

Identificando Oportunidades de Melhoria em Tecnologias de Inspeção de Usabilidade para o Contexto Web Móvel através de um Experimento Controlado

Guto Kawakami², Luis Rivero^{1,2} e Tayana Conte^{1,2}

guto.kawakami@gmail.com, {luisrivero, tayana}@icomp.ufam.edu.br

¹Grupo de Usabilidade e Engenharia de Software (USES)

²Instituto de Computação – Universidade Federal do Amazonas (UFAM)

Abstract. Um software com alto grau de usabilidade pode ser entendido e utilizado com maior facilidade pelos usuários. No caso específico de aplicações Web móveis, a usabilidade pode fazer a diferença entre o uso ou abandono da aplicação pelos usuários finais. Neste artigo, descreve-se um conjunto de tecnologias (técnicas e ferramenta) de inspeção de usabilidade que podem ser aplicadas por engenheiros de software para avaliar aplicações Web durante o seu design, diminuindo o custo de correção de problemas antes da escrita do código fonte da aplicação. Para identificar oportunidades de melhoria destas tecnologias para o contexto Web móvel, foi executado um estudo experimental comparando o seu desempenho e coletando dados sobre os fatores que dificultam a sua utilização. Espera-se com essas medidas adequar as tecnologias avaliadas ao contexto Web móvel, encorajando a indústria a utilizá-las no início do processo de desenvolvimento, garantindo a qualidade do software desenvolvido.

Keywords: Aplicações Web Móveis, Inspeção de Usabilidade, Mockups, Estudo Experimental

1 Introdução

Uma aplicação Web móvel deve executar em dispositivos, cujas dimensões e processamento são relativamente menores que as do computador convencional [17]. Aplicações Web móveis consideram, portanto, características específicas dos dispositivos, as limitações do ambiente de comunicação sem fio e o contexto de uso dinâmico durante o seu desenvolvimento [20]. Estes fatores podem influenciar na facilidade de uso da aplicação e por isso garantir a usabilidade torna-se fundamental para atingir a qualidade no produto de software. Nesse contexto, usabilidade é a capacidade do produto de software de ser entendido, aprendido, operado, atraente para o usuário e aderente a padrões/guias de usabilidade, quando usado sob condições específicas [9].

Segundo Bastien [1], o método de avaliação mais utilizado para identificar problemas de usabilidade em aplicações Web móveis é o teste de usabilidade, que foca na interação direta dos usuários com o sistema. No entanto, o custo deste tipo de avaliação é alto devido aos equipamentos e laboratórios necessários para sua aplicação [15].

adfa, p. 1, 2011.

© Springer-Verlag Berlin Heidelberg 2011

Para apoiar a avaliação de usabilidade a um baixo custo, a indústria e a academia têm proposto métodos de inspeção que permitam a detecção de problemas de usabilidade e a análise de aplicações Web móveis na real situação de uso [4].

Em uma pesquisa sobre características indispensáveis em novas propostas de métodos de inspeção de usabilidade, identificou-se a necessidade de permitir a inspeção de artefatos durante as primeiras etapas do processo de desenvolvimento [13]. Esta característica é necessária para que problemas de software sejam identificados antes da escrita do código fonte da aplicação, visto que o custo de correção de problemas chega a ser cem vezes maior no final do processo de desenvolvimento [5]. No entanto, mesmo com as tecnologias existentes, existe uma carência de métodos de inspeção que possam ser aplicados durante as etapas de projeto de aplicações Web móveis.

Este artigo apresenta um estudo experimental para avaliar a adequação de técnicas de inspeção de usabilidade no contexto de aplicações Web móveis. Neste estudo, foram comparadas duas técnicas: a Avaliação Heurística (AH) [12] e a técnica *Web Design Usability Evaluation* (Web DUE) [13]; ao serem aplicadas para identificar problemas de usabilidade em uma aplicação Web móvel. As duas técnicas foram escolhidas, pois podem ser aplicadas na avaliação de usabilidade de protótipos de baixa fidelidade (ou mockups) a um baixo custo e esforço, através da utilização da ferramenta Mockup DUE. A ferramenta disponibiliza funcionalidades como mapeamento de mockups e geração automática de relatórios que diminuem a carga cognitiva em cima dos inspetores para que estes foquem na identificação de defeitos [13]. O estudo permitiu coletar dados sobre: (a) o desempenho das técnicas de inspeção em termos de eficácia e eficiência, (b) a opinião dos inspetores sobre a adequação das mesmas, e (c) problemas na aplicação das tecnologias no contexto de inspeções de usabilidade de aplicações Web móveis.

O restante deste trabalho está dividido em outras seis seções. A Seção 2 apresenta os trabalhos relacionados, e a motivação para esta pesquisa. Na Seção 3, são descritas as tecnologias de inspeção aplicadas no estudo experimental. A Seção 4 discorre sobre o planejamento do estudo experimental realizado enquanto na Seção 5 são apresentados os resultados do mesmo. Além disso, a Seção 6 descreve as ameaças à validade do estudo e por fim as conclusões e perspectivas futuras são apresentadas na Seção 7.

2 Usabilidade em Aplicações Web Móveis

2.1 Avaliações de Usabilidade

Métodos de avaliação de usabilidade (*Usability Evaluation Methods* - UEMs) são procedimentos compostos por um conjunto definido de atividades que são usados para avaliar a usabilidade de uma aplicação [7]. Existem duas categorias de UEMs: testes com usuários e inspeções com especialistas [14]. Um teste de usabilidade é uma avaliação centrada no usuário, em que métodos de observação e entrevistas/questionários são usados para medir a usabilidade. Inspeções, por outro lado, são avaliações em que especialistas em usabilidade, ou até a equipe de desenvolvimento, verificam regras ou heurísticas de usabilidade em artefatos de software. Estas heurísticas permitem aos

inspetores verificar o grau de aderência do sistema com atributos de usabilidade e, com base nessa análise, prever se ocorrerá um problema de usabilidade.

Ambas as abordagens de avaliação de usabilidade tem suas vantagens e desvantagens. O teste com usuários, quando aplicado, permite a identificação de mais problemas de usabilidade que afetam os usuários finais [14]. Porém, Matera et al. [11] afirmam que ele possui um custo elevado, devido ao fato que é necessário realizar o teste com diferentes perfis de usuários. Além disso, o teste de usabilidade só pode ser aplicado quando já existe uma parte inicial do sistema pronta [14], dificultando sua execução nas primeiras etapas do processo de desenvolvimento. Por outro lado, os métodos de inspeção de usabilidade podem ser aplicados nas primeiras etapas do processo de desenvolvimento. Além disso, requerem menos recursos, podendo diminuir os custos relacionados à identificação de problemas de usabilidade no sistema [7].

2.2 Inspeção de Aplicações Web Móveis

Devido à necessidade de avaliar a usabilidade das aplicações Web móveis a um baixo custo, novos métodos de inspeção de usabilidade específicos para este tipo de aplicação têm surgido [3]. No artigo de Bonifácio et al. [4], os autores descrevem a técnica *Usability-Based Inspection CUsTomizable Approach* (UBICUA), que utiliza fatores específicos de aplicações Web móveis para orientar os inspetores durante o processo de avaliação. Os fatores: características do usuário, características do dispositivo e características do canal de comunicação, são empregados em conjunto com itens de verificação que detalham o que deve ser avaliado. A técnica UBICUA também é customizada dependendo do nível de experiência do inspetor, para que inspetores com alta, média e baixa experiência possam aplicá-la com diferentes níveis de detalhamento e assim facilitar a identificação de problemas de usabilidade.

Outra técnica de inspeção, chamada *Mobile Cognitive Heuristic*, foi proposta por Bertine et al. [2]. Esta técnica foi criada através da análise de problemas de usabilidade de aplicações Web móveis e a análise das heurísticas propostas por Nielsen [12] (ver Subseção 3.1) para gerar um novo conjunto de heurísticas adequado ao contexto móvel. Nesse sentido, a técnica possui um conjunto de heurísticas específicas como: facilidade de entrada de dados, leitura de tela e relance. Outras heurísticas foram modificadas, tais como: estética, privacidade e convenções sociais; e flexibilidade, eficiência de uso e personalização. Os inspetores podem utilizar a descrição destas heurísticas para identificar problemas de usabilidade em aplicações Web móveis.

2.3 Motivação para Avaliação de Mockups em Aplicações Web Móveis

Apesar da crescente demanda por métodos que permitam a avaliação de usabilidade de aplicações Web móveis, apenas 25% dos métodos propostos refere-se a inspeções que podem ser aplicadas por especialistas ou a equipe de desenvolvimento [3]. Além disso, Fernandez et al. [7] afirmam que perto de 91% dos métodos de avaliação de usabilidade são aplicados na etapa de implementação das aplicações. Nesse sentido, é necessário propor métodos de inspeção de usabilidade de aplicações Web móveis que

possam ser aplicados nas etapas iniciais do desenvolvimento, para assim diminuir os custos da correção dos problemas identificados [14].

No presente trabalho, para garantir a qualidade das aplicações Web móveis em termos de usabilidade a um baixo custo, optou-se pela inspeção de mockups, que são modelos de software que utilizam imagens para mostrar qual seria a aparência da aplicação quando o desenvolvimento fosse concluído (antes da escrita do código fonte). Ao aplicar inspeções em mockups espera-se diminuir os custos decorrentes tanto da avaliação quanto da correção dos problemas de usabilidade identificados. Nas próximas seções são descritas as tecnologias de inspeção avaliadas neste trabalho para identificar melhorias no contexto de avaliação de aplicações Web móveis.

3 Tecnologias de Inspeção de Usabilidade Avaliadas

Neste trabalho, foram escolhidas a técnica Avaliação Heurística [12] e a Técnica Web DUE [13] para a avaliação de usabilidade de mockups de aplicações Web móveis. A Avaliação Heurística foi escolhida devido à sua disseminação como técnica de inspeção e também devido ao fato que pode ser aplicada no início do processo de desenvolvimento [7]. Já a técnica Web DUE foi escolhida porque é específica para a avaliação de mockups de aplicações Web [13]. Nesta seção, descrevem-se as técnicas escolhidas e seu apoio ferramental, a ferramenta Mockup DUE.

3.1 Avaliação Heurística

A avaliação heurística é um método de inspeção proposto por Nielsen [12] baseado em diretrizes que tem como objetivo auxiliar o inspetor durante a avaliação. Para aplicar o método, um conjunto de avaliadores inspeciona a interface e detectam defeitos de usabilidade, caracterizando-os de acordo com uma lista composta por 10 heurísticas. A Tabela 1 apresenta as dez heurísticas propostas por Nielsen.

Tabela 1. Heurísticas propostas por Nielsen [12].

Nº	Heurística	Nº	Heurística
1	Visibilidade do status do sistema	6	Reconhecer ao invés de lembrar
2	Concordância do sistema com o mundo real	7	Flexibilidade e eficiência de uso
3	Controle e liberdade do usuário	8	Estética e design minimalista
4	Consistência e padrões	9	Ajudar os usuários a reconhecer, diagnosticar e corrigir erros
5	Prevenção de erros	10	Ajuda e documentação

3.2 A Técnica Web DUE

A técnica de inspeção *Web Design Usability Evaluation* (Web DUE), proposta por Rivero et al. [13], tem como objetivo orientar os inspetores no processo de avaliação de mockups de aplicações Web. A técnica sugere a divisão dos mockups em zonas de páginas Web que, segundo Fons et al. [8], referem-se a uma parte da página que tem

um tipo específico de conteúdo. Ao utilizar as zonas para direcionar os inspetores durante a avaliação, a técnica Web DUE faz com que o inspetor foque nos elementos que realmente estão presentes na aplicação, ou seja, os inspetores só deverão avaliar os atributos de usabilidade para as zonas que compõem o mockup avaliado.

O processo de avaliação da técnica Web DUE é composto de três etapas: dividir mockups em zonas de páginas web, avaliar itens de verificação e apontar problemas de usabilidade [13]. A Figura 1 apresenta um exemplo da aplicação da técnica Web DUE sendo usada para avaliar um dos mockups da aplicação móvel “PeopleBus” que está atualmente sendo projetada na Universidade Federal do Amazonas. A aplicação “PeopleBus” tem como objetivo melhorar o acesso a informações sobre o transporte público na cidade de Manaus, informando aos usuários o tempo de espera até a chegada da linha de transporte desejada. Inicialmente no processo de inspeção, o mockup é dividido em zonas: institucional, navegação, estado do sistema, serviços e entrada de dados (ver Etapa 1). Para cada zona, os itens de verificação da técnica são avaliados (ver Etapa 2). Caso aconteça uma não conformidade com algum dos itens, o inspetor deverá julgar se esta afeta a usabilidade do sistema. Por exemplo, na zona de entrada de dados, o item de verificação não está sendo respeitado, pois a aplicação não indica como preencher os dados. Assim, quando for identificado um problema de usabilidade, este deve ser apontado no mockup (ver Etapa 3).

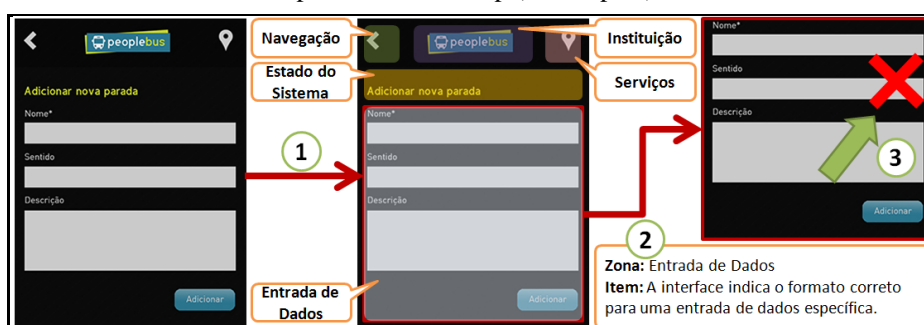


Fig. 1. Processo de inspeção da técnica Web DUE.

Os itens de verificação devem ser avaliados para cada zona presente nos mockups. A listagem completa das zonas e seus respectivos itens podem ser encontrados no trabalho de Rivero et al. [13].

3.3 A Ferramenta de Apoio Mockup DUE

A ferramenta *Mockup Design Usability Evaluation* (Mockup DUE) foi proposta para apoiar a utilização da técnica Web DUE e da Avaliação Heurística reduzindo o esforço cognitivo em cima dos inspetores ao executarem uma inspeção de usabilidade [13]. Para diminuir a dificuldade do inspetor de ter de simular a interação com o sistema enquanto realiza a inspeção, a ferramenta Mockup DUE abstrai o processo de mapeamento dos mockups para que o inspetor possa se concentrar exclusivamente na identificação de problemas. A ferramenta também fornece outras funcionalidades especí-

ficas como [13]: apontar defeitos de usabilidade e notas diretamente nos mockups; e gerar automaticamente um relatório contendo os defeitos identificados.

A ferramenta Mockup DUE apoia duas etapas do processo de inspeção: a etapa de planejamento da inspeção e a detecção de problemas de usabilidade [13]. Durante a etapa de planejamento, os moderadores responsáveis por preparar os mockups devem carregá-los e conectá-los através de links. Um link na ferramenta Mockup DUE serve para ligar um mockup ao outro, e assim simular a interação no sistema. Já na etapa de detecção de problemas, os inspetores utilizam os mockups que foram mapeados pelos moderadores e realizam uma inspeção usando a técnica Web DUE ou a técnica Avaliação Heurística. Nesta etapa, os itens de verificação da técnica Web DUE ou as heurísticas de Nielsen [12] são apresentados. Os inspetores podem tirar dúvidas sobre os itens/heurísticas enquanto apontam problemas de usabilidade nos mockups. A ferramenta Mockup DUE e seu guia de utilização estão disponíveis para download no site *DUE Technologies*¹.

4 Estudo Experimental no Contexto de Aplicações Web Móveis

Neste estudo experimental espera-se obter dados sobre a aplicabilidade de tecnologias de inspeção de usabilidade nas primeiras etapas do processo de desenvolvimento de aplicações Web móveis. Para isso, a técnica Web DUE, a Avaliação Heurística (AH) e a ferramenta Mockup DUE foram utilizadas na avaliação de uma aplicação Web móvel. Nesse contexto, o estudo experimental visava responder a seguinte questão de pesquisa: “*Qual técnica tem melhor desempenho para avaliar mockups de aplicações Web móveis e quais fatores dificultam a sua utilização nesse contexto?*” Este estudo foi planejado seguindo o processo experimental descrito por Wohlin et al. [21] e será descrito nas próximas subseções fornecendo informações para avaliar sua qualidade e a aplicabilidade dos seus resultados.

4.1 Objetivo

Este estudo possui dois objetivos principais: (a) avaliar as técnicas Avaliação Heurística e Web DUE para saber qual das duas tem melhor desempenho quando utilizadas para inspecionar mockups de aplicações Web móveis; e (b) identificar oportunidades de melhorias nas tecnologias (técnicas e ferramenta de inspeção) avaliadas quando aplicadas nesse contexto. Para atingir estes objetivos, o desempenho das técnicas foi avaliado quantitativamente através dos indicadores de eficácia e eficiência. Esses indicadores foram escolhidos, pois fornecem informações sobre a capacidade de uma técnica de inspeção em auxiliar os inspetores durante o processo de identificação de problemas [18]. Neste contexto, eficácia é a razão entre o número de defeitos que foram detectados e o total de defeitos existentes, enquanto eficiência é a razão entre o número de defeitos apontados e o tempo gasto em identificá-los.

¹ <https://sites.google.com/site/duetechnologies/>

Em relação à análise qualitativa, foi utilizado o Modelo de Aceitação da Tecnologia (*Technology Acceptance Model* - TAM) [19], para verificar até que ponto os usuários estariam inclinados a adotar as tecnologias em seu ambiente de trabalho. Adicionalmente, para entender que fatores podem facilitar a utilização destas tecnologias, foram aplicadas questões abertas para complementar as respostas ao modelo TAM.

4.2 Hipóteses

Em relação aos indicadores de eficácia e eficiência, foram testadas as seguintes hipóteses, nula e alternativa respectivamente:

H₀₁: Não existe diferença em termos de eficácia ao comparar a técnica Web DUE com a AH quando aplicadas em mockups no contexto de aplicações Web móveis.

H_{A1}: Existe diferença em termos de eficácia ao comparar a técnica Web DUE com a AH quando aplicadas em mockups no contexto de aplicações Web móveis.

H₀₂: Não existe diferença em termos de eficiência ao comparar a técnica Web DUE com a AH quando aplicadas em mockups no contexto de aplicações Web móveis.

H_{A2}: Existe diferença em termos de eficiência ao comparar a técnica Web DUE com a AH quando aplicadas em mockups no contexto de aplicações Web móveis.

4.3 Participantes

O estudo foi realizado com 20 estudantes (alunos de graduação e pós-graduação) que cursavam uma disciplina sobre a aplicação de modelos de melhoria da usabilidade em atividades de engenharia de software. Os estudantes assinaram um termo de consentimento indicando sua participação voluntária no estudo, e preencheram um formulário de caracterização, para obter informações sobre sua experiência em termos de: (a) grau de escolaridade, (b) experiência no desenvolvimento de software, (c) experiência no desenvolvimento de aplicações Web Móveis, e (d) experiência em usabilidade. Cada participante foi classificado com base nas suas respostas como tendo: Baixa, Média ou Alta experiência. Por exemplo, em relação ao grau de escolaridade, o participante foi classificado como: (a) Baixo, se o voluntário era aluno da graduação, (b) Médio, se tivesse graduação completa ou estivesse cursando o mestrado e (c) Alto, se fosse aluno de doutorado. Da mesma forma, a experiência em desenvolvimento ou usabilidade foi medida, mas levando em consideração o número de anos de experiência: (a) Baixo, se for abaixo de um ano, (b) Médio, entre um e três anos e (c) Alto, acima de três anos. A Tabela 2 apresenta o resumo da divisão dos participantes.

Com base nesta caracterização, os voluntários foram divididos em dois grupos de dez participantes cada. Para evitar o viés de ter participantes mais ou menos experientes em um dos grupos, os participantes foram alocados nas equipes de forma balanceada de acordo com o nível de experiência.

4.4 Materiais e Procedimento

Todos os participantes tiveram treinamentos sobre a utilização das tecnologias de inspeção de usabilidade: a ferramenta Mockup DUE, e uma das técnicas de inspeção,

ou a Avaliação Heurística ou a técnica Web DUE segundo a equipe à que pertenciam. Para evitar o viés do treinamento de influenciar o desempenho dos inspetores, estes apresentaram os mesmos exemplos e atividades.

Após o treinamento, cada participante recebeu: (a) a ferramenta Mockup DUE com uma das técnicas (AH ou Web DUE) embutida, (b) o arquivo com o mapeamento dos mockups, (c) o roteiro de inspeção, e (d) o questionário pós-inspeção. Nesse contexto, foi avaliada a aplicação Web móvel “*PeopleBus*” (ver Figura 1) devido a que esta ainda se encontra nas etapas iniciais do seu desenvolvimento e, além disso, a equipe de desenvolvimento desejava obter *feedback* sobre a interface que tinha sido proposta.

Os participantes instalaram a ferramenta Mockup DUE e a utilizaram para carregar o mapeamento dos mockups. Este mapeamento, como descrito na Subseção 3.3, continha os mockups da aplicação “*PeopleBus*”. Os participantes utilizaram este arquivo para visualizar os mockups e simular a interação seguindo o roteiro de inspeção. O roteiro continha a descrição de possíveis tarefas que podiam ser realizadas na aplicação. Nesse sentido, os participantes realizavam as tarefas nos mockups enquanto os avaliavam usando a técnica que tinha sido designada a eles, porém sem supervisão.

Ao concluir a inspeção, cada participante gerou um relatório de forma automática usando a ferramenta. Este relatório continha o tempo da avaliação, os defeitos apontados e sua localização nos mockups. Finalmente, o participante respondeu perguntas sobre a aplicabilidade das tecnologias de inspeção no contexto de aplicações Web móveis. Vale ressaltar que todas estas atividades foram realizadas no prazo de uma semana e que todos os materiais e artefatos necessários para replicar o estudo, encontram-se na seção de avaliação experimental no site *DUE technologies*².

4.5 Coleta de Dados

Todos os participantes entregaram o relatório de inspeção e o questionário pós-inspeção dentro do prazo disponibilizado e nenhum foi descartado. Um dos autores que era responsável pelo estudo experimental atuou como moderador da inspeção realizando: (a) a análise de todos os relatórios, (b) a verificação das informações incorretas ou ausentes, e (c) a coleção das discrepâncias contidas neles. Durante a análise dos relatórios o moderador classificava os defeitos apontados pelos inspetores como duplicatas ou não. Neste contexto, duplicatas são defeitos apontados mais de uma vez em um mesmo mockup por um inspetor ou inspetores diferentes. Após esta atividade, o moderador gerou uma nova tabela que continha todas as discrepâncias, porém sem duplicatas. Vale ressaltar que caso houvesse dúvida ao julgar se uma discrepância era duplicata ou não, esta era resolvida ao consultar com os demais pesquisadores envolvidos no estudo experimental.

Após a atividade de coleta, houve uma reunião com quatro pesquisadores experientes em usabilidade e não relacionados com esta pesquisa. Nesta reunião, foi executada a discriminação das discrepâncias, que consiste em classificá-las como defeitos reais ou falso-positivos. Realizada a discriminação, obteve-se a tabela final de defeitos, que continha apenas os problemas que afetavam a usabilidade da aplicação *PeopleBus*,

² <https://sites.google.com/site/duetechnologies/estudos-experimentais>

sem falso-positivos. Considerando todas as discrepâncias, houve um total de 101 defeitos reais distribuídos ao longo de 11 mockups.

5 Discussão dos Resultados

Foi realizada uma análise quantitativa para testar as hipóteses sobre o grau de eficiência e eficácia das técnicas de inspeção aplicadas; e uma análise qualitativa para conhecer a opinião dos usuários em relação à aplicabilidade das tecnologias. Os dados quantitativos referem-se ao tempo necessário para realizar a inspeção e o número de discrepâncias por inspetor, enquanto os dados qualitativos foram coletados através das respostas ao questionário pós-inspeção.

5.1 Análise Quantitativa

A Tabela 2 apresenta os resultados por inspetor e técnica. Para calcular o indicador de eficácia foi utilizado o valor 101, número total de defeitos identificados no estudo. Já o indicador de eficiência foi calculado dividindo o número de defeitos por inspetor pelo tempo em horas que este levou para identificar os defeitos.

Tabela 2. Resultados obtidos por inspetor e técnica.

Equipe Avaliação Heurística											
ID Inspetor		1	2	3	4	5	6	7	8	9	10
Experiência	Grau de Escolaridade	B	B	B	B	B	M	A	M	B	B
	Desenvolvimento de Software	B	M	M	B	B	M	M	A	B	B
	Desenvolvimento Web Móvel	B	B	B	B	B	B	B	B	B	B
	Usabilidade	B	B	B	B	B	B	A	B	B	B
Total de Discrepâncias		16	6	17	31	10	9	41	37	26	21
Total de Falsos Positivos		7	2	5	3	1	3	1	13	11	7
Total de Defeitos		9	4	12	28	9	5	40	24	15	14
Tempo (min.)		66	22	28	48	17	62	98	26	25	58
Eficácia (%)		8,9	4,0	11,9	27,7	8,9	5,0	39,6	23,8	14,9	13,9
		Média: 15,84 e Mediana: 12,87									
Eficiência (defeitos por hora)		8,2	10,9	25,7	35,0	31,8	4,8	24,5	55,4	36,0	14,5
		Média: 24,68 e Mediana: 25,10									
Equipe Web DUE											
ID Inspetor		11	12	13	14	15	16	17	18	19	20
Experiência	Grau de Escolaridade	B	B	B	B	M	M	B	B	B	B
	Desenvolvimento de Software	B	M	M	B	M	A	M	B	B	B
	Desenvolvimento Web Móvel	B	B	B	B	B	A	B	B	B	B
	Usabilidade	B	B	B	B	B	B	B	B	B	B
Total de Discrepâncias		16	22	40	13	72	40	19	36	56	25
Total de Falsos Positivos		4	3	10	1	18	13	6	17	21	2
Total de Defeitos		12	19	30	12	54	27	13	19	35	23
Tempo (min.)		104	80	119	13	66	52	53	55	88	48
Eficácia (%)		11,9	18,8	29,7	11,9	53,5	26,7	12,9	18,8	34,7	22,8
		Média: 24,16 e Mediana: 20,79									
Eficiência (defeitos por hora)		6,9	14,3	15,1	55,4	49,1	31,2	14,7	20,7	23,9	28,8
		Média: 26,00 e Mediana: 22,30									

Para comparar os indicadores de eficácia e eficiência em ambas as amostras, foi aplicado o teste estatístico Mann-Whitney. Para um valor de $\alpha = 0.05$, o teste apontou valores $p = 0,123$ a favor de H_{01} e $p = 1,000$ a favor de H_{02} . Estes resultados indicam que não existe diferença do ponto de vista estatístico. No entanto, devido ao pequeno tamanho da amostra, as hipóteses nulas não puderam ser rejeitadas e consequentemente, os resultados obtidos devem ser considerados indicativos e não conclusivos. Os gráficos de boxplot dos indicadores apresentados na Figura 2 apontam que a técnica Web DUE teve uma mediana maior em termos de eficácia, mas uma mediana menor em termos de eficiência. Isto pode ser explicado devido ao número de itens de verificação da técnica Web DUE em comparação com a técnica AH. Como a técnica Web DUE é mais específica, esta permite identificar mais problemas. Entretanto, o fato de ter que verificar um número maior de itens, faz com que os inspetores levem mais tempo em identificar problemas de usabilidade.

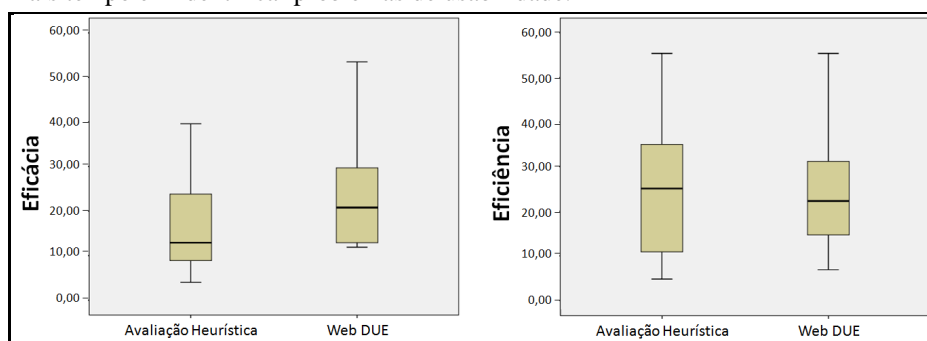


Fig. 2. Gráfico de Boxplots comparando os indicadores de eficácia e eficiência.

5.2 Análise Qualitativa

O questionário pós-inspeção foi formulado para coletar dados sobre a viabilidade de utilizar as tecnologias (técnicas AH e Web DUE, e ferramenta Mockup DUE) no contexto de mockups de aplicações Web móveis; e identificar oportunidades de melhoria para aumentar o desempenho da inspeção em termos de eficácia e eficiência. Para avaliar a aceitação das tecnologias no contexto descrito, foram feitas perguntas com base no Modelo de Aceitação da Tecnologia (*Technology Acceptance Model - TAM*) [19]. O TAM serve como instrumento para prever a possibilidade de que uma nova tecnologia seja adotada por um grupo ou organização, ao fazer perguntas sobre a utilidade percebida, a facilidade de uso percebida e a intenção comportamental de uso. As respostas às perguntas do TAM foram fornecidas em uma escala de quatro níveis (discordo fortemente – DF e parcialmente – DP, e concordo parcialmente - CP e fortemente - CF) sem considerar um nível neutro já que este não fornece informações sobre o grau de concordância do participante com as afirmativas [10].

A fim de entender que fatores influenciaram negativamente a opinião dos participantes com relação à adoção das tecnologias, foram incluídas perguntas abertas no questionário pós-inspeção. Por exemplo, após as perguntas do modelo TAM sobre facilidade de uso, foram incluídas perguntas abertas em relação a fatores que tenham

facilitado/dificultado a utilização das tecnologias. A Tabela 3 apresenta o resumo das respostas obtidas para as afirmativas do modelo TAM relacionadas com Utilidade, Facilidade de Uso e Intenção de Uso. Apesar de ter coletado dados para a melhoria de ambas as técnicas, esta seção foca em oportunidades de melhoria para a técnica Web DUE e a ferramenta Mockup DUE, visto que estas tecnologias foram projetadas exclusivamente para avaliar a usabilidade em mockups [13].

As respostas das questões abertas para entender a opinião dos participantes em relação aos fatores do modelo TAM foram analisadas seguindo procedimentos de Grounded Theory [16], um método qualitativo que permite a coleta e análise de dados de forma sistemática. Nesse sentido, foram analisadas as respostas dos participantes que discordaram forte ou parcialmente com alguma das afirmativas do modelo TAM com o intuito de identificar os fatores que influenciaram o participante a discordar.

Tabela 3. Número de respostas para as afirmativas do TAM da equipe Web DUE.

Fatores	Afirmativas	Sobre a Técnica Web DUE					Sobre a Ferramenta Mockup DUE				
		DF	DP	CP	CF	Total	DF	DP	CP	CF	Total
Utilidade	A1	0	2	5	3	10	1	1	7	1	10
	A2	1	1	6	2	10	0	0	5	5	10
	A3	0	1	7	2	10	0	0	7	3	10
	A4	0	3	3	4	10	1	1	6	2	10
Facilidade de Uso	A1	0	1	5	4	10	0	0	5	5	10
	A2	0	3	3	4	10	0	1	5	4	10
	A3	0	1	7	2	10	0	0	5	5	10
	A4	0	1	6	3	10	0	0	6	4	10
Intenção de Uso	A1	0	4	5	1	10	0	2	5	3	10

Afirmativas sobre Utilidade: (A1) Melhora a eficácia, (A2) Melhora a eficiência, (A3) Melhora a Precisão, (A4) É útil. **Afirmativas sobre Facilidade de Uso:** (A1) É fácil aprender a usar, (A2) É fácil fazer o que preciso, (A3) É fácil se tornar habilidoso, (A4) É fácil de usar. **Afirmativa sobre Intenção de Uso:** (A1) Pretendo usar.

Em relação à técnica Web DUE, quatro inspetores discordaram forte ou parcialmente de que a técnica é útil; três inspetores discordaram de que fosse fácil de usar e quatro discordaram de que a utilizariam no seu ambiente de trabalho. Com relação à utilidade e intenção de uso, três fatores que afetaram a opinião dos inspetores foram: o tamanho da técnica, a experiência dos inspetores e a falta de aplicabilidade de alguns itens de verificação. Nesse sentido, o Inspetor 11 afirmou que a técnica é muito longa e cansativa e que tinha itens repetidos (ver afirmativa do inspetor 11). Já o Inspetor 17 indicou que a falta de conhecimento afetou sua opinião em termos do uso da técnica. Este inspetor também indicou que alguns itens não estavam complementa relacionados ao contexto de aplicações Web móveis. Vale ressaltar que ambos os inspetores possuem baixa experiência em usabilidade e inspeções, e que também identificaram um baixo número de defeitos em comparação com o restante da equipe.

“Achei um pouco cansativa por ser longa... a técnica possibilita focar em uma área do sistema, mas era um pouco repetitiva” – Inspetor 11.

“A dificuldade foi nos conceitos e termos utilizados, posteriormente tive que ler todos os itens de verificação de todas as zonas, a fim de melhorar o entendimento, isto demandou muito tempo... Acredito que alguns itens de verificação não são totalmente aplicáveis a aplicações moveis.” – Inspetor 17.

Em relação à facilidade de uso da técnica Web DUE os inspetores que discordaram das afirmações, apontaram a dificuldade em entender alguns itens de verificação, pois estes se repetiam (ver afirmativa do inspetor 11). Além disso, devido ao fato que alguns conceitos não estavam claros, isso também dificultou a aplicação da técnica Web DUE (como indica a afirmativa do Inspetor 17 acima).

“Alguns itens de uma zona eram parecidos com itens de outras zonas. Às vezes fiquei em dúvida.” - Inspetor 11.

A ferramenta Mockup DUE também foi avaliada usando o modelo TAM e questões abertas. Deste modo, três inspetores discordaram sobre sua utilidade, enquanto um e dois inspetores discordaram sobre sua facilidade de uso e intenção de uso respectivamente. Com relação à utilidade e intenção de uso, os inspetores ressaltaram que o principal problema é a falta de adequação de alguns itens de verificação (ver afirmativa do Inspetor 17), ou seja, os inspetores não conseguem diferenciar a técnica da ferramenta, e visualizam ambas como uma tecnologia só. Além disso, apesar de gostarem da interface da ferramenta e a forma em que esta apoia o processo de inspeção, os inspetores afirmaram que é necessário facilitar o acesso às funcionalidades mais utilizadas para facilitar seu uso (ver afirmativa do Inspetor 20).

“Adequação de alguns itens de verificação a web móveis” - Inspetor 17.

“Eu gosto do design da ferramenta, a única coisa que eu mudaria é colocar os ícones mais usados (adicionar erro e nota) mais próximos da tela, onde fica atualmente a navegação entre mockups, porque os botões de adicionar são bem mais utilizados...” – Inspetor 20.

6 Ameaças à Validade

Com relação à validade interna, que analisa o tratamento dos resultados [21], foram consideradas quatro ameaças. Em relação ao efeito de treinamento, pode haver um risco se a qualidade do treinamento da técnica Web DUE for diferente ao da Avaliação Heurística. No entanto, esse risco foi mitigado apresentando os mesmos exemplos de problemas de usabilidade. Além disso, a fim de controlar o efeito da experiência dos participantes no desempenho das técnicas, estes foram divididos em grupos balanceados. Para evitar o viés de uma divisão inadequada, esta divisão também foi feita com base na experiência acadêmica e em desenvolvimento dos participantes, que foi mensurada através de um questionário objetivo. Finalmente, outra ameaça é a medição do tempo, que apesar de ter sido obtida automaticamente através da ferramenta, não é possível garantir que os participantes realizaram uma medição cuidadosa.

Com relação à validade externa, que se preocupa com a generalização dos resultados [21], serão discutidas três ameaças. Com relação à primeira, a participação de alunos como inspetores, é possível argumentar que estes foram distribuídos em equipes balanceadas. Além disso, nenhum dos alunos tinha experiência previa na utilização das técnicas no contexto de mockups de aplicações Web móveis para evitar o risco de alterar os resultados. A segunda ameaça é a utilização de um ambiente acadêmico para a realização do estudo experimental. No entanto, foi avaliada uma aplicação Web móvel real, nas fases iniciais do seu desenvolvimento, ajudando a simular

um ambiente industrial. Finalmente, a avaliação da aplicação “*PeopleBus*” pode ser considerada uma ameaça, pois existem diferentes tipos de aplicações Web móveis.

A principal ameaça do estudo é o tamanho da amostra que afeta a validade de conclusão sobre a relação entre o tratamento e os resultados [21]. Nesse sentido, visto que o número de participantes é baixo, os resultados obtidos no estudo devem ser considerados indicativos e não conclusivos. Adicionalmente, pode haver uma ameaça à validade de constructo (sobre a relação entre a teoria e a observação [21]) com relação aos critérios utilizados medir o desempenho das técnicas. No entanto, eficácia e eficiência são critérios comuns usados na avaliação de técnicas de inspeção [18].

7 Conclusões e Perspectivas Futuras

Neste artigo, compararam-se as técnicas Web DUE e Avaliação Heurística quando utilizadas dentro da ferramenta Mockup DUE para avaliar a usabilidade de mockups de aplicações Web móveis. Devido ao tamanho da amostra, os resultados quantitativos devem ser considerados indícios de que não existe diferença em utilizar as técnicas no contexto de aplicações Web móveis. No entanto, como os inspetores indicaram, a técnica Web DUE possui mais itens de verificação, o que permite identificar mais problemas de usabilidade, porém aumentando também o tempo de inspeção. Os resultados de eficácia da técnica Web DUE são similares aos de outras técnicas de inspeção [6]. Porém, deve-se tomar cuidado ao generalizar estes resultados, pois as aplicações avaliadas foram diferentes, assim como o número e experiência dos inspetores que participaram do estudo experimental. A Tabela 4 apresenta um resumo das modificações a serem feitas em cima da técnica Web DUE e da ferramenta Mockup DUE, que foram obtidas através da análise qualitativa neste estudo experimental.

Tabela 4. Melhorias a serem incorporadas em futuras versoes das tecnologias DUE.

Tecnologia	Problemas Identificados	Possíveis Pontos de Melhoria
Técnica Web DUE	- Esforço em realizar uma inspeção. - Baixo desempenho na inspeção.	- Diminuir o número de itens de verificação. - Criar uma lista única contendo os itens que podem ser avaliados independentes das zonas de páginas Web. - Incluir ou refinar os itens de verificação atuais ao contexto Web móvel.
Ferramenta Mockup DUE	- Facilitar o acesso às funcionalidades. - Melhorar Design.	- Reposicionamento de botões específicos. - Aumento do número de dicas disponibilizado no momento da inspeção. - Melhorar acesso e consulta de itens de verificação.

Como trabalho futuro, pretende-se repetir este estudo com um maior número de participantes; e continuar a análise dos itens de verificação para adequá-los ao contexto Web móvel, conforme sugerido na análise qualitativa. Espera-se com essas medidas, melhorar as tecnologias avaliadas e encorajar a indústria de desenvolvimento de software a utilizá-las nas primeiras etapas do processo de desenvolvimento.

Referências

1. Bastien, J.: Usability testing: a review of some methodological and technical aspects of the method. In: International Journal of Medical Informatics, pp. 18 - 23 (2010)

2. Bertini, E, Gabrielli, S., Kimani, S.: Appropriating and assessing heuristics for mobile computing. Proc. of the working conference on Advanced visual interfaces (2006)
3. Bonifácio, B., Cruz, M.J., Conte T.: Extração e análise de características para avaliação de usabilidade em aplicações móveis baseadas em evidências da literatura. Relatório Técnico UsES-RT-003-2010 (2010)
4. Bonifácio, B., Fernandes, P., Santos, F., Oliveira, H. A., Conte, T.: Usabilidade de Aplicações Web Móvel: Avaliando uma Nova Abordagem de Inspeção através de Estudos Experimentais. Anais do XV CibSE, pp. 236-249 (2012)
5. Charette, R.: Why software fails. IEEE Spectrum, Volume 42, Issue 9 (2005)
6. Fernandes, P., Rivero, L., Bonifácio, B., Santos, D., Conte, T.: Avaliando uma nova Abordagem para Inspeção de Usabilidade através de Análise Quantitativa e Qualitativa. Proc. of the VIII ESELAW, pp. 67-76 (2011)
7. Fernandez, A., Insfran, E., Abrahao, S.: Usability evaluation methods for the Web: A systematic mapping study. Information and Software Technology, Volume 53, Issue 8 (2011)
8. Fons, J., Pelechano, V., Pastor, O., Valderas, P., Torres, V.: Applying the OOWS model-driven approach for developing Web applications: The internet movie database case study. In: Rossi, G., Schwabe, D., Olsina, L., Pastor, O.: Web Engineering: Modeling and Implementing Web Applications, Springer (2008)
9. International Organization for Standardization - ISO/IEC 9126-1: Software Engineering – Product Quality – Part 1: Quality Model (2001)
10. Laitenberger, O., Dreyer, H.: Evaluating the Usefulness and the Ease of Use of a Web-based Inspection Data Collection Tool. Proc. 10th Software Metrics Symposium, Bethesda - MD, pp. 122-132 (1998)
11. Matera, M., Rizzo, F., Carughi, G.: Web Usability: Principles and Evaluation Methods. In: Mendes, E., Mosley, N.: Web Engineering, Springer (2006)
12. Nielsen, J.: Finding usability problems through heuristic evaluation. Proc. of the CHI'92, UK, pp. 373-380 (1992)
13. Rivero, L., Barreto, R., and Conte, T.: Characterizing Usability Inspection Methods through the Analysis of a Systematic Mapping Study Extension. In: Latin-american Center for Informatics Studies Electronic Journal, Volume 16, Issue 1, (2013)
14. Rocha, H., Baranauska, M.: Design e avaliação de interfaces humano-computador. (Livro) Instituto de Computação, Universidade Estadual de Campinas (2003)
15. Souza, C., Prates R., Barbosa S.: Adopting information technology as a first step in design. In: Journal of Interactions, Volume 6, pp. 72-79 (2003)
16. Strauss A., Corbin, J.: Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. 2 ed. London, SAGE Publications (1998)
17. Tarasewich, P.: Designing mobile commerce applications. In: Communications of the ACM, Volume 46, Issue 12, pp. 57-60 (2003)
18. Triacca, L., Inversini, A., Bolchini, D.: Evaluating Web usability with MiLE+. Proc. 7th IEEE International Symposium on Web Site Evolution, Hungary, pp. 22-29 (2005)
19. Venkatesh, V., Davis, F.: A theoretical extension of the technology acceptance model: four longitudinal field studies. In: Management Science, Volume 46, pp. 186-204 (2000)
20. Weiss, S.: Handheld usability: design, prototyping and usability testing for mobile phones. Proc. of the 7th International Conference on Human Computer Interaction with Mobile Devices and Services, Volume 12, pp. 367-369 (2005)
21. Wohlin, C. Runeson, P., Host, M., Ohlsson, M., Regnell, B., Wessl, A.: Experimentation in software engineering: an introduction (Book), Kluwer Academic Publishers (2000)