprimoramento. A Tabela 3 mostra a taxa de acerto de cada equipe, estabelecendo o nível de eficácia dos métodos neste experimento.

Observa-se que os métodos USARP e ACUX apresentaram resultados semelhantes em relação à média de eficácia. Dessa forma, o número médio de artefatos especificados corretamente ficou em torno de 9 a 10 Requisitos/ACs para uma sessão de brainstorming com 2 USs.

<sup>3</sup><https://jasp-stats.org/>

<sup>4</sup><https://jasp-stats.org/>

<sup>5</sup><https://flourish.studio/>

<sup>6</sup><https://docs.google.com/spreadsheets>

**Tabela 3: Eficácia dos métodos**

Método	Oráculo	Equipe	Quantidade de artefatos corretos	Eficácia
USARP	38	E1	9	24%
		E2	4	11%
		E3	16	42%
		Média:	9,67	25,67%
ACUX	43	E4	18	42%
		E5	5	12%
		E6	9	21%
		Média:	10,7	25%

**6.1.2 Eficiência.** A abordagem definida para avaliar a eficiência dos métodos consistiu em calcular a taxa de especificação de artefatos na atividade de refinamento das USs. Esse cálculo foi realizado dividindo o número de Requisitos ou ACs corretamente especificados pelo tempo, em horas, do início da atividade até a entrega da solução. A Tabela 4 mostra o tempo decorrido e a taxa de especificação de artefatos de cada equipe, estabelecendo o nível de eficiência dos métodos neste experimento.

**Tabela 4: Eficiência dos métodos**

Método	Equipe	Hora de entrega	Tempo de especificação	Eficiência (Artefatos/h)
Hora de início:		10:30:00		
USARP	E1	11:55:00	01:25:00	6,35
	E2	11:55:00	01:25:00	2,82
	E3	12:00:00	01:30:00	10,00
	Média:		1h 27min	6,39
ACUX	E4	11:56:00	01:26:00	12,56
	E5	11:35:00	01:05:00	8,31
	E6	11:38:00	01:08:00	4,41
	Média:		1h 13min	8,43

O método USARP apresentou um tempo médio de especificação maior, estabelecido pela média de 1 hora e 27 minutos. Como resultado, a eficiência do ACUX, medida em artefatos por hora, mostrou-se superior ao método USARP uma vez que os participantes produziram mais artefatos em menos tempo.

**6.1.3 Percepção de Facilidade de Uso.** Ao examinar o gráfico de PFU da USARP, conforme ilustrado na Figura 7, constata-se que a maioria dos participantes expressa concordância (Concordo e Concordo Totalmente) em relação à clareza e compreensibilidade da interação com a USARP, especialmente no que diz respeito à elicitación de requisitos de usabilidade e facilidade de uso. Entretanto, ao analisar os dados associados ao ACUX, observa-se uma distribuição mais equilibrada das respostas, com um número considerável de participantes posicionados de forma neutra em relação à facilidade de uso, clareza da interação e esforço mental. Estes resultados sugerem que os participantes tendem a perceber a USARP como sendo mais clara e fácil de usar em comparação ao ACUX.

**6.1.4 Percepção de Utilidade.** Ainda na Figura 7, nos gráficos que retratam a PU das técnicas USARP e ACUX no contexto do refinamento de USs, evidencia-se que ambas as técnicas são consideradas úteis pelos participantes. Há uma forte concordância de opiniões, tanto para a USARP quanto para o ACUX, no sentido de que essas técnicas contribuem para a melhoria do desempenho, aumento da produtividade e da eficácia no processo de desenvolvimento de USs. Contudo, a USARP demonstra uma vantagem marginal, apresentando uma proporção menor de respostas neutras em comparação com o ACUX. Este cenário sugere que os participantes têm uma percepção mais positiva e confiante na utilidade da USARP em comparação com o ACUX.

## 6.2 Teste de Hipóteses

O teste de hipóteses permite determinar se as diferenças observadas nos dados são estatisticamente significativas ou se podem ser atribuídas ao acaso [46]. O teste estatístico de Shapiro-Wilk foi empregado para verificar se o conjunto de dados segue uma distribuição normal (Tabela 5), destacando-se por sua eficácia em amostras pequenas [41].

**Tabela 5: Teste de normalidade (Shapiro-Wilk)**

		W	p
Facilidade de Uso (PFU)	USARP	0.919	0.160
	ACUX	0.909	0.114
Utilidade (PU)	USARP	0.855	0.016
	ACUX	0.947	0.450
Eficácia	USARP	0.787	0.002
	ACUX	0.770	0.001
Eficiência	USARP	0.787	0.002
	ACUX	0.804	0.003

*Nota.*  $p < 0.05$  (nível de significância), rejeita-se a hipótese nula, indicando evidências de que os dados não seguem uma distribuição normal.

Os valores de  $p$  associados à variável PFU foram 0.160 e 0.114 para a USARP e para a ACUX, respectivamente. Estes valores foram os únicos a passar no teste de normalidade, que indica a ausência de desvios significativos da normalidade, uma vez que a hipótese nula não foi rejeitada ( $p > 0.05$  não rejeita  $H^0$ ) [13]. Dessa forma, o teste paramétrico  $t$  de Student, também conhecido como teste  $t$ , foi aplicado na análise dos dados de PFU. Essa técnica é utilizada para comparar as médias de dois grupos amostrais independentes com distribuição normal.

Ao conduzir o teste correspondente com um nível de significância de 5%, a análise da Tabela 6 revela um valor de  $p$  igual a 0,039, resultando na rejeição da hipótese nula ( $p < 0,05$ ). Essa hipótese nula afirmava que os níveis de facilidade de uso eram equivalentes do ponto de vista estatístico. Dessa forma, constata-se a presença de diferenças significativas entre as amostras da USARP e ACUX.

A mediana é frequentemente empregada como medida de referência em testes não paramétricos, juntamente com medidas percentis, devido à sua capacidade de lidar adequadamente com a assimetria dos dados [15]. Dessa forma, ao examinar a Tabela 7, constata-se

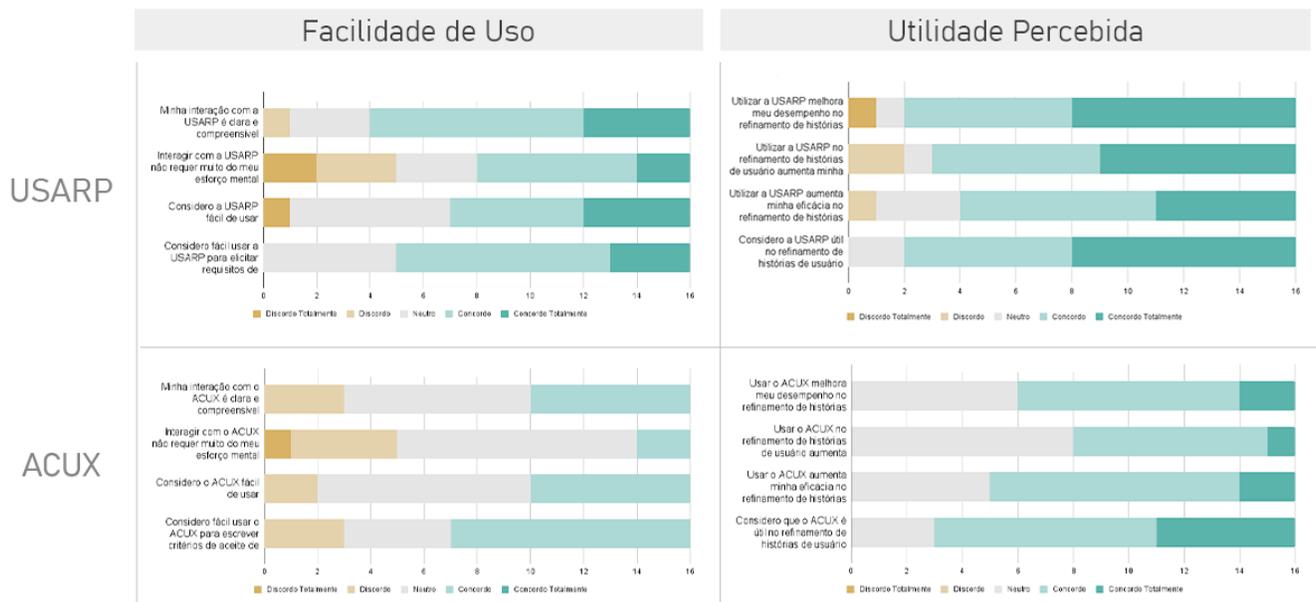


Figura 7: Percepção de Facilidade de Uso e Utilidade - Gráfico de barras empilhadas

Tabela 6: Teste t de Student - PFU

	t	df	p
Facilidade de Uso	2.164	30	0.039

que a mediana das pontuações atribuídas às respostas relacionadas à eficácia da USARP supera aquela associada à ACUX. Portanto, há evidências estatisticamente suficientes para afirmar que a USARP é percebida como mais fácil de utilizar que a ACUX.

Tabela 7: Mediana e Percentis da PFU

	Facilidade de Uso	
	USARP	ACUX
N	16	16
Mediana	3.750	3.000
Perc 25	3.500	2.875
Perc 75	4.063	3.750

Para analisar as variáveis que apresentaram a possibilidade de não normalidade dos dados, foi utilizado o teste não paramétrico de Mann-Whitney. Esse teste é uma alternativa ao teste t em situações que a normalidade ou a homogeneidade de variâncias não pode ser assumida [13, 46].

Ao conduzir o teste de Mann-Whitney para os fatores PU, Eficácia e Eficiência com um nível de significância de 5%, os resultados revelaram valores de p iguais a 0,074, 0,893 e 0,329 (Tabela 8). Nesse contexto, a não rejeição da hipótese nula ( $p > 0,05$ ) sugere que não há variações estatisticamente significativas entre as amostras.

Portanto, observa-se uma equivalência estatística na Percepção de Utilidade, Eficácia e Eficiência dos métodos USARP e ACUX.

Tabela 8: Teste de Mann-Whitney - PU, Eficácia e Eficiência

	W	df	p
Utilidade	175.500		0.074
Eficácia	132.000		0.893
Eficiência	102.000		0.329

### 6.3 Análise qualitativa de codificação fechada

De acordo com Gibbs [12], a Análise Qualitativa pode ser conduzida em dois níveis de profundidade. No primeiro nível, denominado de análise de codificação aberta ou indutiva, os dados são comparados e códigos são atribuídos para explicar um conjunto de dados. O pesquisador inicia o processo sem um conjunto prévio de códigos ou conceitos, mantendo-se aberto a novas direções teóricas à medida que avança na análise. No segundo nível, o pesquisador organiza um conjunto de códigos com conceitos pré-definidos e a marcação é orientada por um conteúdo anterior. Esse segundo nível, aplicado neste trabalho, é conhecido como análise de codificação fechada ou dedutiva [12, 26]. Nesta seção, será fornecida evidências experimentais sobre outros aspectos identificados nos artefatos gerados pelos participantes.

Para traçar algum tipo de similaridade entre os métodos, como também, o poder de cobertura do seus artefatos, o *framework* de Hassenzahl [14] foi selecionado como base para identificar características de UX nos requisitos e critérios de aceite especificados pelos participantes. Ele define um modelo que diz que a UX é formalizada

por três níveis de ação de interação, "o quê", "por quê" e "como", e um "objetivo", compreendido como o alcance do bem-estar do usuário. O nível "por quê" centra-se no significado que a interação pode proporcionar aos usuários, abordando as motivações e necessidades que impulsionam o usuário a utilizar o produto. Nesta análise, optou-se por adotar os níveis "o quê" e "como", uma vez que estes apresentam uma alta capacidade de rastrear a especificação de requisitos funcionais e as interações do usuário com o sistema nos artefatos produzidos pelos participantes. A Tabela 9 apresenta os códigos e suas definições, dos quais as pesquisadoras utilizaram para realizar a análise de codificação fechada.

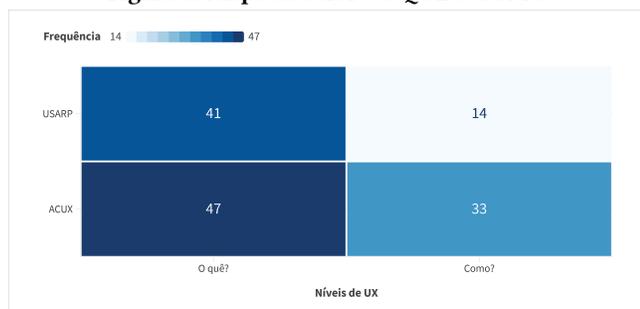
**Tabela 9: Livro de Códigos**

Níveis de UX	Definições
O quê?	Define as funcionalidades que o produto oferece aos usuários para atender suas necessidades.
Como?	Explora ações concretas dos usuários para interagirem com o produto, por exemplo, clicar em um botão ou ler instruções.

A análise transcorreu em duas etapas primordiais: a análise individual e o alinhamento das soluções. Na primeira etapa, duas pesquisadoras receberam uma cópia impressa contendo todos os artefatos especificados pelas equipes, que incluíam histórias de usuário e critérios de aceite. A tarefa de cada pesquisadora consistia em destacar trechos nos dados e atribuir códigos que melhor representassem as extrações. A primeira pesquisadora é concluinte em Engenharia de Software com experiência de 1 ano e 8 meses como Analista de Qualidade e, a segunda pesquisadora é docente, com doutorado em Informática e uma vasta experiência em análise qualitativa de dados em Engenharia de Software.

Após a conclusão da análise individual, as pesquisadoras reuniram-se para discutir os trechos extraídos por cada uma. Nessa segunda etapa, observou-se poucas discordâncias entre as duas análises, sendo estas resultantes de interpretações distintas entre as pesquisadoras. Tais divergências foram discutidas e chegou-se a um consenso. Ao término do alinhamento, um documento contendo uma solução consensual foi gerado e transferido para a ferramenta ATLAS.ti, como ilustrado na Figura 8.

**Figura 8: Mapa de calor - O QUÊ e COMO**



No mapa de calor da Figura 8, observa-se que ambos os métodos apresentaram números semelhantes no nível de UX "O quê". Isso

indica que ambos os métodos possuem uma alta capacidade de especificar funcionalidades a serem implementadas em um software.

Quanto ao nível de UX "Como", o método ACUX demonstra uma maior capacidade de especificar ações de interação do usuário com o software, apresentando uma diferença significativa em comparação com o método USARP. Isso pode ser explicado pela estrutura base Dado-Quando-Então [33] utilizada pela ACUX para especificar critérios de aceite. Nesse modelo, o leitor é guiado a fazer uma declaração do tipo: "Dado [um contexto ou precondição] Quando [realizo uma ação] Então [tenho um resultado esperado]". Dessa forma, a seção "Quando" induz o leitor a declarar interações específicas do usuário com o sistema em todas as especificações.

#### 6.4 Análise qualitativa de codificação aberta

Este tópico se concentra na análise dos comentários fornecidos de forma voluntária pelos participantes no Questionário de Feedback, após o uso da USARP e do ACUX. Ao explorar as preferências e desafios destacados pelos participantes, busca-se fornecer uma perspectiva que sirva como base para aprimorar a aplicabilidade e a eficácia dos métodos. Para isso, duas perguntas-chave orientaram nossa investigação, e são elas:

**6.4.1 Pergunta: O que você mais gostou no método?** Vários comentários dos participantes expressavam que os métodos eram **fáceis de utilizar e práticos**. Algumas citações podem reforçar esse resultado, como ["A facilidade" – P7, USARP] e ["Praticidade para desenvolver os critérios de aceitação" – P26, ACUX]. Outra prática bem avaliada por eles foi a **interação entre o time**, que pode ser destacada nos comentários ["Acredito que o fato desse método nos fazer interagir e discutir o que seria ou não necessário para o sistema..." – P8, USARP] e ["A interação e a discussão para decidir quais [diretrizes] usar no modelo [US] específico" – P27, ACUX].

Além dessas, os participantes também avaliaram positivamente o **fornecimento de exemplos e padrões de escrita**, o **uso de cartas/cartões**, a **boa documentação dos métodos**, o **apoio a melhoria da qualidade do software** e o **aumento da compreensão no assunto**.

**6.4.2 Pergunta: Você sentiu dificuldade em utilizar o método? Se sim, em qual ou quais partes do processo?** Houveram comentários que relatavam **dificuldade de interpretação de algumas cartas/cartões**, são citações ["Na clareza de algumas ações [mecanismos] que o sistema desempenhava" – P5, USARP] e ["Senti dificuldade em interpretar o que estava descrito nos critérios de aceitação." – P25, ACUX]. Outra dificuldade citada foi quanto a **identificação dos Requisitos ou ACs adequados** para a US, encontrada nos comentários ["Sim, para identificar as USs que se encaixam com o cartão." – P33, ACUX] e ["[outra dificuldade] e do que faz sentido para as USs" – P10, USARP].

Outras barreiras citadas pelos participantes foram **dificuldade inicial no processo** que no decorrer da atividade foi reduzida, **algumas diretrizes do ACUX que pareciam ambíguas ou parciais**, o **baixo conhecimento do assunto**, a **falta de prática** e a **quantidade de informações e descrição da carta do método USARP**.

## 6.5 Ameaças à validade

A experimentação de software com humanos requer cuidado, planejamento meticuloso e a consideração de uma série de fatores para garantir a confiabilidade dos resultados obtidos [46]. As principais ameaças à validade e seus tratamentos foram:

*Validade interna:* Para reduzir a ameaça de falta de motivação para resolução da atividade, as equipes com melhor desempenho nesta atividade e a equipe mais engajada foram premiadas com uma pontuação extra, como parte da metodologia de gamificação aplicada na disciplina. Houveram estudantes que não participaram do treinamento antecipado de prática dos métodos. Estes participantes compuseram a Equipe 7, a qual foi desconsiderada durante a análise dos resultados. Os estudantes não foram sinalizados sobre essa particularidade e realizaram a atividade normalmente. Com isso, o número de equipes por método durante a análise ficou equilibrado.

*Validade externa:* A maior ameaça à validade externa é o uso de estudantes como participantes. No entanto, essa ameaça é reduzida ao selecionar estudantes do quarto semestre ou que estão no último ano de conclusão, considerados engenheiros novatos [37], além dos quais alguns já iniciaram atividade na indústria.

*Validade de conclusão:* A normalidade das amostras foi verificada para a seleção dos testes estatísticos adequados. Porém, o tamanho da amostra reduziu o poder do teste para rejeitar as hipóteses nulas de eficácia, eficiência e utilidade percebida. Uma solução seria a adoção do design pareado, porém, devido à programação das aulas da disciplina, o experimento deveria ser realizado em dois dias e optou-se utilizar um dos dias para treinamento visando nivelar o conhecimento dos participantes.

*Validade de constructo:* O cenário proposto no experimento pode ser relativamente breve e a quantidade de USs em projetos reais é significativamente maior. Entretanto, em projetos ágeis, os times de desenvolvimento costumam trabalhar com subconjuntos de USs, similar a como foi realizado no experimento.

## 7 DISCUSSÃO

Os resultados do experimento indicaram que a percepção sobre a facilidade de uso da USARP foi maior em relação à ACUX. Por outro lado, a ACUX possibilitou especificar mais informações sobre ações dos usuários ao interagirem com o sistema.

As diretrizes da ACUX indicam os aspectos a serem especificados e utiliza o padrão Gherkin para a especificação dos critérios de aceitação. Os participantes podem ter associado um maior nível de esforço para a adoção da ACUX, uma vez que a USARP não define um padrão, mas sugere um texto para especificação do aspecto de usabilidade.

Embora ambos os métodos (ACUX e USARP) permitam descrever as funcionalidades que o sistema deve fornecer para atender aos aspectos de usabilidade e UX, o padrão Gherkin adotado pela ACUX permite detalhar interações, o que favorece a completude dos requisitos de usabilidade e UX. Os métodos especificam diferentes campos da história de usuário, o que favorece o seu uso conjunto e potencializa os benefícios de sua utilização.

Com base na revisão da literatura realizada e descrita na subseção 2.2, este é o primeiro experimento relatado que comparou dois métodos para elicitación e especificação de requisitos de usabilidade e UX, explorando seus benefícios e oportunidades de melhoria. Os

métodos propostos em outras pesquisas como Keshk et al. [21] e Sousa et al. [44] não focavam em requisitos de usabilidade e UX, enquanto Pereira et al. [35] focava apenas na especificação. Assim, este estudo contribui para o avanço da integração entre Engenharia de Requisitos e UX.

A amostra adotada neste experimento não permitiu identificar diferença significativa entre os métodos em relação à utilidade percebida, eficácia e eficiência. Estas hipóteses podem ser investigadas em replicações deste experimento. Outro aspecto não investigado é em relação à utilidade das informações fornecidas pelos diferentes métodos do ponto de vista de profissionais de diferentes times como desenvolvimento, design e qualidade.

## 8 CONCLUSÃO E TRABALHOS FUTUROS

Este artigo investigou o uso de métodos para dar suporte a elicitación e especificação de requisitos de UX e de usabilidade. Um experimento controlado comparou a eficácia, eficiência, percepção de facilidade de uso e utilidade percebida dos métodos USARP e ACUX. Uma descoberta relevante foi alcançada por meio do teste de hipótese paramétrico, o qual demonstrou que a USARP foi percebida com um nível mais elevado de facilidade de uso em comparação com o ACUX. Por outro lado, não foram observadas diferenças significativas entre os métodos em relação aos aspectos de eficácia, eficiência e utilidade percebida pelos usuários.

A partir de dados qualitativos sobre a opinião dos participantes em relação aos dois métodos, observou-se que os métodos são fáceis de usar e práticos, os exemplos fornecidos e padrões de escrita apoiam sua adoção. Dificuldades foram percebidas na interpretação de algumas orientações descritas pelos métodos e no mapeamento das diretrizes adequadas a cada requisito. Além disso, outras características subjetivas foram avaliadas, como a capacidade dos métodos de especificar funcionalidades do sistema e descrever ações concretas de interação do usuário, através do *framework* de Hassenzahl. A ACUX permitiu especificar ações concretas dos usuários durante a interação com o software, o que não é muito apoiado pela USARP. Ambos os métodos permitem especificar as funcionalidades que o software deve fornecer aos usuários.

Este artigo apresenta contribuições relevantes para área de engenharia de requisitos e qualidade de software por investigar o uso de métodos que auxiliam a especificação de requisitos de UX e usabilidade. A investigação mostra resultados que apontam as potencialidades de cada método, e também suas lacunas. Como trabalho futuro sugere-se investigar o uso conjunto desses métodos buscando captar o melhor dos dois métodos. Também sugere-se a replicação deste experimento com uma amostra maior de participantes e com profissionais de desenvolvimento de software de diferentes áreas como qualidade, desenvolvimento e design.

O balanceamento das equipes para o experimento foi uma etapa difícil do estudo. A falta de determinados dados sobre as características dos participantes foi o principal obstáculo encontrado.

## AGRADECIMENTOS

A equipe de autoria agradece ao apoio da Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP BP5-0197-00016.01.00/22). O estudo foi desenvolvido com apoio parcial

do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 309497/2022-1).

## REFERÊNCIAS

- [1] ISO 25010:2011. 2011. Systems and software Quality Requirements and Evaluation (SQuaRE). *ISO/IEC 25010:2011 1* (2011), 34. <https://doi.org/10.1109/IEEESTD.2017.8016712>
- [2] S. Barbosa and B. Silva. 2010. *Interação Humano-Computador*. Elsevier Brasil.
- [3] V.R. Basili and H.D. Rombach. 1988. The TAME project: towards improvement-oriented software environments. *IEEE Transactions on Software Engineering* 14, 6 (1988), 758–773. <https://doi.org/10.1109/32.6156>
- [4] K. Beck and C. Andres. 2004. *Extreme Programming Explained: Embrace Change*. Pearson Education.
- [5] Johan Kaj Blomkvist, Johan Persson, and Johan Åberg. 2015. Communication through boundary objects in distributed agile teams. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1875–1884.
- [6] M. Cohn. 2004. *User Stories Applied: For Agile Software Development*. Addison-Wesley.
- [7] Mike Cohn. 2004. *User Stories Applied: For Agile Software Development*. Addison Wesley Longman Publishing Co., Inc., USA.
- [8] Karina Curcio, Tiago Navarro, Andreia Malucelli, and Sheila Reinehr. 2018. Requirements engineering: A systematic mapping study in agile software development. *Journal of Systems and Software* 139 (2018), 32–50.
- [9] Karina Curcio, Rodolfo Santana, Sheila Reinehr, and Andreia Malucelli. 2019. Usability in agile software development: A tertiary study. *Computer Standards Interfaces* 64 (2019), 61–77. <https://doi.org/10.1016/j.csi.2018.12.003>
- [10] Bruna Ferreira, Tayana Conte, and Simone Diniz Junqueira Barbosa. 2015. Eliciting requirements using personas and empathy map to enhance the user experience. In *2015 29th Brazilian Symposium on Software Engineering*. IEEE, 80–89.
- [11] J.J. Garrett. 2003. *The Elements of User Experience: User-centered Design for the Web*. New Riders.
- [12] Graham Gibbs. 2009. *Análise de dados qualitativos: Coleção Pesquisa Qualitativa*. Bookman.
- [13] Mark Goss-Sampson. 2021. *Statistical analysis in JASP: a guide for students* (4 ed.). JASP. <https://doi.org/DOI:10.6084/m9.figshare.9980744> Versão em português..
- [14] Marc Hassenzahl. 2018. *The Thing and I (Summer of '17 Remix)*. Springer International Publishing, Cham, 17–31. [https://doi.org/10.1007/978-3-319-68213-6\\_2](https://doi.org/10.1007/978-3-319-68213-6_2)
- [15] Heitor Marques Honório. 2019. *Teste de Mann-Whitney (SigmaPlot 12.0)*. <https://youtu.be/xnfpghqg0sw> Fornece dicas de análise dos dados resultantes do teste..
- [16] ISO-24765. 2017. ISO/IEC/IEEE 24765:2017 Systems and software engineering Vocabulary. *ISO/IEC/IEEE 24765:2017(E)* (2017), 1–541. <https://doi.org/10.1109/IEEESTD.2017.8016712>
- [17] R. Jeffries. 2001. Essencial XP: card, conversation and confirmation. *XP Magazine* 30 (2001).
- [18] James Garrett Jesse. 2011. The elements of user experience: User-centered design for the web and beyond.
- [19] Natalia Juristo, Ana Moreno, and Maria-Isabel Sanchez-Segura. 2007. Guidelines for Eliciting Usability Functionalities. *IEEE Transactions on Software Engineering* 33, 11 (2007), 744–758. <https://doi.org/10.1109/TSE.2007.70741>
- [20] Steven Loh Mun Keong and Zarina Che Embi. 2022. A Systematic Review on Non-Functional Requirements Documentation in Agile Methodology. *Journal of Informatics and Web Engineering* 1, 2 (2022), 19–29.
- [21] Nader Keshk, Mohammad El-Ramly, and Akram Salah. 2022. A Proposal for Enhancing Agile Requirements Engineering with Prototyping and Enriched User Stories. In *Proceedings of the Federated Africa and Middle East Conference on Software Engineering (Cairo-Kampala, Egypt) (FAMECSE '22)*. Association for Computing Machinery, New York, NY, USA, 59–63. <https://doi.org/10.1145/3531056.3542773>
- [22] S. Krug. 2013. *Don't Make Me Think, Revisited: A Common Sense Approach to Web Usability*. Pearson Education.
- [23] Anna Beatriz Marques, Alex Felipe Costa, Ismayle Santos, and Rossana Andrade. 2022. Enriching User Stories with Usability Features in a Remote Agile Project: A Case Study. In *Proceedings of the XXI Brazilian Symposium on Software Quality (Curitiba, Brazil) (SBQS '22)*. Association for Computing Machinery, New York, NY, USA, Article 24, 10 pages. <https://doi.org/10.1145/3571473.3571496>
- [24] Anna Beatriz Marques and Maria Victoria Fiori. 2023. Evolving the USARP Method to Support Usability Requirements Elicitation and Specification. (2023), 14 pages. <https://doi.org/10.29327/1298356.26-3>
- [25] Anna Beatriz Marques, Alex Alan Santos, Maria Victoria Fiori, Natalia Coelho, and Victor Feitosa. 2022. Integrando técnicas de IHC e Engenharia de Software na especificação de requisitos de uma ferramenta de modelagem. In *Anais Estendidos do XXI Simpósio Brasileiro de Fatores Humanos em Sistemas Computacionais*. SBC, 40–47.
- [26] S Martinelli, J Choma, J Saad, and L Zaina. 2023. Análise Qualitativa em IHC: da codificação a criação de visualizações. , 72 pages.
- [27] Suellen Martinelli, Nicolas Nascimento, Jonathan Souza, Afonso Sales, and Luciana Zaina. 2022. UX Requirements Matters: Guidelines to Support Software Teams on the Writing of Acceptance Criteria. In *Proceedings of the XXXVI Brazilian Symposium on Software Engineering*. 398–408.
- [28] Suellen Martinelli, Nicolas Nascimento, Jonathan Souza, Afonso Sales, and Luciana A. M. Zaina. 2022. UX Requirements Matters: Guidelines to Support Software Teams on the Writing of Acceptance Criteria. In *SBES 2022: XXXVI Brazilian Symposium on Software Engineering, Virtual Event Brazil, October 5 - 7, 2022*, Marcelo de Almeida Maia, Fabiano A. Dorça, Rafael Dias Araújo, Christina von Flach, Elisa Yumi Nakagawa, and Edna Dias Canedo (Eds.). ACM, 398–408. <https://doi.org/10.1145/3555228.3555230>
- [29] Ana M Moreno and Agustín Yagüe. 2012. Agile user stories enriched with usability. In *Agile Processes in Software Engineering and Extreme Programming: 13th International Conference, XP 2012, Malmö, Sweden, May 21–25, 2012. Proceedings* 13. Springer, 168–176.
- [30] J. Nielsen. 1994. *Usability Engineering*. Elsevier Science.
- [31] D.A. Norman. 2004. *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Books.
- [32] D.A. Norman and A. Deiró. 2006. *O design do dia-a-dia*. Rocco.
- [33] Dan North. 2006. Behavior Modification. *Better Software Magazine* 2006-03 (2006).
- [34] Gabriel F. de Oliveira, Bruna Ferreira, and Anna Beatriz Marques. 2020. USARP Method: Eliciting and Describing Usability Requirements with Personas and User Stories. In *Proceedings of the XXXIV Brazilian Symposium on Software Engineering (Natal, Brazil) (SBES '20)*. Association for Computing Machinery, New York, NY, USA, 437–446. <https://doi.org/10.1145/3422392.3422435>
- [35] Anathan Pereira, Abner Cleto Filho, Eduardo Guerra, and Luciana Zaina. 2021. Towards a Pattern Language to Embed UX Information in Agile Software Requirements. In *26th European Conference on Pattern Languages of Programs*. 1–8.
- [36] Roger S. Pressman. 2011. *Engenharia de Software: uma abordagem profissional* (7ª ed. ed.). Bookman.
- [37] Luis Rivero and Tayana Conte. 2016. How novice software engineers apply user interface design patterns: an empirical study. In *Proc Int Conf Softw Eng Knowl Eng SEKE*. 600–604.
- [38] Y. Rogers, H. Sharp, and J. Preece. 2013. *Design de Interação - 3ed*. Bookman Editora.
- [39] Alex Alan Santos, Maria Elanne Mendes, and Anna Beatriz Marques. 2022. Inserindo um olhar de IHC no ensino de Engenharia de Requisitos: um relato de experiência. In *Anais do XIII Workshop sobre Educação em IHC*. SBC, 13–18.
- [40] Eva-Maria Schön, Jörg Thomaschewski, and Maria José Escalona. 2017. Agile Requirements Engineering: A systematic literature review. *Computer Standards Interfaces* 49 (2017), 79–91. <https://doi.org/10.1016/j.csi.2016.08.011>
- [41] Tassio Ferenzini Martins Sirqueira, Marcos Alexandre Miguel, Humberto Luiz de Oliveira Dalpra, and Marco Antônio Pereira Araújo. 2020. Aplicação de Métodos Estatísticos em Engenharia de Software: Teoria e Prática. 18 (2020), 228–246.
- [42] I. Sommerville. 2011. *Engenharia de software*. Pearson Prentice Hall.
- [43] I. Sommerville and P. Sawyer. 1997. *Requirements Engineering: A Good Practice Guide*. Wiley.
- [44] Gabriel Ferreira Barros de Sousa and Simone da Silva Amorim. 2021. Speech-To-Story: Gerando Histórias de Usuário. In *Proceedings of the XXXV Brazilian Symposium on Software Engineering*. 309–314.
- [45] Viswanath Venkatesh and Hillol Bala. 2008. Technology acceptance model 3 and a research agenda on interventions. *Decision sciences* 39, 2 (2008), 273–315.
- [46] C. Wohlin, P. Runeson, M. Höst, M.C. Ohlsson, B. Regnell, and A. Wesslén. 2012. *Experimentation in Software Engineering*. Springer Berlin Heidelberg.
- [47] Luciana AM Zaina, Helen Sharp, and Leonor Barroca. 2021. UX information in the daily work of an agile team: A distributed cognition analysis. *International journal of human-computer studies* 147 (2021), 102574.

## A DISPONIBILIDADE DE ARTEFATOS

Os materiais utilizados no experimento estão disponíveis em um repositório e anonimizados para esta submissão: <https://zenodo.org/records/10995302>.