

# Estudo Comparativo entre Técnicas de Inspeção de Usabilidade sobre Diagramas de Atividades

**Natasha M. C. Valentim**  
Universidade Federal do Amazonas  
Av. Gen. Rodrigo Octávio  
3000, Manaus – AM –  
Brasil  
natasha\_costa16@yahoo.  
com.br

**Tiago Silva da Silva**  
Universidade de São Paulo  
Av. Arlindo Bétio,  
1000, São Paulo – SP  
– Brasil  
silvadasilva@usp.br

**Milene S. Silveira**  
Pontifícia Universidade  
Católica do Rio Grande  
do Sul  
Av. Ipiranga, 6681,  
Porto Alegre – RS –  
Brasil  
milene.silveira@pucrs.br

**Tayana Conte**  
Universidade Federal do Amazonas  
Av. Gen. Rodrigo  
Octávio 3000, Manaus –  
AM – Brasil  
tayana@icomp.ufam.  
edu.br

## ABSTRACT

The goal of this paper is to compare the performance of two usability inspection methods that aim at anticipating the evaluation process by inspecting Activity Diagrams. In order to achieve this goal, we carried out a feasibility study by following an established inspection process. According to quantitative results, the Usability Guidelines method outperformed the MIT 3 with regards to efficacy and had similar performance regarding efficiency. However, through a qualitative analysis, the MIT 3 was considered easier to remember and to learn and also more useful than the Usability Guidelines. We found out that the Usability Guidelines method could be used for more experienced professionals on usability issues and perhaps has an even better performance. Notwithstanding, the MIT 3 method could be suggested for the use by a little less experienced professionals with regards to usability issues, supporting the formation and training of design and development teams.

## RESUMO

O objetivo deste trabalho é comparar dois métodos de inspeção de usabilidade os quais visam antecipar o processo de avaliação através da inspeção de Diagramas de Atividades. A fim de atingir este objetivo, conduziu-se um estudo de viabilidade seguindo um estabelecido processo de inspeção. De acordo com resultados quantitativos, o método de Diretrizes de Usabilidade superou o método MIT 3 no que diz respeito à eficácia, e obteve desempenho similar em relação à eficiência. No entanto, de acordo com a análise qualitativa, o método MIT 3 foi considerado mais fácil de lembrar, de aprender, e mais útil que as Diretrizes de Usabilidade. Notou-se que as Diretrizes de Usabilidade poderiam ser utilizadas por profissionais mais experientes

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. IHC'13, Brazilian Symposium on Human Factors in Computing Systems. October 8-11, 2013, Manaus, AM, Brazil. Copyright 2013 SBC. ISSN 2316-5138 (pendrive). ISBN 978-85-7669-278-2 (online).

em usabilidade, podendo talvez apresentar um melhor desempenho. Não obstante, o método MIT 3 pode ser sugerido para uso de profissionais um pouco menos experientes no que diz respeito à usabilidade, apoiando a formação e capacitação de equipes de projeto e desenvolvimento.

## Keywords

Usability evaluation; inspection; activity diagrams.

## ACM Classification Keywords

H.5 INFORMATION INTERFACES AND PRESENTATION / H.5.2 User Interfaces/Evaluation/methodology.

## General Terms

Human factors; design; experimentation; verification.

## INTRODUÇÃO

O processo de avaliação de Interação Humano-Computador (IHC) é um processo de coleta de dados que informa como um usuário ou grupo de usuários usará o produto para atingir determinado objetivo em um certo contexto [16]. Seus principais objetivos são:

- Prever a usabilidade de um produto ou algum aspecto dele;
- Verificar se a equipe de *design* entende os requisitos dos usuários;
- Testar ideias de forma rápida e informal;
- Identificar problemas que os usuários estão enfrentando e melhorar o produto para satisfazer as necessidades dos usuários.

Com o intuito de atingir estes objetivos, existem diversos métodos e técnicas de avaliação que são classificados, em linhas gerais, como métodos de avaliação formativa e somativa [9].

A avaliação formativa é realizada durante o processo de desenvolvimento, usando como base para avaliação artefatos como cenários, *storyboards* e modelos de interação. Por outro lado, a avaliação somativa é realizada nas etapas finais do processo de desenvolvimento, a fim de

avaliar o sucesso de um produto ou verificar sua conformidade com determinados padrões.

Dentre estes métodos, os mais amplamente utilizados são os de avaliação somativa. Na maioria dos casos, estes métodos são fáceis de entender e de aplicar, pois são utilizados quando se tem, pelo menos, um protótipo do sistema a ser avaliado, tornando assim o processo de avaliação mais tangível [2].

No entanto, apesar da maior utilização de métodos de avaliação somativa, o fato desta ser realizada apenas nas etapas finais do processo de desenvolvimento pode acarretar altos custos para reparar falhas. Assim, quanto mais cedo os problemas de interação forem identificados, menor será o esforço com retrabalho e, por conseguinte, mais baixos serão os custos com eventuais reparos.

Dentre os objetivos da avaliação de IHC apresentados, este trabalho está inserido no objetivo de “prever a usabilidade de um produto ou algum aspecto dele” com foco na identificação antecipada de problemas por meio da aplicação de métodos de Inspeção de Usabilidade. Nestes métodos, inspetores examinam aspectos da aplicação para detectar violações de princípios de usabilidade estabelecidos [15]; eles foram propostos como uma alternativa com bom custo-benefício em comparação aos testes com usuários.

O presente trabalho apresenta um estudo de viabilidade que objetiva analisar o desempenho de duas técnicas de inspeção de usabilidade. As técnicas avaliadas neste estudo visam levar o processo de avaliação de usabilidade para a etapa de projeto, quando se tem apenas a modelagem – neste caso, diagramas de Atividades da UML [3] – do sistema a ser desenvolvido e não um protótipo funcional. A opção por este tipo de diagramas – da Engenharia de Software (ES) – deve-se ao fato de ser um artefato bastante utilizado pela indústria de *software* nas etapas iniciais do processo de desenvolvimento [3].

O restante deste artigo está organizado da seguinte forma: a Seção 2 define técnicas de inspeção e apresenta as técnicas que visam avaliar a usabilidade através de diagramas de atividades. A Seção 3 apresenta o estudo de viabilidade realizado. A Seção 4 apresenta os resultados quantitativos e qualitativos do estudo; e a Seção 5 apresenta as ameaças à validade deste estudo. Por fim, a Seção 6 apresenta as conclusões e trabalhos futuros.

## **INSPEÇÃO DE USABILIDADE APLICADA A DIAGRAMAS DE ATIVIDADES**

Durante a fase de projeto do processo de desenvolvimento de *software*, são construídos, dentre outros, os Diagramas de Atividades, que são diagramas utilizados pela indústria para representar a interação dentro das funcionalidades de forma simples e clara [3]; assim sendo, torna-se importante buscar avaliar a usabilidade por meio destes.

Do ponto de vista da ES, dentre as diferentes formas de avaliação, “a inspeção de software é uma das melhores práticas para a detecção e remoção de defeitos encontrados durante o processo de desenvolvimento” [11]. Em IHC, a avaliação por inspeção é utilizada para buscar problemas de usabilidade em um projeto de interface existente, e analisar estes problemas com vistas a fazer recomendações para consertá-los e assim melhorar a usabilidade do projeto [15]. Os principais objetivos deste tipo de avaliação são identificação de problemas de usabilidade e seleção dos problemas que devem ser corrigidos, promovendo também a formação e capacitação da equipe com relação a projetos de interfaces centrados no usuário [14].

Pesquisas apontam tentativas de apoiar a usabilidade por meio do diagrama de atividades. Em um mapeamento sistemático realizado em 2011 [21], foi citado o estudo inicial da abordagem proposta por Atterer [1]. Neste estudo, criaram-se diagramas de atividades que tiveram como objetivo mostrar o fluxo de atividades em um único processo. A abordagem de Atterer [1] buscou usar modelos abstratos de abordagens de engenharia *web* durante a validação automática de usabilidade e por isso o diagrama de atividades serviu apenas como base para a criação manual de *links* adequados e conteúdo de página.

Recentemente, duas técnicas de inspeção de usabilidade em diagrama de atividades foram propostas: a MIT 3 [22] e as Diretrizes de Usabilidade [18], as quais serão a seguir apresentadas.

Ambas as técnicas têm como propósito uma fácil adoção pela indústria por não necessitar de apoio ferramental nem de especialistas com muita experiência em IHC, contribuindo para a formação e capacitação da equipe no que diz respeito a usabilidade de sistemas interativos. Além disso, estas devem ser utilizadas de forma complementar a outras técnicas de avaliação de usabilidade.

### **MIT 3**

A MIT 3 faz parte de um conjunto de técnicas, chamada MIT (*Model Inspection Technique for Usability Evaluation*), que são técnicas de leitura propostas para inspeção de usabilidade em modelos de projeto. Segundo Travassos *et al.* [20], técnicas de leitura são um tipo específico de técnica de inspeção que contém uma série de passos para a análise individual de um produto de *software*, de forma a alcançar a compreensão necessária para uma tarefa específica. Segundo os autores, estas têm como objetivo aumentar a eficácia das inspeções, fornecendo diretrizes que podem ser utilizados pelos inspetores para analisar um determinado artefato de *software* e identificar defeitos de *software*.

A MIT 3 possui 14 itens de verificação que guiam o inspetor e que estão agrupados nas seguintes heurísticas [23]: concordância entre o sistema e o mundo real; controle e liberdade ao usuário; consistência e padrões; prevenção de erros; reconhecer em vez de lembrar; flexibilidade e

eficiência de uso; reconhecimento, diagnóstico e recuperação de erros; e proporcionar funções que são úteis para o usuário. A Tabela 1 apresenta alguns dos itens de verificação da MIT 3.

MIT-3C. Controle e liberdade ao Usuário	
3C1	Verifique se o usuário através de alguma atividade consegue alterar dados que ele inseriu através de outra atividade.
MIT-3D. Consistência e padrões	
3D1	Verifique se todas as atividades indicam funcionalidades diferentes, ou seja, não há atividades com nomes diferentes, mas com a mesma funcionalidade.
MIT-3E. Prevenção de Erros	
3E1	Verifique se há alguma atividade em que o sistema previne algum tipo de erro.
MIT-3G. Flexibilidade e eficiência de uso	
3G1	Verifique se o usuário possui diferentes formas de acesso às atividades principais.
3G2	Verifique se o usuário pode navegar com facilidade pelas diferentes atividades, ou seja, as possibilidades de navegação para atingir um objetivo estão claras para o usuário.

Tabela 1. Exemplos de itens de verificação da MIT 3.

### Diretrizes de Usabilidade

As Diretrizes de Usabilidade consistem em conjuntos de diretrizes definidos com o objetivo de levar a avaliação de usabilidade para a etapa de projeto de sistemas interativos. Originalmente, foram definidos quatro conjuntos de diretrizes – para Diagramas de Casos de Uso, Diagramas de Atividades, Modelo de Tarefas e Modelo de Interação –, no entanto, no presente estudo – para fins comparativos – são utilizadas apenas as diretrizes para os Diagramas de Atividades.

As Diretrizes para o Diagrama de Atividades possuem 16 questões, agrupadas nas seguintes heurísticas [23]: Correspondência entre o sistema e o mundo real, Visibilidade do estado do sistema, Reconhecimento em vez de lembrança e Prevenção de erros.

A Tabela 2 apresenta algumas questões que compõem o método Diretrizes de Usabilidade.

Diretrizes de Usabilidade – Visibilidade do Estado do Sistema	
AD06	Há algum feedback do sistema para cada ação do usuário?
Diretrizes de Usabilidade –Prevenção de Erros	
AD15	Opções (tarefas) com nomes similares desempenham funções opostas e potencialmente perigosas?

AD16	Ao final de uma sessão de trabalho o sistema informa sobre o risco de perda de dados?
Diretrizes de Usabilidade – Flexibilidade e Eficiência de Uso	
AD18	Se o usuário pode voltar a uma atividade anterior, ele pode mudar sua escolha anterior?
AD19	O usuário pode reverter facilmente suas ações?

Tabela 2. Exemplos de questões das Diretrizes de Usabilidade.

### ESTUDO DE VIABILIDADE

Anteriormente, em outros estudos, tanto a técnica MIT 3 [22] quanto as Diretrizes de Usabilidade [18] foram avaliadas individualmente. No entanto, como as duas técnicas têm o mesmo propósito – avaliar a usabilidade através de diagramas de atividades –, realizou-se este estudo com o objetivo de verificar qual das técnicas obtém maiores resultados em termos de eficácia – razão entre o número de defeitos detectados e o total de defeitos existentes – e eficiência – razão entre o número de defeitos detectados e o tempo gasto na inspeção.

Para atingir esta finalidade, seguiu-se um processo de inspeção que foi dividido em quatro atividades, tendo por base o processo de inspeção de usabilidade sugerido por Sauer *et al.*[17]. Essas atividades são apresentadas na Figura 1. As atividades e papéis que compõem o processo de inspeção de usabilidade serão descritas nas subseções a seguir.

### Planejamento do Estudo

O planejamento é a atividade na qual é feita a definição do escopo da inspeção, preparação do material para a inspeção, seleção dos inspetores, treinamento dos inspetores nas técnicas e atribuição das tarefas a cada inspetor. Esta atividade é realizada pelo líder da inspeção (ou moderador). Neste estudo, esse papel foi exercido por uma pessoa que possui conhecimento e experiência em avaliações de usabilidade.

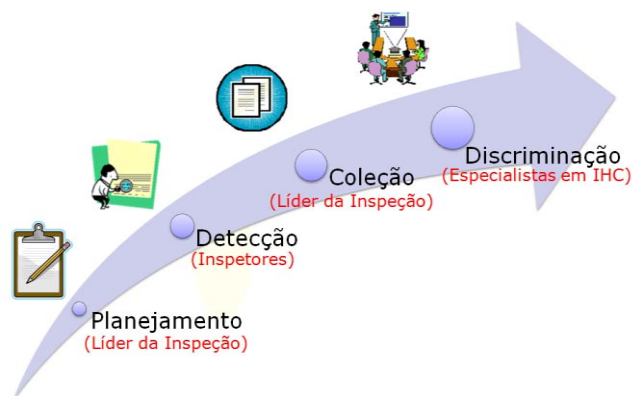


Figura 1 - Processo de inspeção baseado no sugerido por Sauer *et al.* [17].

Os participantes deste estudo foram alunos voluntários da disciplina Modelagem e Projeto de Sistemas do 6º período do curso de graduação em Ciência da Computação da Universidade Federal do Amazonas.

Ao todo, 27 alunos concordaram inicialmente em participar do estudo. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) e preencheram o Formulário de Caracterização que continha perguntas sobre o conhecimento dos participantes em relação a Avaliação de Design em IHC. Os 27 participantes foram divididos em dois grupos, de forma a distribuir igualmente os participantes com conhecimento alto, médio e baixo no que diz respeito à IHC, Avaliação e Design.

Um dos grupos utilizou as Diretrizes de Usabilidade para a inspeção de usabilidade de um Diagrama de Atividades e o outro grupo usou a MIT 3 para o mesmo fim. O Diagrama de Atividades avaliado neste experimento faz parte da especificação de um sistema real, que realiza cadastro, exclusão e edição de dados referentes à doações<sup>1</sup>.

Uma apresentação sobre usabilidade foi realizada para todos os participantes do estudo. Para cada grupo foi realizada uma breve apresentação sobre a técnica que o grupo iria utilizar para fazer a inspeção.

#### **Deteção**

Em seguida é feita a atividade de deteção de defeitos, na qual cada inspetor executa individualmente essa atividade, a qual consiste na busca de problemas de usabilidade nas interações selecionadas.

Inicialmente, o líder da inspeção foi responsável por passar as informações para execução da avaliação aos inspetores. Uma das orientações dada foi que todos teriam que anotar o horário inicial e o horário final da inspeção.

Em seguida, os grupos foram divididos e ficaram em salas diferentes. Cada participante recebeu as instruções para inspeção, a técnica de inspeção (MIT 3 ou Diretrizes de Usabilidade) e uma planilha para a anotação das discrepâncias encontradas, além do modelo (Diagrama de Atividades) a ser inspecionado.

Ao todo, participaram 13 inspetores usando a MIT 3 e 12 inspetores usando as Diretrizes de Usabilidade, pois dois alunos não compareceram no dia da realização do estudo. A atividade de deteção de defeitos foi realizada individualmente pelos inspetores e vale ressaltar que durante esta atividade os inspetores não receberam nenhum auxílio dos pesquisadores envolvidos no estudo.

#### **Coleção**

Após a atividade de deteção, é realizada a atividade de coleção, na qual é feita a redução de discrepâncias repetidas (problemas encontrados por mais de um inspetor), gerando

uma lista com possíveis discrepâncias únicas (sem duplicatas), retirando-se a referência do inspetor e da técnica utilizada por ele. Esta atividade também é realizada pelo líder da inspeção.

#### **Discriminação**

Por fim, é realizada a atividade de discriminação. Nesta atividade, a lista de discrepâncias foi passada para uma equipe de dois outros pesquisadores da área de IHC. Esta equipe realizou uma reunião para decidir quais dessas discrepâncias (problemas encontrados pelos inspetores) eram únicas e quais eram duplicatas se houverem (problemas equivalentes apontados por mais de um inspetor), além de classificar em defeitos ou falso-positivos. Nesta atividade, os especialistas também podem classificar as discrepâncias em duplicatas, pois mesmo tentando reduzir o número de duplicatas na atividade de Coleção, algumas duplicatas permanecem.

#### **RESULTADOS OBTIDOS**

A seguir serão apresentados os resultados quantitativos e qualitativos do estudo de viabilidade.

#### **Resultados Quantitativos**

Os níveis de conhecimento relacionados à IHC, Avaliação e Design e os resultados para cada inspetor que avaliou o diagrama de atividades após a classificação durante a reunião de discriminação são mostrados na Tabela 3.

Ao analisar esta tabela, pode-se verificar que as Diretrizes de Usabilidade ajudaram os inspetores a detectarem mais defeitos na inspeção. Pode-se observar também que os inspetores que utilizaram a MIT 3 tiveram o tempo de inspeção variando entre 0,38 horas e 1,35 horas e encontraram entre 8 e 22 defeitos. Já os inspetores que utilizaram as Diretrizes de Usabilidade empregaram entre 0,58 a 1,32 horas na deteção, encontrando entre 11 e 23 defeitos.

A Tabela 4 apresenta as médias calculadas para os indicadores de eficácia e eficiência.

Para comparar a eficiência das duas amostras, utilizou-se análise de *boxplots* e o teste não-paramétrico Mann-Whitney [13], dado o tamanho limitado das amostras. Para realização da análise estatística foi utilizado o *software* SPSS com  $\alpha=0.05$ . A Figura 2 mostra os *boxplots* comparando a distribuição de eficiência por técnica.

Ao analisar a Figura 2, pode-se notar que a mediana do Grupo Diretrizes de Usabilidade está um pouco mais alta que a mediana do Grupo MIT 3. Ao comparar as duas amostras usando o teste Mann-Whitney, não foi encontrada diferença estatística significativa entre os dois grupos ( $p = 0,503$ ). Esses resultados sugerem que a MIT 3 e as Diretrizes de Usabilidade proveram eficiência similar quando utilizados para inspecionar o Diagrama de Atividades.

<sup>1</sup> Maiores informações sobre os diagramas utilizados não podem ser divulgadas por questões de confidencialidade.



Nº	Conhecimento em IHC	Conhecimento em Avaliações	Conhecimento em Design	Número de Discrepâncias	Total de Falso-positivo	Total de Defeitos	Tempo (h)	Defeitos/Hora	Total de Defeitos
01	Baixo	Baixo	Médio	22	1	21	0,83	25,20	Diretrizes de Usabilidade = 215
02	Baixo	Baixo	Médio	17	2	15	0,60	25,00	
03	Baixo	Baixo	Baixo	23	1	22	1,32	16,71	
04	Baixo	Baixo	Baixo	22	0	22	0,85	25,88	
05	Médio	Baixo	Baixo	21	4	17	1,00	17,00	
06	Baixo	Baixo	Médio	20	0	20	1,10	18,18	
07	Baixo	Baixo	Baixo	20	0	20	1,25	16,00	
08	Médio	Baixo	Baixo	16	5	11	0,58	18,86	
09	Alto	Alto	Médio	17	0	17	1,27	13,42	
10	Médio	Baixo	Baixo	25	2	23	1,12	20,60	
11	Baixo	Baixo	Baixo	12	0	12	0,87	13,85	
12	Médio	Médio	Baixo	16	1	15	1,05	14,29	
13	Médio	Médio	Médio	16	4	12	0,88	13,58	
14	Médio	Médio	Baixo	13	5	8	1,10	7,27	
15	Baixo	Baixo	Baixo	13	1	12	0,93	12,86	
16	Baixo	Baixo	Médio	19	5	14	0,85	16,47	
17	Baixo	Baixo	Baixo	25	3	22	0,75	29,33	
18	Baixo	Baixo	Baixo	22	6	16	0,40	40,00	
19	Baixo	Baixo	Baixo	22	4	18	1,35	13,33	
20	Médio	Baixo	Médio	14	2	12	1,12	10,75	
21	Baixo	Baixo	Baixo	13	3	10	0,63	15,79	
22	Médio	Baixo	Baixo	16	4	12	0,58	20,57	
23	Médio	Baixo	Baixo	13	0	13	0,38	33,91	
24	Baixo	Baixo	Médio	18	4	14	0,83	16,80	
25	Baixo	Médio	Baixo	15	2	13	0,67	19,50	

Tabela 3. Resultados por Inspetor – Diretrizes de Usabilidade X MIT 3

Técnica	Total de Defeitos	Média de Defeitos	Eficácia Média (%)	Tempo Total (h)	Eficiência (Def./Hora)
Diretrizes de Usabilidade	215	17,92	46	11,83	18,17
MIT 3	176	13,54	35	10,48	16,79

Tabela 4: Resultados dos indicadores de Eficiência e Eficácia – Diretrizes de Usabilidade X MIT 3

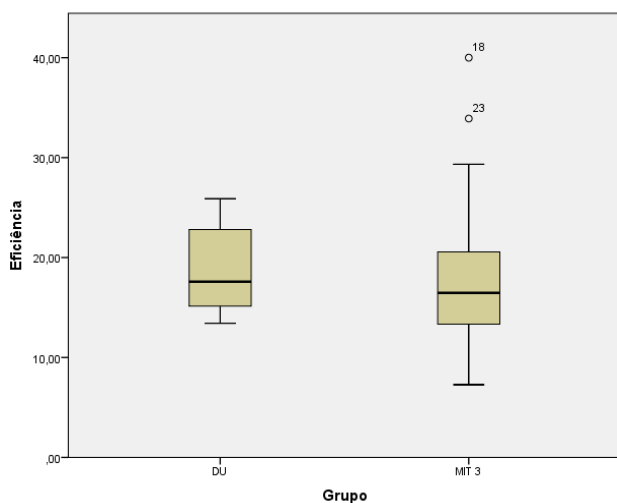


Figura 2 - Boxplots de eficiência por técnica

A mesma análise foi aplicada para verificar se houve diferença significativa em relação ao indicador eficácia das duas técnicas na detecção de defeitos de usabilidade. Os *boxplots* apresentados na Figura 3 mostram que o grupo de inspetores que utilizou as Diretrizes de Usabilidade obteve um desempenho significativamente superior ao do grupo que utilizou a MIT 3. Isto foi confirmado pelo teste Mann-Whitney ( $p = 0,016$ ). Estes resultados apontam que as Diretrizes de Usabilidade foram mais eficazes do que a MIT 3 para inspecionar o Diagrama de Atividades.

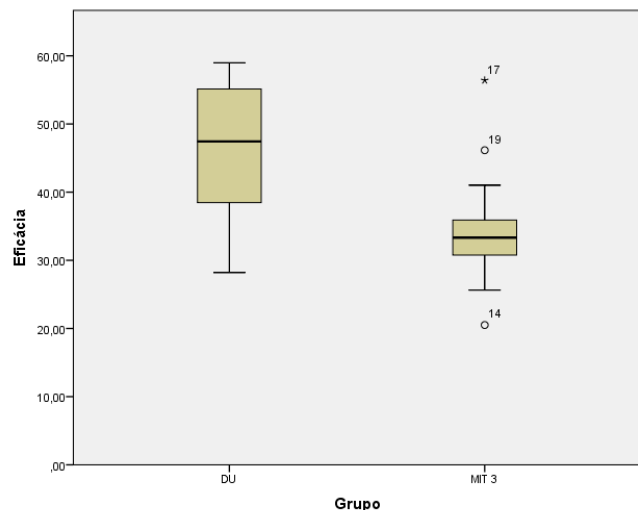


Figura 3 - Boxplots de eficácia por técnica

Neste primeiro estudo experimental, os resultados quantitativos obtidos mostram que as Diretrizes de Usabilidade apresentaram uma eficiência de 18,17 defeitos por hora, e eficácia de 46%, já a MIT 3 mostrou uma eficiência de 16,79 defeitos por hora e eficácia de 35%. De acordo com esses resultados, é possível ter indícios de que o uso das Diretrizes de Usabilidade como técnica de inspeção de usabilidade em Diagramas de Atividades é viável para auxílio na detecção de defeitos nesta linha.

No entanto, devido à pequena amostra, não é possível considerar estes resultados conclusivos, sendo necessário repetir este estudo com uma amostra maior e mais heterogênea de participantes.

### Resultados Qualitativos

Para Mallardo e Calefato [12], investigar a aceitação dos usuários para uma tecnologia requer um modelo que explique as atitudes e comportamentos das pessoas. Neste estudo utilizou-se o modelo de aceitação de tecnologia (*Technology Acceptance Model – TAM*) proposto por Davis [8] para investigar a percepção do usuário sobre as Diretrizes de Usabilidade e MIT 3. A estrutura do modelo TAM tem como base dois fatores: (1) Percepção sobre utilidade, definida como “o grau no qual uma pessoa acredita que utilizar uma tecnologia específica melhoraria seu desempenho no trabalho” e (2) Percepção sobre facilidade de uso, definida como “o grau no qual uma pessoa acredita que utilizar uma tecnologia específica seria livre de esforço” [10].

Foram utilizados, na análise qualitativa, os questionários preenchidos após a inspeção, através dos quais os participantes relataram suas ações e impressões em relação à utilização de cada técnica. Foi utilizada uma escala de seis pontos, tendo como base os questionários aplicados por Mallardo e Calefato [12], que são: concordo totalmente, concordo amplamente, concordo parcialmente, discordo parcialmente, discordo amplamente e discordo totalmente. Neste questionário, os inspetores respondiam qual o seu grau de concordância em relação à utilidade e facilidade de uso, além do seu grau de concordância em relação à facilidade de compreender.

#### *Percepção sobre Facilidade de Uso*

A Figura 4 exibe as respostas referentes à Percepção sobre facilidade de uso das Diretrizes de Usabilidade e MIT 3. Pode-se destacar que nas Diretrizes de Usabilidade, três dos doze inspetores discordaram da questão “foi fácil aprender

a utilizar a técnica”. Dois dos doze discordaram das questões “consegui utilizar a técnica da forma que eu queria”, “foi fácil ganhar habilidade no uso da técnica” e “é fácil lembrar como utilizar a técnica para realizar uma inspeção de usabilidade”. Houve ainda, um participante que discordou totalmente da questão “considero a técnica fácil de lembrar”. Já com relação à MIT 3, pode-se destacar que em cada uma das questões houve apenas um inspetor – dos treze – que discordou.

#### *Percepção sobre a Utilidade da Técnica*

As respostas relacionadas à Percepção sobre utilidade das técnicas são exibidas na Figura 5. As questões relatam sobre a percepção do inspetor se a técnica permite detectar defeitos mais rápido, se usá-la ajuda a melhorar o desempenho na inspeção, se facilita a inspeção e se o inspetor a considera útil para inspeções de usabilidade. Observa-se que, no que diz respeito às Diretrizes de Usabilidade, dos doze participantes, um discordou totalmente na questão “a técnica permitiu detectar defeitos mais rápidos”, um na questão “usar a técnica melhorou o desempenho na inspeção” e também um na questão “usar a técnica facilitou a inspeção”, o que indica que em algum momento o usuário encontrou dificuldade de utilização. Além disso, outros dois inspetores discordaram parcialmente das questões “usar a técnica melhorou o desempenho na inspeção” e “considero a técnica útil para inspeção de usabilidade”. Ainda sobre esta percepção, nenhum inspetor discordou com relação a MIT 3, indicando que houve facilidade na utilização dessa técnica.

#### *Percepção sobre Facilidade de Compreender*

Através do questionário pós-inspeção foi verificado um indicador adicional que não está relacionado ao modelo TAM, chamado de Facilidade de Compreensão. As respostas relacionadas à Facilidade de Compreensão das técnicas são exibidas na Figura 6. Observa-se que apenas um usuário discordou parcialmente na seguinte questão referente às Diretrizes de Usabilidade: “as questões são fáceis de compreender”, o que indica que em algum momento o usuário encontrou dificuldade de compreender. Nota-se que não houve nenhum inspetor que discordasse da questão referente à MIT 3, mostrando que no geral a técnica é fácil de compreender.

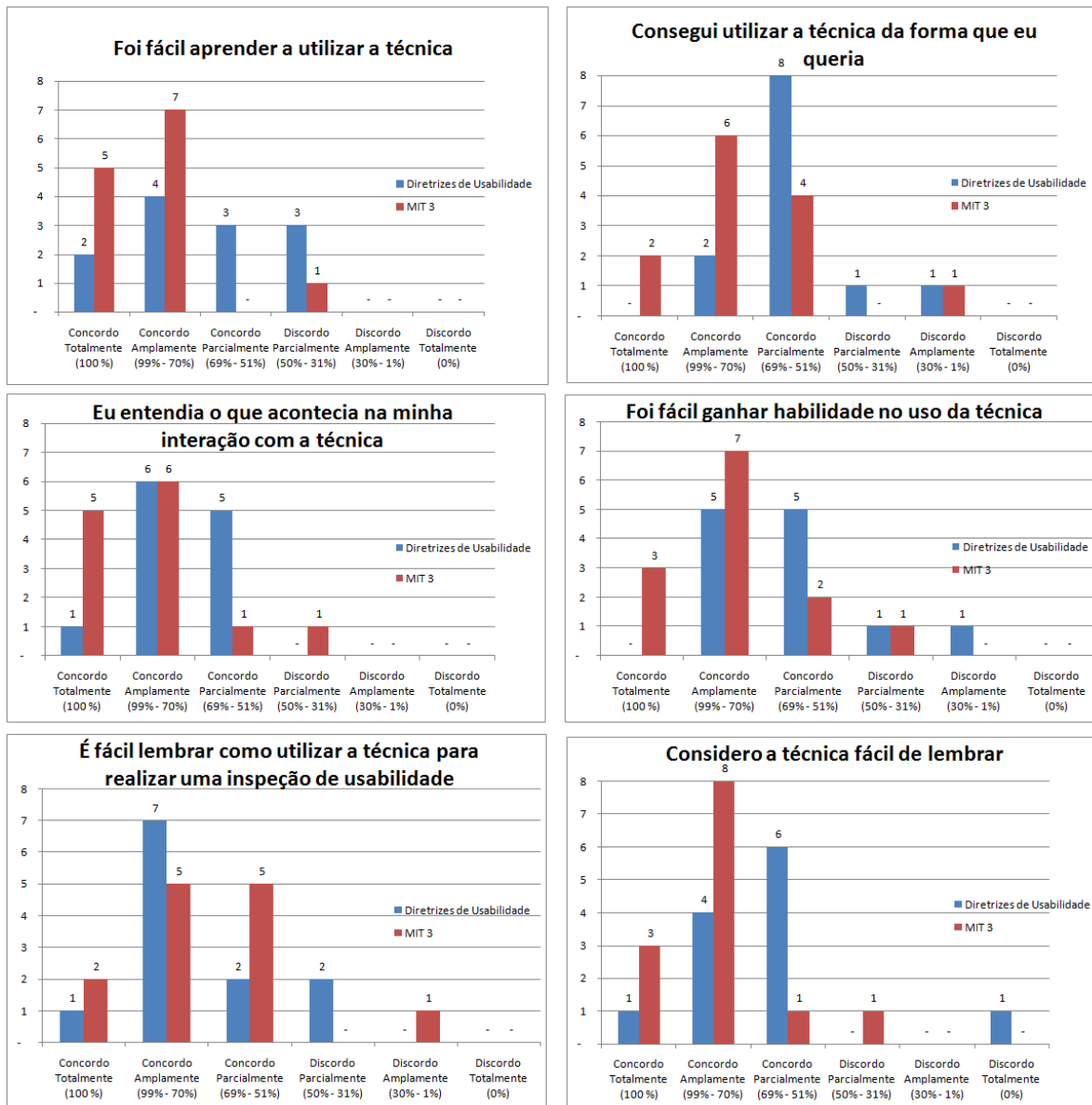


Figura 4 - Resultado das questões relacionadas à Percepção sobre a Facilidade de Uso das técnicas

### AMEAÇAS À VALIDADE DO ESTUDO

Ameaças que podem afetar a validade dos resultados existem em todos os estudos experimentais. As ameaças relacionadas a este estudo são apresentadas a seguir, classificadas em quatro categorias: validade interna, validade externa, validade de conclusão e validade de constructo [24].

#### Validade Interna

Neste estudo, consideraram-se três principais ameaças que representavam um risco de interpretação imprópria dos resultados: (1) efeitos de treinamento, (2) classificação de experiência e (3) medição de tempo. Em relação à primeira ameaça, poderia haver um efeito causado pelo treinamento, caso o treinamento das Diretrizes de Usabilidade tivesse

qualidade inferior ao treinamento da MIT 3 e vice-versa. Este risco foi controlado preparando o mesmo treinamento sobre usabilidade para os dois grupos. Além disso, para cada grupo foi realizada uma apresentação específica sobre a técnica que seria utilizada contendo apenas instruções e exemplos das questões das Diretrizes de Usabilidade e itens de verificação da MIT 3. Em relação à classificação do conhecimento dos participantes, ela foi uma auto classificação, com base no tipo de conhecimentos anteriores (em IHC, avaliações e design). Finalmente, sobre a medição do tempo, o moderador conferia o tempo anotado por cada participante a cada entrega da planilha com a descrição das discrepâncias.

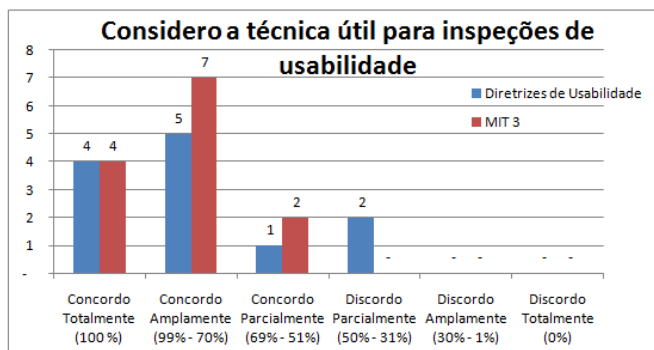
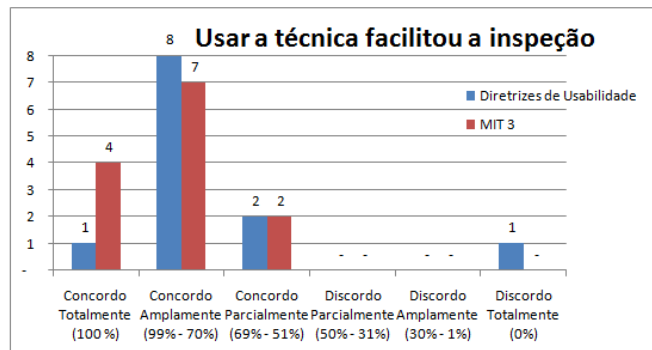
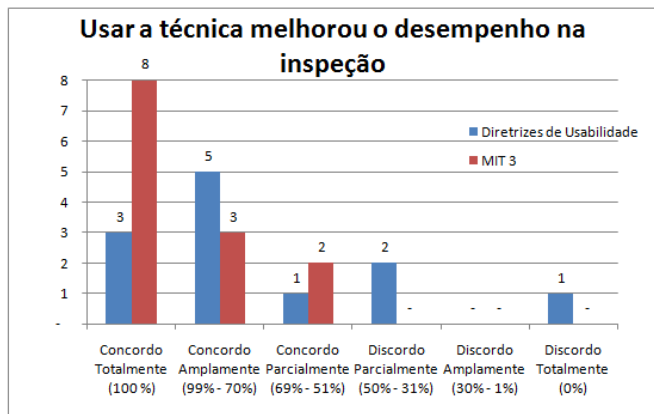
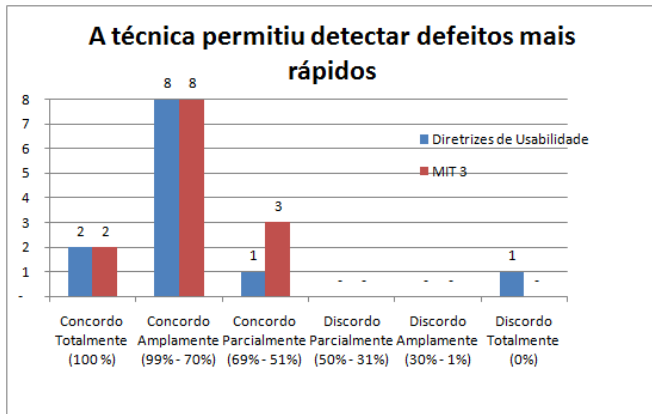


Figura 5 - Resultado das questões relacionadas à Percepção sobre a Utilidade das Técnicas

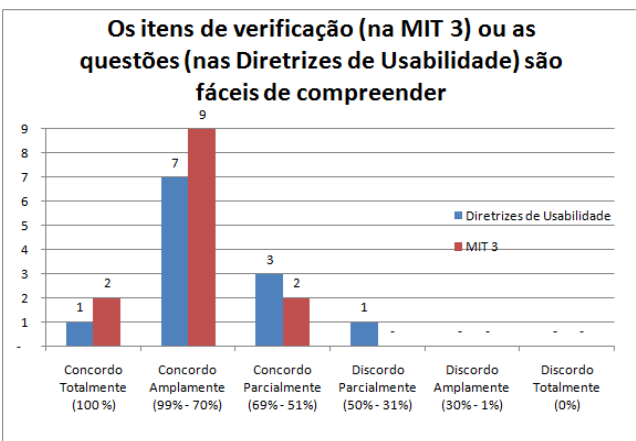


Figura 6 - Resultado das questões relacionadas à Facilidade de Compreensão das Técnicas

#### Validade Externa

Quatro questões foram consideradas: (1) os participantes do estudo foram estudantes de graduação; (2) o estudo foi realizado em ambiente acadêmico; e (3) a validade do modelo como representante. Segundo Carver *et al.* [5], estudantes que não possuem experiência em aplicações na indústria podem apresentar habilidades similares a inspetores menos experientes, sendo este o perfil desejado para os participantes. Em relação à questão (2) o objeto da inspeção (Diagrama de Atividades) é um modelo que faz parte da especificação de um sistema real. Sobre a questão

(3), não é possível afirmar que o modelo utilizado na inspeção represente todo tipo de Diagrama de atividades.

#### Validade de Conclusão

Neste estudo, o maior problema é o tamanho e a homogeneidade da amostra. Pois a quantidade de participantes não é a ideal do ponto de vista estatístico e por serem todos alunos de graduação de uma mesma instituição. O tamanho reduzido de amostras é um problema conhecido em estudos em IHC e ES [4, 6]. Devido a estes fatos, há limitação nos resultados, sendo estes considerados indícios e não conclusivos.

#### Validade de Constructo

Neste tipo de ameaça, considerou-se a definição dos indicadores eficiência e eficácia. Esses indicadores são comumente adotados em estudos que investigam técnicas de detecção de defeitos e estes indicadores foram medidos utilizando a mesma abordagem aplicada em [4, 7].

#### CONCLUSÕES E TRABALHOS FUTUROS

Este artigo apresentou um estudo de viabilidade que teve como objetivo comparar duas técnicas de inspeção de usabilidade, MIT 3 [22] e Diretrizes de Usabilidade [18], em termos de eficiência e eficácia. Através da análise dos resultados quantitativos do experimento, observou-se que as Diretrizes de Usabilidade apresentaram uma eficiência de 18,17 defeitos por hora enquanto a MIT 3 apresentou uma eficiência de 16,79 defeitos por hora, entretanto, sem diferença estatística significativa. As Diretrizes de Usabilidade apresentaram uma eficácia de 46% e a MIT 3



uma eficácia de 35%, notando-se assim, que as Diretrizes de Usabilidade obtiveram um desempenho significativamente superior, confirmado pelo teste Mann-Whitney, que o do grupo que utilizou a MIT 3. No entanto, conforme já mencionado, devido à pequena amostra e às ameaças à validade também já apresentadas, não é possível considerar estes resultados conclusivos, sendo necessário repetir este estudo com uma amostra maior e mais heterogênea.

No que diz respeito à análise dos resultados qualitativos, pode-se observar que, no geral, uma quantidade menor de inspetores discordaram das questões de Percepção sobre a Facilidade de Uso, Utilidade da Técnica e Facilidade de Compreender com respeito a MIT 3, mostrando indícios de facilidade ao utilizar a MIT 3. Com relação às Diretrizes de Usabilidade, vários inspetores discordaram das mesmas questões de Percepção, mostrando indícios de dificuldades ao utilizar as Diretrizes de Usabilidade.

Isto nos leva a seguinte questão: como que as Diretrizes de Usabilidade foram mais eficientes e eficazes e ao mesmo tempo foram consideradas mais difíceis de serem utilizadas e percebidas como menos úteis? Isto pode estar relacionado ao fato delas exigirem um pouco mais de experiência por parte do avaliador, o que também deve ser analisado pois foram consideradas fáceis de compreender assim como a MIT 3. O fato da MIT 3 ter maior aceitação dos inspetores do experimento por apresentar itens de verificação mais fáceis de aprender e lembrar pode reforçar a ideia que esta técnica é mais indicada para inspetores com pouco conhecimento em inspeção de usabilidade. Desta maneira, presume-se, que as Diretrizes de Usabilidade podem ser utilizadas por equipes que tenham profissionais com algum grau de conhecimento em IHC e a MIT 3 por inspetores com baixo grau de conhecimento, já que são fáceis de aprender e lembrar.

Este trabalho traz à tona uma discussão sobre a importância de se ter métodos que antecipem a avaliação de usabilidade e, neste caso, métodos que possam ser adotados e utilizados na indústria de forma “fácil” e complementar à outros métodos, contribuindo também para a formação e capacitação das equipes de projeto e desenvolvimento. Adicionalmente, este trabalho contribui para a aproximação das áreas de IHC e ES, mostrando a aplicabilidade de métodos de avaliação de IHC em artefatos oriundos da ES.

As questões levantadas confirmam a necessidade, já identificada com a análise dos dados quantitativos, de um estudo com uma amostra mais heterogênea. Estudo este que é parte dos trabalhos futuros desta pesquisa. Cabe ressaltar ainda que, a fim de aprofundar a análise qualitativa do estudo, está sendo realizada uma análise específica dos dados qualitativos que foram coletados por meio dos questionários pós-inspeção, através do método *Grounded Theory* [19].

## AGRADECIMENTOS

Os autores agradecem a CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), a SUFRAMA e a FAPEAM pelo apoio financeiro para este trabalho, aos participantes do estudo de viabilidade e aos pesquisadores do USES-UFAM pelas contribuições na execução desta pesquisa.

## REFERÊNCIAS

1. Atterer, R. Where Web Engineering Tool Support Ends: Building Usable Websites. In Proceedings of the 20th Annual ACM Symposium on Applied Computing, Santa Fe, New Mexico (2005), 12–17.
2. Bias, R., Mayhew, D.: Cost-Justifying Usability: An Update for the Internet Age. Addison Wesley (2005).
3. Booch, G., Rumbaugh, J., Jacobson, I.: The Unified Modeling Language. Addison Wesley (1998)
4. Bonifácio, B., Fernandes, P., Oliveira, H. A. B. F., Conte, T. UBICUA: A customizable usability inspection approach for web mobile applications. In Proceedings of the IADIS International Conference on WWW/Internet (2011).
5. Carver, J., Jaccheri, L., Morasca, S., Shull, F. Issues in Using Students in Empirical Studies in Software Engineering Education. In Proceedings of the 9th International Symposium on Software Metrics (METRICS'03), Sydney, Australia (2003), 239 – 249.
6. Conte, T., Massolar, J., Mendes, E., et al. Usability Evaluation based on Web Design Perspectives. In Proceedings of the First International Symposium on Empirical Software Engineering and Measurement (2007), 146 – 155.
7. Conte, T., Massolar, J., Mendes, E., et al. Web Usability Inspection Technique Based on Design Perspectives. In IET Software Journal, 3 (2009), 106-123.
8. Davis, F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. In MIS Quarterly, 13(1989), n. 3, 319-339.
9. Hix, D., Hartson, H.: Developing User Interface: Ensuring Usability Through Product & Process. John Wiley & Sons (1993).
10. Laitenberger, O., Dreyer, H. M. Evaluating the usefulness and the ease of use of a web-based inspection data collection tool. In Proceedings of the 5th International Symposium on Software Metrics (1998), page 122.
11. Lanubile, F., Mallardo, T., Calefato, F. Tool support for Geographically Dispersed Inspection Teams. In Software Process Improvement and Practice, 8(2003), 217-231.
12. Mallardo, T., Calefato, F. Tool support for Geographically Dispersed Inspection Teams. In Software Process Improvement and Practice, 8(2003), 217-231.

13. Mann, H. B., Whitney, D. R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. In *Annals of Mathematical Statistics* 18 (1947), 50 - 60.
14. Nielsen, J., Mack, R. L.: *Usability Inspection Methods*. John Wiley & Sons (1994).
15. Prates, R. O., Barbosa, S. D. J. Avaliação de Interfaces de Usuário – Conceitos e Métodos. Juan Manuel Adán Coelho; Sandra C. P. Ferraz Fabbri. (Org.). *Jornada de Atualização em Informática do Congresso da Sociedade Brasileira de Computação*. Campinas: SBC, 2 (2003), 245-293.
16. Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., Carey, T.: *Human-Computer Interaction*. Addison Wesley (1994).
17. Sauer, C., Jeffery, D.R., Land, L., Yetton, P. The Effectiveness of Software Development Technical Review: A Behaviorally Motivated Program of Research. In *IEEE TSE*, 26 (1) (2000), 1-14.
18. Silva, T. S., Silveira, M. S. Validação de um método para identificação de problemas de usabilidade a partir de diagramas UML. In *Proceedings of the IX Symposium on Human Factors in Computing Systems* (2010), 179 – 188.
19. Strauss, A., Corbin, J., *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. 2 ed. SAGE Publications (1998), London.
20. Travassos, G. H., Shull, F., Fredericks, M., Basili, V. Detecting defects in object-oriented designs: using reading techniques to increase software quality. In *Proceedings of the 14th ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications*, 34 (1999), 47-56.
21. Valentim, N. M. C., Conte, T. Mapeamento Sistemático para a geração da Técnica de Inspeção de Modelos para garantir a Usabilidade. In: *USES Technical Report Número RT-USES-2011-0006* (2011), Manaus.
22. Valentim, N. M. C., Oliveira, K. M., Conte, T. Definindo uma Abordagem para Inspeção de Usabilidade em Modelos de Projeto por meio de Experimentação. In: *XI Simpósio Brasileiro de Fatores Humanos em Sistemas Computacionais (IHC 2012)*, Cuiabá – MT (2012), p. 165 - 17.
23. Valentim, N. M. C. MIT - Um conjunto de técnicas de leitura para inspeção de usabilidade em modelos de projeto. *Dissertação de Mestrado, Programa de Pós-Graduação em Informática, UFAM, Manaus* (2013).
24. Wöhlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., Wessl, A. *Experimentation in software engineering: an introduction*. (2000) Kluwer Academic Publishers.